More than Health: the Role and Value of Meta-Health Effects in Health Care Decisions

Richard De Abreu Lourenço

Doctor of Philosophy

Centre for Health Economics Research and Evaluation

University of Technology Sydney

February 2017

Certification of Original Authorship

I certify that the work in this thesis has not been submitted previously for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signed: _____

R De Abreu Lourenço

Date: 15.02.2017

Acknowledgements

Over the last three years, many have helped me to complete this thesis. To all of you, thank you.

Thank you to my supervisory panel, Professors Rosalie Viney, Marion Haas and Jane Hall for their unending counsel throughout this PhD, and indeed throughout my career as a health economist. Their guidance on this research, keeping my ideas within the realms of possibility and continuous support, particularly through some challenging times in my families' life, will always be appreciated. As will their advice on how to see the bigger picture, while not losing site of the details. I am eternally grateful.

I'd like to acknowledge the support of the Breast Cancer Network of Australia (BCNA) for the mastectomy research in this thesis. Thank you to Lisa Morstyn of the BCNA who helped recruit participants for the mastectomy focus groups. Facilitation of those groups would not have been possible without the support of Rosemary Hannan. To the women with experience of breast cancer who participated in those focus groups, thank you. Particular thanks to Kim Parish and Domini Stuart, patient representatives from the BCNA, for helping me place my mastectomy research in the context of the consumer. Thanks also to Dr Nicholas Zdenkowski for his clinical advice.

A special thanks to Patsy Kenny for her assistance in the design, facilitation and analysis of the findings from the mastectomy focus groups; and for listening to my musings on discrete choice analysis. Similar thanks to Denzil Fiebig, Richard Norman and Chunzhou Mu for their insights into the world of discrete choice analysis. Thank you to Liz Chinchen and Jenny Edwards for proof-reading a draft of this document, and to Liz for her expertise in searching the published literature and alerting me to new publications of relevance.

Of course to my friends who have offered their advice, support and special insights on life during a PhD, thank you – in particular, Kirsten Howard and Marian Shanahan for their advice on discrete choice, managing PhDs, time and life! Special thanks to Phillip Browning, friend and copywriter, for distilling my ramblings on the concept of interest into the term 'meta-health effect' and for proof-reading an earlier version of this thesis.

Most importantly to my partner, my best friend, my in-built supervisor, Dr Jason Abbott, for all of his unwavering support and belief in me over the last three years. Even when I lost both brothers in consecutive years, his strength and support kept me going. Without him by my side this work might never have happened. Thank you does not seem adequate. And of course, thank you to our ever faithful and mischievous fur-babies, Jack and Belle. For keeping me grounded and distracting me when I needed the distraction.

Contents **Page Number**

Table Headings **Page Number**

Figure Heading **Page Number**

Appendix List

Abstract

One of the most visible functions of government is to make decisions about funding health care treatments. This thesis investigates the role and value of meta-health effects in such decisions. Meta-health effects are effects other than health that result from the consumption of health care, and have value in their own right regardless of health status.

The research in this thesis is facilitated via four inter-related case studies. The first examines the available information on decisions made by the Pharmaceutical Benefits Advisory Committee (PBAC) in Australia for evidence of the use of meta-health effects in drug reimbursement decisions. This is supplemented in that same case study by a systematic review of the methods used to value meta-health effects for use in economic evaluations.

Three empirical case studies are subsequently presented which focus on the role of meta-health effects in individuals' decisions regarding health care as a means of informing what might be considered in public decision making. All three case studies use survey-based methods: a general community survey on experiences and attitudes on general practitioner use, and two discrete choice experiment (DCE) surveys (one on ongoing therapy for rheumatoid arthritis, the other for the management of breast cancer recurrence risk). Together these three case studies explore how differences in the decision-making context, and methods of elicitation (such as attitudes or preferences) influence the role and value of meta-health effects. Within the DCEs those values are explored using willingness to pay, investigating how they are affected by framing.

The results show that meta-health effects do influence choice. The review of PBAC decisions and the systematic review show that gains in convenience (e.g. gains in mode of administration) are investigated most often, but that differences in study methods

influence the values derived. An important finding of the results of the empirical case studies is that meta-health effects do influence individual choices and the extent of that influence declines the greater the health implications of that decision. Similarly, they find that the amount and type of information presented influences the values derived in studies eliciting values for meta-health effects. This is not only a contribution to the literature, but highlights the importance to government decision makers of understanding how values for meta-health effects have been derived; careful attention needs to be paid to the manner in which such values have been derived lest they misrepresent the resulting value to society.

Abbreviations List

Abbreviation	Description
ABCR	Intramuscular interferon beta-1a, subcutaneous interferon beta-1a, interferon beta-1b and glatiramer acetate
ADHD	Attention deficit hyperactivity disorder
AE	Adverse event
AIC	Akaike information criterion
AIDS	Acquired immune deficiency syndrome
AIHW	Australian Institute of Health and Welfare
AMD	Age related macular degeneration
ART	Assisted reproductive technology
ASC	Alternative specific constant
ATSI	Aboriginal & Torres Strait Islander
AUD	Australian dollar
BC	Breast cancer
BCNA	Breast Cancer Network Australia
BD	Twice daily
bDMARD	Biological disease modifying anti-rheumatic drug
BIC	Bayesian information criterion
BRCA	Breast cancer
BSA	Body surface area
BSC	Best supportive care
BWS	Best-worst scaling
CA	Conjoint analysis
CBA	Cost-benefit analysis
CBC	Contra-lateral breast cancer
CEA	Cost-effectiveness analysis
CF	Cystic fibrosis
CHERE	Centre for Health Economics Research and Evaluation
CI	Confidence interval
CKD	Chronic kidney disease
CMV	Cytomegalovirus
COPD	Chronic obstructive pulmonary disease
CPM	Contralateral prophylactic mastectomy
CRC	Colorectal cancer
CUA	Cost-utility analysis
CV	Contingent valuation
CVD	Cardio-vascular disease
DCE	Discrete choice experiment
d.f.	Degrees of freedom
EDSS	Expanded disability status scale
EGFR	Epidermal growth factor receptor
EQ-5D	European Quality of Life (EUROQoL) 5 Dimensions
5-FU	5-fluorouracil
FDC	Fixed dose combination
FF	Fluticasone

Abbreviation	Description
FRF	French francs
FS	Fibroscan
Ft	Fortnightly
GBP	Great Britain pounds
GCSF	Growth colony stimulating factor
GMNL	Generalised multinomial logit regression
Govt	Government
GP	General practitioner
HAART	Highly active antiretroviral therapy
HbA1C	Glycosylated haemaglobin
HCV	Hepatitis C virus
HITAP	Health Intervention and Technology Assessment Program
HIV	Human immunodeficiency virus
HPV	Human papillomavirus
HREC	Human Research Ethics Committee
ICER	Incremental cost-effectiveness ratio
ICS	Inhaled cortico-steroid
ICU	Intensive care unit
IIA	Independence of irrelevant alternatives
iid	Independent and identically distributed
Inc	Income
IUD	Intra-uterine device
IV	Intravenous
IVF	In-vitro fertilisation
Kras	Kirsten rat sarcoma
LABA	Long-acting beta agonist
LB	Liver biopsy
LLH	Log-likelihood
LR	Likelihood ratio
MAUI	Multi-attribute utility instrument
MBS	Medicare Benefits Schedule
MDS	Myelodysplastic syndrome
MESH	Medical subject headings
MHE	Meta-health effects
MM	Multiple myeloma
MNL	Multinomial logit
MNP	Multinomial probit
MOGA	Medical Oncology Group of Australia
MRI	Magnetic resonance imaging
MRS	Marginal rate of substitution
MS	Multiple sclerosis
mth	Month(ly)
mWTP	Marginal willingness to pay
n.a.	Not applicable
NHMRC	National Health & Medical Research Council
NPR	Nepalese rupees
NSCLC	Non-small cell lung cancer

Abbreviation	Description
OLS	Ordinary least squares
OMEP	Orthogonal main effects plan
ONJ	Osteonecrosis of the jaw
OOP	Out-of-pocket
PA	Pseudomonas aeruginosa
PBAC	Pharmaceutical Benefits Advisory Committee
PBS	Pharmaceutical Benefits Scheme
PC	Primary care
PhD	Doctor of Philosophy
PSA	Prostate specific antigen
PSD	Public summary document
QALY	Quality adjusted life year
QoL	Quality of life
RA	Rheumatoid arthritis
rnk	Rank
SAL	Salbuterol
s.d.	Standard deviation
s.e.	Standard error
SF-36	Short Form 36
SF-6D	Short Form 6 Dimensions
SG	Standard gamble
SRE	Skeletal related event
t.i.d	Three times daily
T2DM	Type 2 diabetes mellitus
TM	Therapeutic mastectomy
TTO	Time-trade-off
UK	United Kingdom
Unk	Unknown
USA	United States of America
USD	United States dollar
VAS	Visual analogue scale
VI	Vilanterol
VIF	Variance inflation factor
Wgt	Weight
Wk	Week(ly)
WTP	Willingness to pay
Yrs	Years

1 Introducing an Investigation of Meta-Health Effects

Chapter Summary

This thesis investigates the role and value of meta-health effects in health care decision-making. These are effects other than health that result from the consumption of health care, and have value in their own right, regardless of health status. The focus of the research is on direct meta-health effects, such as convenience and autonomy.

If we accept that health care has effects other than health, understanding the value of meta-health effects is important if health care decision-making is to consider the full suite of costs and outcomes associated with alternative interventions. To date, cost-utility analysis (CUA) has been the main form of economic evaluation used to inform decision-making to allocate public sector resources to health care. But the metric used in CUA, quality adjusted life years (QALYs), focuses on health maximisation, largely excluding meta-health effects. Cost-benefit analysis (CBA) applies monetary valuations to the outcomes of interest and does not rely on health maximisation. Typically monetary valuations for outcomes such as meta-health effects are derived using willingness to pay (WTP), estimated using stated preference methods such as discrete choice experiments (DCEs).

Framing effects are likely to influence the values obtained for meta-health effects within preference elicitation studies. Whether meta-health effects are described as gains or losses, how they are described in the overall context of a condition and in comparisons between alternatives, all influence the importance they are given and the values attributed to them. Understanding the magnitude of these framing effects is therefore important in assessing the value of meta-health effects.

The question regarding the role and value of meta-health effects in health care decision making is explored using four related case studies: a review of the existing evidence on the use of meta-health effects, including by a public sector decision-maker; the role of choice and convenience in

influencing patient-general practitioner relationships; the role and value of convenience in affecting treatment choices for rheumatoid arthritis; and the role and value of convenience and reassurance in decisions regarding the management of ongoing breast cancer recurrence risk.

1.1 Introduction

Decisions about health care are made constantly and at many levels; by the woman deciding whether she is prepared to undergo fertility enhancement in her attempt to become pregnant, or the doctor deciding whether to remove an infected gall bladder via laparoscopic or open surgery, or a government deciding whether to include the latest monoclonal antibody for colorectal cancer on its national drug subsidy list. Whether explicit or not, all these decisions compare what goes into the production of health care (its costs) with its expected outcomes.

The comparison of costs and outcomes is particularly important in the context of allocating public sector resources to health care. Ideally, in that context the decision-maker should ensure that all relevant costs and outcomes are considered. Commonly, outcomes in this case are those that are specific to a change in health[i]. However, health care has effects that extend beyond its impact on health. These include effects such as providing individuals with reassurance or ensuring autonomy, the (in)convenience associated with health care, or increased knowledge (information). In this thesis, these effects are referred to as meta-health effects; effects other than health resulting from the consumption of health care, and that have value in their own right, regardless of health status.

This thesis is concerned with the role and value of meta-health effects in health care decision-making. Specifically, it investigates whether such effects are being considered in current reimbursement decisions, and the role and value placed on meta-health

[i] Throughout this thesis, the term health is used interchangeably with health status, where the latter refers to more than the presence or absence of disease, but is a more holistic characterisation of health that reflects mental well-being, functioning and physical illness.[1]

effects by individuals in a number of different health care contexts. Given that values for meta-health effects might exist, this thesis also investigates how the manner in which they are measured influences the resulting values for use in informing decision-makers allocating public sector resources to health care.

This chapter proceeds by presenting a definition of meta-health effects and a rationale for why they are important. Subsequently, the use of meta-health effects in health care decision-making is explored, including an overview of current methods of preference and value elicitation. The research questions, methods to be applied and thesis structure are then described.

1.2 Meta-Health Effects

From a health production viewpoint, the outputs of health care are the expected health outcomes. However, there are other outcomes from health care. Mooney (1998)[2] uses the term process utility to describe such outcomes, and suggests that they arise from four possible sources: information; autonomy and the freedom to choose; the circumstances around being treated; and from being community minded citizens who desire a decent health service. Other authors, have offered similar definitions of outcomes other than health, including: the transitory, unintended consequences of health care delivery that result from the process of consuming care[3,4]; outcomes pertaining to the process of care[5]; and broader community and social impacts resulting from health care (such as productivity, education, equity, reductions in crime, social participation, environmental effects, and caring).[6-9]

In the past these effects have been collectively termed non-health outcomes[2,5,10]. The term meta-health effect is introduced in this thesis for these effects on the basis that describing effects as "non-health" presupposes that health effects have primacy. As described below, it allows for those effects, such as reassurance or autonomy for example, to have value regardless of whether there is a corresponding change in health status.

There are several aspects to note regarding the definition of meta-health effects. Firstly, these effects can be separated into: (i) those arising from the process of consuming health care; and (ii) outcomes other than health arising due to the consumption of that care. Examples of the former include convenience associated with an improved medication regimen, the trust engendered by improved clinician-patient relationships, or the impact of the location of where care is provided. Examples of the latter include reassurance from being vaccinated against a serious condition, and obtaining knowledge about one's genetic status for a condition.

Secondly, meta-health effects can impact on patients directly, irrespective of the effect of care on health, or indirectly through an intermediate health effect. Direct meta-health effects arise from the consumption of health care, regardless of the impact on health. For example, the act of administering daily injections for patients with insulin dependent Type I diabetes might have different impacts for different patients. For some, daily injections provide reassurance (the meta-health effect) that their underlying condition will remain controlled; for others the discomfort of the needle and negative association of the reminder that they are diabetic represents a disutility. Neither of these impacts is necessarily reliant on the effect those injections has on the patient's blood sugar levels.

Indirect meta-health effects are those that arise from a change in health status, for example, improved productivity, changes in education performance, social participation and crime. Carer and equity effects might be exceptions where in some instances elements of the impact might be direct rather than indirect. For example, carers might derive a direct positive effect from knowing that patients are receiving appropriate care (reassurance, as a form of positive externality), and equity might be affected merely through the act of accessing care regardless of the outcome of that care. However, in both cases there are also likely to be indirect effects (such as a reduction in carer burden associated with improved health status, and improved social and health

equality through improved health status). Due to this potential duality in effect, in this thesis carer and equity effects are classified as indirect meta-health effects.

This thesis focuses on meta-health effects that arise either through the process of care or due to the consumption of care, and it is restricted to the consideration of direct meta-health effects: those which have an impact on the individual without first requiring an effect on their health.

1.2.1 Why do meta-health effects matter?

It is likely that patients value not only being in good health, but also their experience of the care that produces that health.[2,11,12] For some individuals, how health care is delivered might be just as important as what is produced. Not just the consequences of care carry value, but so too does the process of care.[2,11-13] For others, the outcomes of intermediate steps in the production of health – such as information arising from diagnostic tests – contribute directly to their well-being.[2,11-13] Thus, assessing and valuing meta-health effects is important to capture the full spectrum of the impact of health care, but it is complicated due to the lack of instruments available currently to value such effects.[9]

Assessing meta-health effects is also becoming more important as the manner in which care is delivered changes and the outputs of health care become more complex. Within the provision of medical care, there is a growing focus on the manner in which care is provided[14], in particular on providing patient centred care[15-17], and on less tangible outcomes of care such as information.[18-21].

Increasingly, advances in the delivery of pharmaceuticals are made in terms of changes in the route or schedule of their administration. These innovations have been used by manufacturers to claim higher prices for pharmaceutical products. This is of particular relevance to public sector decision-makers such as the Pharmaceutical Benefits Advisory Committee (PBAC) in Australia, whose role is to evaluate the claims being

made by manufacturers to determine whether or not to recommend public subsidy of a pharmaceutical product. Such decisions are informed by economic evaluations that address the question of whether the additional costs of subsidising the new pharmaceutical product justify the additional outcomes. There is the potential for, and increasing evidence that, gains in meta-health effects have been proposed as justification for higher prices within economic evaluations submitted to the PBAC (see Chapter 2).

A pragmatic search of the published guidelines of eight public health care reimbursement agencies around the world showed that four (the Netherlands, the UK, Australia and Canada) refer to the measurement of meta-health effects[ii] as being potentially relevant in demonstrating value, two (New Zealand and Germany) do not mention meta-health effects but also do not restrict benefits to health effects, and two (Taiwan and Thailand) restrict benefits to health effects (see Table 1).[22-30] The guidelines that refer to meta-health effects offer convenience as a potential example, as well as mentioning productivity gains, reassurance, access to care and satisfaction with care. Only Canada's guidelines suggest how such effects might be incorporated into an evaluation to inform subsequent decision-making, explicitly recommending the use of cost-benefit analysis (CBA).[28] The Institute for Quality and Efficiency in Health Care (IQWiG)[23] in Germany do not mention meta-health effects but note that any effect can be included as a measure of benefit, provided that it can be measured on a cardinal scale. The PBAC guidelines[24] refer to the use of meta-health effects as a potential source of value for use in reimbursement submissions, noting that valuations of these effects are not straightforward to conduct and that the results might not be as influential as health outcomes in the decision-making process.

[ii] The term meta-health effect has been defined within this thesis. Accordingly, searches of health technology assessment guidelines focused on terms referring to non-health outcomes, or outcomes other than health.

Table 1: Reimbursement authority guidelines referring to meta-health effects

Agency	Description	Source
PBAC: Australia	"PBAC may also consider non-health outcomes, including aspects of the delivery of a health care intervention beyond the health gain obtained; for example, greater convenience or production gains to society beyond those valued by the population benefiting with improved health. However, the valuation of non-health outcomes is not straightforward and those outcomes might not be as influential in decision-making as health outcomes."	Throughout guidelines, particularly in defining outcomes.[24]
Canadian Agency for Drugs and Technology in Health: Canada	In justifying use of CBA, some meta-health effects such as convenience, reassurance, reductions in anxiety are listed. "It may be appropriate to use a CBA in certain situations, such as when: • a consequence of an intervention is difficult to value using QALYs (e.g., short-term symptom relief, patient reassurance or anxiety from screening) • an attribute of an intervention is difficult to value using any health outcome (e.g., shorter or less frequent treatment, a more convenient dose form) • a process outcome are major factors in analysing an intervention (e.g., access to or satisfaction with care)."	Description of types of economic evaluation.[28]
IQWiG: Germany	Any type of effect can be included in an assessment of benefit provided that is measurable on a cardinal scale. Health effects take precedence over impacts on time, effort and patient satisfaction.	Definition of "benefit" that can be included in assessments of health technology. Page 35 and 78 of the General Methods.[23]
College voor Zorgverzekeringen: Netherlands	In determining therapeutic value between drugs, "ease of use" (which includes mode and frequency of administration) can be used as a differentiating criterion where intended and unintended (health) effects are comparable.	Definition of therapeutic value and sub-criteria in assessment procedures. Page 28.[29]
PHARMAC: New Zealand	Health benefits should be restricted to HRQoL, measured using QALYs. There is no mention of meta-health effects.	Description of estimating health benefits, Chapter 6.[27]
National Health Insurance Administration: Taiwan	Only health effects, including QALYs, should be used as the outcome measure for assessing cost-effectiveness; these are limited to life years saved and QALYs. There is no mention of meta-health effects.	Description of efficacy versus effectiveness.[22]
HITAP: Thailand	Focus on health effects (effectiveness); measured using QALYs (recommend EQ-5D valuation). There is no mention of meta-health effects.	Description of measuring clinical effects and utility.[25,26]
NICE: UK	If characteristics of healthcare technologies have a value to people independent of any direct effect on health, the nature of these characteristics should be clearly explained and, if possible, the value of the additional benefit should be quantified. These characteristics may include convenience and the level of information available for patients.	Definition of perspective, item 5.1.8.[30]

Abbreviations: CBA, cost-benefit analysis; HITAP, Health Intervention and Technology Assessment Program; HRQoL, health related quality of life; IQWiG, Institute for Quality and Efficiency in Health Care; NICE, National Institute of Health and Clinical Excellence; PBAC, Pharmaceutical Benefits Advisory Committee; PHARMAC, Pharmaceutical Management Agency; QALY, quality adjusted life year; UK, United Kingdom.

Given the scope for meta-health effects to represent value to individuals and to be used to justify higher prices for health care products, it is important to understand the circumstances under which meta-health effects might be considered relevant, and the factors influencing the value placed on those effects. If individuals do indeed value meta-health effects, it is desirable that we can measure their value accurately if decision-making in health care, and economic evaluations in particular, are to be capable of considering the full suite of costs and outcomes.[10]

1.3 Methods to Inform Decisions About Resource Allocation

The use of economic evaluations to inform the allocation of public resources in health care has become well established. Economic evaluation is the systematic, comparative analysis of alternative courses of action (in this case, competing allocations of resources in health care) in terms of their costs (what is required for their provision) and their consequences (the positive and negative outcomes of those actions).[31]

The type of economic evaluation used is determined by the expected differences between the alternative programmes under comparison, and the manner in which the outcomes of those programmes are measured. Where two alternative health care interventions produce the same outcomes, both in terms of the magnitude and nature of those outcomes, a cost-minimisation analysis is used in which only differences in their costs are considered.[31] Cost-effectiveness analyses (CEA) are used where one intervention produces more of the outcome of interest than the alternative. In CEA, there is a single outcome of interest per comparison, measured in natural units such as life years gained or the proportion of patients that achieve a tumour response. However, this may not capture all relevant differences between interventions and the outcomes used are not necessarily comparable across indications. In this case it is difficult to determine what constitutes an acceptable incremental cost per additional outcome, or cost-effectiveness threshold.[31]

Cost-utility analyses (CUA) are an extension of CEA in which the outcomes are generally expressed as quality adjusted life years (QALYs).[31] The QALY has become the predominant metric used within CUA to inform public decision-making in health care. QALYs combine the length of life and the quality of life (QoL) in a single metric that is used to assess the impact of a health care intervention or condition. By expressing outcomes in a common metric, the QALY, the use of CUA provides estimates of cost-effectiveness that can be compared across indications. Thus, it is possible to form a threshold for what is considered an acceptable cost-per QALY across indications.[31] Finally, CBA expresses both costs and outcomes in a common metric, money. This allows the comparison of costs and outcomes to be expressed as a net cost-benefit ratio, with those that are positive being cost-effective.[31]

1.3.1 Informing cost-utility analysis

In forming QALYs, quality of life (QoL) is typically measured from a health related perspective[iii] and is assigned a utility or QALY weight between 0 (death; weights less than 0 are possible) and 1 (perfect health). There are two approaches to measuring utility (QALY weights) for the purposes of CUA: generic or custom built.

1.3.1.1 Generic approaches to generate QALY weights

The most commonly used means of assessing QALY weights is via the use of multi-attribute utility instruments (MAUIs). These are survey-based questionnaires that collect information from individuals (patients or members of the general community) on health-related QoL (HRQoL) across a number of domains of interest e.g. role function, physical function, emotional and health functional.[36] The combination of possible responses to the questions within MAUIs describes a series of health states in which individuals can exist. Utility values are subsequently applied to those health

[iii] Instruments have been proposed for the assessment of QoL for the estimation of QALYs that reflect aspects other than health, such as satisfaction, well-being and individuals' capabilities to function.[32-35] Whether these instruments capture QoL as conceptualised in QALYs is matter for further debate outside of this thesis.

states based on scoring algorithms developed using direct valuations of those health states, typically by members of the general community.[36,37] Those scoring algorithms have been constructed using purpose built approaches, or direct elicitation methods, from which a utility value can be calculated directly from the trade-offs embedded in the task. These methods are described in the following section.

The use of MAUIs to capture the impact of meta-health effects is generally not noted in the literature. Given their focus on HRQoL, it is usually assumed that MAUIs will not capture meta-health effects except where there is overlap with those effects and health effects, such as role functioning or usual activities. However, analyses by Kauf et al. (2008)[38] suggest that this might not be the case. In an examination of the individual and treatment specific factors influencing utility values measured using the SF-6D in patients with human immunodeficiency virus (HIV), they observed that the daily pill burden and the influence of the pill regimen on dietary flexibility did affect utility values.[38] Their results suggest that MAUIs might be sensitive to meta-health effects, although there remains the potential for those effects to be swamped by health effects occurring at the same time.[38]

The potential for MAUIs to capture the influence of meta-health effects introduces one avenue by which attempting to include a separate value for those effects into the estimation of a QALY would constitute some degree of double-counting: value is being ascribed from the change in health status as well as the meta-health effect. Another potential source of such double-counting arises if individuals incorporate health effects when evaluating meta-health effects.[39] For example, if an improved mode of treatment administration (say the introduction of a daily tablet to replace a daily subcutaneous injection) results in better treatment compliance – then *ceteris paribus* – there should be an improvement in patients' health status. In most instances, this will result in an improvement in their HRQoL and be associated with a higher utility value.[iv] To

[iv] It is possible that in such a situation the individual might derive utility from two sources – one a direct meta-health effect attributable to the convenience of the oral medication, the other

introduce a separate utility value to account for the utility from the change in administration might therefore constitute double counting. The potential for such double-counting led Brennan and Dixon (2013)[39] to advise a degree of caution on the use of separate utility values within the one analysis for meta-health effects and health effects for estimating QALYs.

1.3.1.2 Custom approaches to generate QALY weights

Purpose built approaches to the assessment of QALY weights utilise survey based direct elicitation tasks in which respondents state a preference between particular health states, conditions or products. The most commonly used direct elicitation methods are time-trade-off (TTO), standard gamble (SG), and discrete choice experiments (DCEs).[36,37,40,41] The premise of the TTO is that individuals are willing to trade between the length of life and quality of life. Typically, respondents are presented with a hypothetical health state scenario and are asked to choose between a period of time in the health state of interest followed by death, and a shorter period of time in full health followed by death. This question is repeated, keeping the length of time in the health state fixed and varying that in full health, until the respondent is indifferent between the two. By imposing certain assumptions, the ratio of the time in full health and the time in the health state of interest is interpreted as the utility attached to that health state for that duration of time.[42]

The SG is also a scenario-based assessment of utility values, and is grounded in von-Neumann and Morgenstern's expected utility theory.[43] Respondents to an SG task are asked to choose between a certain prospect (the health state of interest for a given period of time) and a lottery (where there is a probability of experiencing full health or immediate death).[43] In each choice, the probabilities in the lottery are varied until the respondent is indifferent between the certain prospect and the lottery. By imposing

an indirect effect resulting from the improvement in health status due to improved compliance associated with the convenient mode of administration.

certain assumptions, this point of indifference represents the respondent's utility value for the health state of interest.[43]

The trade-offs in the TTO and SG focus on the duration of survival and the risk of death respectively. This poses particular challenges for the assessment of value for meta-health effects using these methods because individuals might well value an effect, say convenience, but be unwilling to forgo survival in its favour. This suggests that methods of assessing value that allow for a broader set of trade-offs, such as DCEs, are more well placed for use in measuring the value of meta-health effects.

DCEs are gaining increasing popularity for understanding preferences for health and health care interventions.[44,45] DCEs are based on Lancaster's theory of choice in which the utility from a good or service is derived from the utility of its component attributes.[46] The combination of attributes expected to deliver the highest utility is preferred by the individual. Within a DCE survey the good or service of interest is described by its relevant attributes, such as the efficacy associated with an intervention, how it might be administered, who provides care and the cost to the individual. [47-50] Individuals are presented with repeated comparisons of alternative combinations of the attributes, and asked to make choices for each set of alternatives.[47-50] Analysis of those repeated choices reveals how individuals trade-off between the attributes, and can be used to infer the value of each attribute in that decision.[47-50] Inclusion in a DCE of a survival or risk attribute potentially allows the estimation of QALY weights on a 0 to 1 scale for use in a QALY analysis.[v]

1.3.2 Informing cost-benefit analysis

Applying monetary values to outcomes in CBA relies on an assessment of individuals' or societies' willingness to pay (WTP) for the outcome of interest. The monetary values

[v] In the absence of a survival or risk attribute, the coefficients derived from a DCE cannot be directly interpreted as utility values as would be required for a QALY analysis.[4051]

for use in CBA can be derived using a number of approaches, such as revealed preference information and via stated preference WTP information.[31]

Revealed preference information, utilising the analysis of actual market choices can be used to determine (or reveal) individuals' WTP for a good or service. The difficulty is that such information might not exist for the question of interest[52], particularly where they relate to choices within a constrained market (such as operates in a publicly funded health care system) or for new health care innovations that have not been observed in the market previously.

The use of DCEs and contingent valuation (CV) overcomes this issue through the use of purpose built questionnaires to ask respondents how much they would be willing to pay for a given good or service based on a hypothetical scenario.[53,54] Such scenarios capture real market situations, but the corresponding surveys explore choices that might not otherwise be observed currently. As such, it might not be possible to test the external validity of the choices observed: whether or not they reflect real market choices. Validity in such scenarios might therefore be considered in terms of whether the observed factors influences choices in the manner expected (e.g. less intrusive modes of drug administration would be expected to increase the probability of an intervention being chosen), or whether respondents to DCE and CV questionnaires are being consistent in the choices they make.[55]

Within CV, the question of interest might be framed as a WTP, typically where the attribute of interest is a benefit (gain), or a willingness-to-accept, where the attribute of interest is a loss. A range of mechanisms exist for eliciting the valuation amount (referred to as WTP for simplicity) such as open-ended questions (a statement of the WTP), bidding games (in which the respondent is made repeated offers, ending with their WTP)[54], the use of payment cards which show alternative amounts that respondents accept or not, and take-it-or-leave-it offers (binary yes or no offers that might be followed up with an open-ended question regarding WTP).[53] Within DCEs,

the inclusion of a cost or price attribute allows the estimation of WTP.[50,56] The use of WTP values is discussed later in this chapter.

1.4 Valuing Meta-Health Effects for Decision Making

The proposition that meta-health effects influence individuals' health care-decisions, and therefore influence public sector resource allocations, implies that we are seeking to maximise more than health when allocating resources to health care. This proposition concords with original concepts of extra-utility maximisation. For extra-welfarists, utility is derived from more than just health, and it is not just the total utility achieved that matters, but how a utility outcome is achieved matters, potentially allowing for a broad set of maximands.[57-59] However, in the field of health economics, this maximand is restricted to health, even within an extra-welfarist framework, so that outcomes beyond health are not sources of value within an individual's utility function. [57-61]

Incorporating the value of meta-health effects into the outcomes used in decision-making necessitates the use of CUA or CBA, the two approaches that incorporate a valuation of outcomes. An extra-welfarist view, and focus on health as the maximand, has led to the predominance of CUA and its use of QALYs to inform decision-making in the allocation of public sector resources to health care. Only HRQoL is included in the QoL aspect of QALYs. This has led to criticisms that QALYs are too narrow and incapable of capturing the full suite of benefits associated with health care – such as meta-health effects - that arise to patients and society.[2,8,11,13,60] Mooney (1989, 1998)[2,11] highlight that the focus on health as the maximand in QALYs results in the process of care being excluded as a potential source of value from the individual's utility function.

Despite their focus on health as the maximand, QALY weights have sometimes been constructed to incorporate meta-health effects.[62-68] That is, by constructing comparisons between health state scenarios that differed only on the basis of a meta-

health effect such as convenience[62-64] or the manner in which care was provided[65-68], some studies have reported QALY weights for meta-health effects. What is not clear from such studies was how individuals traded between meta-health effects and other potential sources of value within those scenarios.

Unlike QALY based CUA, CBA adopts a welfarist view of the world, allowing for more than health as the maximand.[57-61] As previously noted, one of the principal means of assessing value for use in CBA is the use of WTP. There are numerous examples reviewed in Chapter 2 of the use of WTP exercises to assess the value of meta-health effects.[vi] Despite the potential for WTP to capture a broader set of benefits beyond health outcomes, there are potential limitations to its use. These include the extent to which individuals protest vote (e.g. choose not to express a WTP over situations for which they currently do not pay lest it result in charges being levied), capture by interest groups, difficulty in framing the payment vehicle (e.g. asking individuals to express their WTP in situations funded by public taxation).[31,55,69] In general, WTP measures exhibit scope bias – they lack sensitivity to the magnitude of the benefit being assessed where one good is part of a larger good.[74] This is potentially problematic in the context of using WTP to assess the value of meta-health effects given that they are likely to be assessed as part of a 'larger good', or health state. Therefore, it is possible that WTP assessments of meta-health effects, considered in isolation of the rest of the good or health state to which they contribute, will be overvalued.

The second challenge is one of 'focus illusion'; individuals inflate the values they place on an intervention they are asked to value by virtue of being asked to consider it in the task.[74] That is, individuals focus their attention on the attribute or characteristic they have been asked to consider, relative to other considerations.[75,76] The implication is

[vi] Including in the area of reproductive health[69], the public health effects of water fluoridation[70], health promotion[71], screening information for life threatening and chronic diseases[72], hope in cancer care[73] and the value of information in multiple health care programmes.[20]

that such 'focus illusion' might result in an overemphasis being placed on the value of meta-health effects, with the resulting values elicited being higher than might have occurred in the absence of such focusing. Cookson (2003) suggests that in general, comparative WTP (between choice alternatives) is preferred to asking individuals for absolute WTP values as a means of overcoming some of these biases (presumably, because in undertaking a comparative WTP the biases will apply to both alternatives under evaluation).[74]

A potential source of 'focus illusion' is that WTP and other stated preference tasks use hypothetical scenarios as a means of asking respondents to make choices as if they are living in a given health state. It might be that such studies are not actually extracting existing preferences that reflect utility (or value) derived at the time of consuming a good or service. Rather, in such situations the respondent is forming preferences based on the information presented.[77,78] In this case, choices are based on the utility expected to arise as a result of a particular decision. Kahneman et al. (1997)[79] note that in such situations, when asked to express their expected utility, individuals will focus on the process of transitioning into rather than being in the health state of interest. The individual therefore focuses on what they would gain or lose from that transition, and fails to take account of adapting to be in the new health state.[75,76,79] This focus on gains or losses might bias the utility effect individuals expect from a particular choice.

Using prospect theory, Kahneman and Tversky (1979)[80] explain that how individuals respond to expected gains and losses is influenced by their reference or starting point. The larger a loss or gain appears to be relative to that starting point, the less or more attractive the choice option. In general, when evaluating possible choices, individuals exhibit loss aversion under which the impact of a loss is greater than a gain of the same magnitude, and they overweight the likelihood of rare events occurring.[80,81]

There is the potential that the impact of focus bias and failure to consider adaptation are magnified in the context of assessing value for meta-health effects where some of

the changes under consideration, such as changes in the process of care, are of a short duration and experienced intermittently. Focusing on meta-health effects within valuation tasks using stated preference methods, such as DCEs, might therefore result in over-valued health states. In particular, failure to consider adaptation, coupled with loss aversion, means individuals might over-state the impact of a change in a health state involving a meta-health effect. The extent of that bias might well be influenced by the ability to affect the perceived extent of the loss (or gain) by altering the information used in tasks designed to elicit preferences.

1.5 The Importance of How the Question is Asked: Framing

de Botton (2013)[82] suggests that art galleries could change the meaning and interpretation placed on artworks by altering the manner in which they are grouped together, and the captions used to introduce collections of works (e.g. "Gallery of Fear"). In so doing, he is articulating the importance of the content and context in which information is delivered to how it is interpreted. This is analogous to the concept of framing within preference elicitation, and the importance of how information is presented to the responses that are elicited from individuals.

In forming preferences or making choices, individuals adopt a frame that is their "*conception of the acts, outcomes, and contingencies associated with a particular choice.*"[83] The frame that the individual adopts depends partly on how the problem they confront is formulated and partly on their norms, habits and personal characteristics.[83,84] Woodhead et al. (2011)[85] and Froberg et al. (1989)[86] found that frame differed depending on whether individuals were more heavily influenced by their own experiences or the data with which they were presented.

Framing effects can influence the decision-making process in two ways: through contextual cues that cast a situation as positive or negative; or through priming, in which the way the task is presented to individuals leads them to draw upon pre-existing information that may or may not be germane to the decision task.[87] The

potential impact of framing is influenced by how individuals combine information in making decisions. Kahneman and Tversky (1984)[88] express this in terms of mental accounting or decision-making heuristics. They describe three approaches to mental accounting when faced with multi-attribute choices: minimal, topical and comprehensive accounts.[88] In the minimal account approach, the individual considers differences between options only and disregards any features they share. In using topical accounts, the individual relates the consequences of possible choices to a reference level determined by the context of the decision problem. Finally, in comprehensive accounts the individual considers all attributes, as well as other relevant factors when choosing between alternatives.[88]

How these approaches are applied is influenced by how the choice problem is framed.[88] Ideally, the manner in which a choice problem is formulated, or how the information is presented, should not alter the choices individuals make in preference elicitation tasks. However, whether the outcomes of a choice are presented as gains or losses, the extent of information provided about the alternatives in a choice problem, and the information provided as background context to a choice problem all have the potential to influence the decisions individuals will make in that task.

Positive or negative framing acts to make the item of interest more (positive framing) or less (negative framing) attractive within a preference or choice task.[89] Levin et al. (1998)[89] describe three types of framing effects that apply in such situations:

1. Risky choice: This presents a set of options with different risk levels measured through a comparison of risky options. Framing effects arise in how those risks are expressed e.g. different risks of individuals surviving (gain) or dying (loss).
2. Attribute: The object of the framing exercise is a single attribute or characteristic. Its effect is measured through the comparison of the attractiveness ratings for the single item. It is distinct from risky choice because a single feature (not an entire option) is being framed, and the riskiness is not being changed.

3. Goal: This refers to the consequence or implied goal of a behaviour. Framing in this context affects the uptake of the desired outcome, and is measured via a comparison of the rate of adoption of the behaviour. The impact of a persuasive message in goal framing depends on whether it stresses the positive outcomes of doing something, or the negative outcomes of it not being done. Both versions of the differently framed goal have the same intent.

There is the potential for framing effects to influence the values obtained for meta-health effects within preference elicitation studies. As previously noted, asking respondents to think about meta-health effects explicitly focuses attention on them as a source of value. This makes the manner in which those attributes are described, and the information provided about them, particularly important. Whether the language used to describe a meta-health effect is positive or negative will impact on whether it is perceived as a gain or loss e.g. "you are required to spend five hours for intravenous administration" compared with "you take a simple daily oral tablet". Similarly, how changes in a meta-health effect are described in the context of the overall health state (either highlighted as the only point of difference, or included as one potential point of difference within an overall health state) influences the importance they are given and the values attributed to them. Understanding the magnitude of these framing effects is important in order to assess the robustness of values determined for meta-health effects.

1.6 Approach to the Research in this Thesis

Meta-health effects have been investigated using a wide variety of approaches in a number of areas within health economics and health services research. Typically preference based approaches, such as TTO, SG and DCEs, have been used to investigate meta-health effects, including from investigating the value of convenience associated with pharmaceuticals[62-64,90-93] to the the influence of patient autonomy on the choice of primary care provider.[94-96] This thesis builds on those approaches in order to

investigate the role and value of meta-health effects in health care decision-making. It does so by addressing the following central question:

What is the influence of meta-health effects in health care decision-making?

This is informed by investigating the role of meta-health effects in individuals' decision-making regarding health care. It considers the value individuals place on meta-health effects in making health care decisions, and how those meta-health effects are likely to influence those decisions. The research addresses the following sub-questions:

How does the influence of meta-health effects compare with that of health effects? This considers the contribution of meta-health effects to the decision to use, or value, a health intervention relative to the other determinants of that decision (such as improvements in health).

What impact do meta-health effects have on values for use in economic evaluations? This considers the contribution of meta-health effects to the values used within an economic evaluation.

How does framing affect values derived for meta-health effects? This considers how framing influences the values derived for a meta-health effect in a stated preference setting.

Who should pay for meta-health effects? This considers whether the WTP for a meta-health effect varies depending on who bears the cost (the patient or the government).

These questions are explored using four case studies: Evidence Review; General Practitioner Loyalty; Rheumatoid Arthritis Therapy; and Mastectomy. These studies are summarised as:

Evidence Review: This study uses two approaches to develop an understanding of how meta-health effects have been investigated in the past. The first presents a review of decisions by the PBAC; it uses publicly available information to assess whether meta-health effects have been considered by the PBAC. This aspect of the research adds a perspective and level of detail that would otherwise not be available from the review of the literature alone.

The second approach focuses on a review of the literature on the exploration and valuation of meta-health effects. The results of that literature review are structured according to the type of meta-health effects considered, how the value of those outcomes was investigated, and whether there were any common drivers of the valuations obtained. The outcomes of both these pieces of research help to inform the remaining empirical research in this thesis.

General Practitioner (GP) Loyalty: This study assesses the role of meta-health effects in influencing the type of relationship individuals form with their GP. It addresses the question of whether meta-health effects influence consumers' decisions about health care. Using data from a community based survey of individuals' experiences and attitudes towards primary care services the decision to remain loyal to a GP practice or use multiple practices was investigated. The influence on that decision of meta-health effects outcomes such as choice, reflecting autonomy, and convenience was assessed.

This study examines the importance of meta-health effects on establishing long-term patient-GP relationships; a reflection of an ongoing situation, not necessarily in the presence of a health care condition (or acute health state).

Rheumatoid Arthritis (RA) Therapy: This study uses a DCE to investigate the value, including WTP, associated with the convenience of different modes of administering treatment for RA. In so doing it examines whether individuals value meta-health effects, and how those values compare with those they ascribe to health effects. The study was designed to not only provide individual valuations of convenience (derived from individuals' assessment of meta-health effects), but also to assess whether or not this differs depending on who pays for treatment (the patient or the government). It also considers the impact of framing effects on the values elicited for meta-health effects.

The role of meta-health effects in the choice of treatment in RA provides an important insight into value determination in the context of choosing medications for long-term treatment. This is particularly relevant for the decisions made by reimbursement bodies such as the PBAC. It provides a perspective on how meta-health effects influence acute decisions (treatment choice) for a chronic condition for which there are longer-term implications.

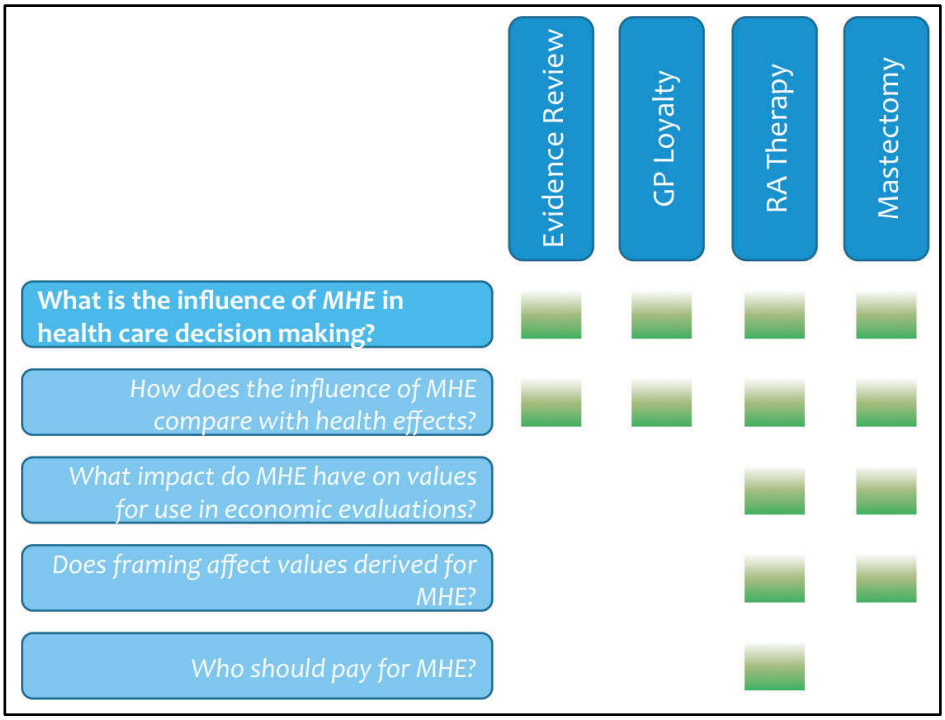
Mastectomy: This research combines qualitative (focus groups) and quantitative (DCE) studies to investigate the value (including WTP) associated with the decision to undergo contralateral prophylactic mastectomy (CPM) following a diagnosis of early stage breast cancer. It focuses on assessing the value to women of differences in the mode and frequency of ongoing monitoring (convenience), participation in decision-making (autonomy) and reductions in the fear of cancer recurrence (reassurance); all compared with changes in possible health outcomes. It investigates whether women value meta-health effects, and how those values compare with those they ascribe to health effects, in the presence of a potentially life threatening health event. The impact of framing effects in terms of the information presented is also tested.

Understanding the role of meta-health effects in this context is of interest insofar as

CPM is an invasive intervention that has long-term implications for both the individual and the health care system.

How each case study corresponds to the central and sub-questions of this thesis is outlined in the evidence map in Figure 1. This also shows where there are common themes across the studies that serve to address the overall research question.

Figure 1: Evidence map – link of research questions to sources of empirical evidence



Abbreviations: GP, general practitioner; MHE, meta-health effects; PSD, public summary document; RA, rheumatoid arthritis.

All four studies investigate the overall influence of meta-health effects in individuals' decision-making about health care. The Evidence Review provides an important starting point for this research: are meta-health effects being considered in current decisions by a Government agency in Australia, and what is the state of play in the literature with regards to their measurement for use in economic evaluations?

The three empirical studies (GP Loyalty, RA Therapy and Mastectomy) investigate the influence of meta-health effects using two different approaches to how those effects are

captured: attitudes and preferences. In this case, the distinction is that attitudes reflect an unconstrained (non-comparative) rating over an attribute whereas preferences are derived from questions that are comparative in nature, requiring a choice between options.[76,97] Within the GP Loyalty study, attitudes are used as a measure of meta-health effects and to assess their role, relative to the importance of other determinants, in the choice of patient-GP relationship. In the RA Therapy and Mastectomy studies, preferences sourced via DCEs are used to assess the role of meta-health effects. Attitude ratings are also used in the Mastectomy study as a means of exploring how preferences might differ between respondent groups based on their degree of cancer concern.

Together, these three studies provide the basis for a comparison of whether different elicitation methods (attitudes or preferences) have an impact on the ability to detect a role for meta-health effects in health care decision-making. All three studies (GP Loyalty, RA Therapy and Mastectomy) assess the importance of their respective meta-health effects to the decisions under investigation.

While the manner in which meta-health effects are measured differs between the GP Loyalty and RA Therapy and Mastectomy studies, all three investigate whether meta-health effects, health effects or the determinants of health, have more influence in the decisions in which they investigate. Common to all three is an assessment of the role of convenience in those decisions. Results of the Evidence Review in Chapter 2 indicate that this is the meta-health effect most commonly presented as a source of value to the PBAC, and most widely investigated in the literature. It was therefore included in all three studies to test whether the different health contexts influenced the importance placed on convenience. Other study-specific meta-health effects - choice in the GP Loyalty study, and autonomy and reassurance in the Mastectomy study – were included as preliminary evidence showed that they were likely to be relevant to those decision contexts.

In addition, the RA Therapy and Mastectomy studies provide values, in the form of marginal WTP, for their respective meta-health effects that can be used within an economic evaluation to inform decision-making. The use of attitude measures in the GP Loyalty study means it is not possible to express the influence of meta-health effects in that decision as a value for use in an economic evaluation. Nonetheless, it informs whether meta-health effects are relevant in decisions related to patient-GP relationships, and thus whether such effects might be important as a source of value for economic evaluations in that area.

The RA Therapy and Mastectomy studies also investigate a second methodological question of interest: how framing impacts on the value of meta-health effects relative to health effects. This was tested across three framing effects:

(1) Goal framing: in terms of the information provided as background (the DCE preamble).
(2) Attribute framing: the amount of information provided describing the relevant attribute of interest (in the mode of treatment administration in the RA Therapy DCE).
(3) Efficacy framing: the size of the treatment effect (e.g. the reduction in cancer risk associated with CPM).

Efficacy framing effects were tested in both DCEs, while goal and attribute framing effects were tested separately, one in each of the two DCEs.[vii] The justification for testing different framing effects in the two DCE studies is provided in Chapters 5 (RA Therapy) and 6 (Mastectomy) respectively.

1.6.1 Structure of the thesis

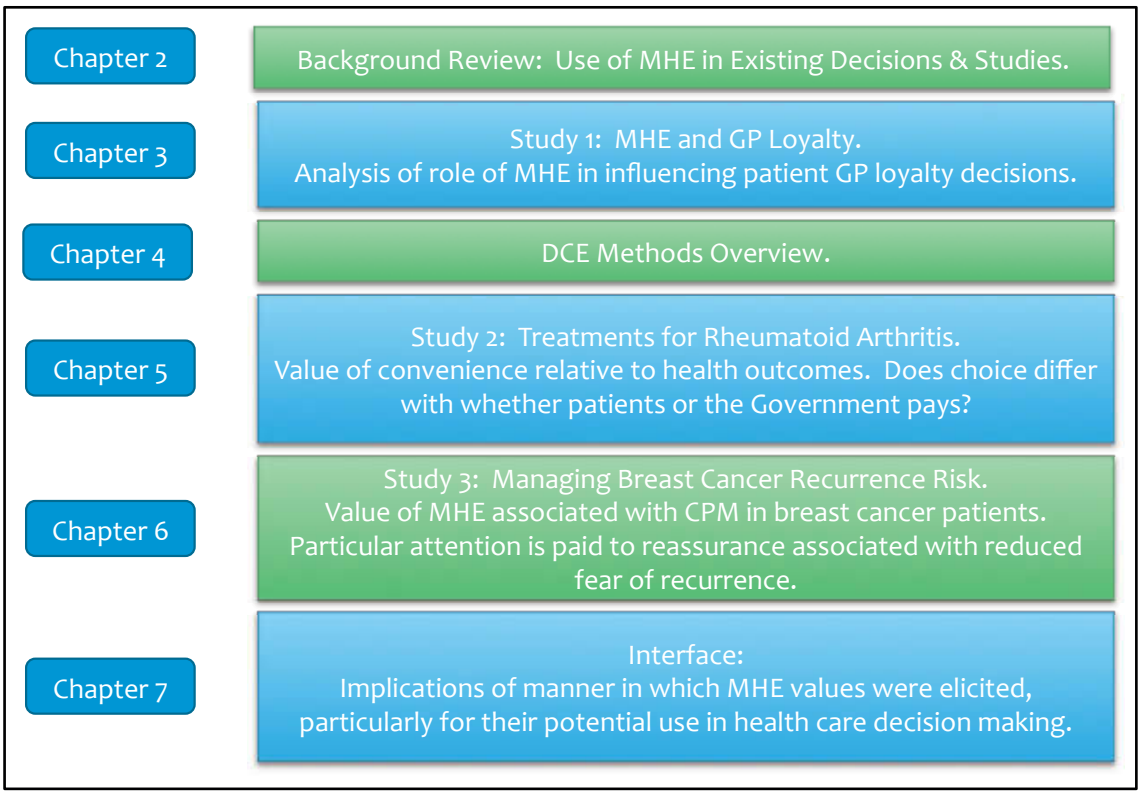
The structure of this thesis is outlined in Figure 2. The methods, results and implications of the Evidence Review are provided in Chapter 2. The role of meta-

[vii] It would have been possible to test all three in each, however this would have increased the number of possible survey versions and therefore the number of respondents required.

health effects in influencing GP loyalty is investigated in Chapter 3. The justification and methodological basis for the use of DCEs in the remaining empirical work is presented in Chapter 4. As well as providing an overview of the theoretical underpinnings of DCEs, the material in this chapter describes the general approach to the development of a DCE survey and the models that are subsequently applied in the analyses of the data presented in Chapters 5 and 6.

The empirical investigation of meta-health effects using DCEs is presented in Chapters 5 (RA Therapy) and 6 (Mastectomy). The implications of the findings from this thesis regarding whether and how meta-health effects might be considered in health care decision-making are discussed in Chapter 7. The importance of WTP for meta-health effects and whether patients or governments should pay for gains attributable to meta-health effects, are also considered.

Figure 2: Thesis structure



Abbreviations: CPM, contralateral prophylactic mastectomy; DCE, choice experiment; GP general practitioner; MHE, meta-health effects; PSD, public summary document.

2 Existing Evidence

Chapter Summary

The research in this chapter explores the extent to which meta-health effects appear in reimbursement decisions made by the PBAC in Australia. It also considers the evidence from the published literature on the consistency in how meta-health effects have been assessed between studies and across indications.

Publicly available information on PBAC applications (March 2008 and July 2014) was reviewed for evidence that meta-health effects had been considered. A separate systemative review of the published literature identified studies that evaluated a meta-health effect and reported utility values or WTP information suitable for use in an economic evaluation. The association between the type of meta-health effect investigated, evaluation method, year of publication, health care intervention and source of study funding was tested. Potential drivers of differences in values for meta-health effects were explored by comparing studies with the highest and lowest values respectively for a given meta-health effect, controlling for the evaluation method.

Of 407 applications to the PBAC reviewed, 35 contained claims of benefits that could be considered to be meta-health effects. The majority of these claims were based on differences in convenience associated with the route or frequency of administration.

The subsequent systematic review of the literature identified 71 studies, 35% of which investigated drug therapies. Convenience, information and process of care were the meta-health effects most often investigated; with approximately half using DCEs or conjoint analysis methods. Associations were found between the meta-health effect being investigated and the evaluation method, intervention topic (e.g. convenience and drug interventions were associated) and the source of funding for research. Controlling for study methods and type of meta-health effect; differences in the magnitude of the effect evaluated, how the meta-health effect was

described, and how it was framed in comparison to overall health might explain differences observed in the values derived between studies.

The research in this chapter contributes to the literature by adopting a broad approach to the consideration of meta-health effects, and placing this evidence in the context of that used by a public reimbursement body, the PBAC. The evidence from the literature shows variability in the methods applied to the valuation of meta-health effects that influences the values obtained and their subsequent interpretation. Some of that variability appears to be due framing effects. This provides an impetus for the investigation in this thesis of the influence of framing effects in the assessment of meta-health effects. Understanding the influence of effects such as framing is important if decision-makers are to have confidence in the methods used to value meta-health effects.

2.1 Introduction

Typically, decision-making for the reimbursement of new medicines or health care services focuses only on the health outcomes of health care. However, there is increasing recognition that the consumption of health care can have broader impacts on individuals' QoL and well-being beyond health outcomes. In this thesis, these impacts are referred to as meta-health effects. Meta-health effects may be used to differentiate new products when seeking reimbursement, but there is considerable debate over how such impacts should be measured and valued for inclusion in decision-making.

Understanding how meta-health effects have been valued for use in economic evaluations and decision-making in the past is important in order to consider how they might be valued to inform such decisions in the future. The research in this chapter develops this understanding in two ways. First, it explores the extent to which meta-health effects are being presented and considered in reimbursement decisions made by the PBAC in Australia. This utilises publicly available information summarising the nature of claims made to the PBAC in applications for reimbursement, and the

subsequent recommendations made by the PBAC regarding those applications. Second, to supplement the available evidence on the inclusion of meta-health effects in decisions by the PBAC, a search of the published literature was conducted as a means of assessing the breadth and consistency across studies and indications of approaches used to measure such effects.

2.2 Meta-Health Effects in PBAC Decisions

2.2.1 Introduction

In considering submissions for drug reimbursement, the PBAC has a preference for the use of CUA (presenting incremental costs-per QALY).[24] The Australian PBAC guidelines explicitly refer to the use of meta-health effects as a potential source of value for use in those reimbursement submissions:

> *PBAC may also consider nonhealth [sic] outcomes, including aspects of the delivery of a health care intervention beyond the health gain obtained; for example, greater convenience or production gains to society beyond those valued by the population benefiting with improved health. However, the valuation of nonhealth [sic] outcomes is not straightforward and those outcomes might not be as influential in decision making as health outcomes*.[24]

This first part of the evidence review draws on the publicly available information on reimbursement decisions made by the PBAC between 2008 and November 2014 to investigate the extent to which claims regarding meta-health effects have been considered by the PBAC, and to identify the nature of those claims. The aim was to identify applications considered by the PBAC in which the impact on a meta-health effect was used to support the requested price/listing claim. This was addressed by summarising the number of applications in which a meta-health effect was included as part of a reimbursement claim, and describing the basis for those claims. The methods,

results and discussion of that review of PBAC decisions are presented in the following sub-sections.

2.2.2 Method

The primary interest of the review of PBAC decisions was in applications that included a value for a meta-health effect within the economic evaluation. Applications in which a potential meta-health effect was identified (either by the sponsor or the PBAC) and discussed as a potential point of difference, regardless of whether or not it was noted as being included in the economic evaluation, were also of interest. The following process was applied in cataloguing public summary documents (PSDs[viii]):

1. PSDs for the seven years (March 2008 to July 2014) were accessed via the Pharmaceutical Benefits Schedule online (http://www.pbs.gov.au)
2. Applications were catalogued as drug-indication pairings (not as individual considerations) by the number of times they had been considered by the PBAC.
3. For a given drug, each PSD was recorded separately, noting if the PBAC had previously considered that drug in that indication at a prior meeting (i.e. whether the latest PSD in that drug-indication pairing could be considered to be a re-submission).
4. In the majority of cases, the review of drug-indication pairings focused on the latest PSD unless there was insufficient material contained in that document to discern if meta-health effects had been considered by the PBAC; or the PSD included language that suggested meta-health effects might have been considered at a previous PBAC meeting. In those instances, prior (re)submissions for that drug-indication pairing were also reviewed.

[viii] PSDs are publicly available (online) summaries of the evidence submitted to the PBAC, and its resulting recommendation, when considering applications for reimbursement. PSDs are only available for major submissions (those with an economic evaluation assessed by external evaluation groups) and not minor submissions (those which are not externally assessed).

All PSDs for the latest PBAC consideration for each drug-indication pairing were compiled into one searchable document. Following an initial manual review and PSD classification, key-terms for meta-health effects were used to search the compiled document. All PSDs identified as potentially relevant were reviewed a second time to ensure they met the inclusion criteria. PSDs were classified as containing a meta-health effect claim if they satisfied one of the following criteria:

1. Specific mention of the terms "non-health gains" or "non-health outcomes" in reference to the comparison under evaluation[ix];
2. Inclusion of terms reflecting meta-health effects (such as mode of treatment administration, frequency of treatment administration, location of treatment, convenience);
3. Reference to satisfaction with treatment, preferences for treatment, acceptability of treatment, or patient emotion and well-being (e.g. comfort, distress, anxiety) as potential differentiators; and
4. Specific reference by the PBAC, either directly or in reference to the submitted consumer comments, to additional value not captured by health gains.

PSDs that referred to the inclusion of cost offsets alone to allow for differences in the mode or frequency of treatment administration were not classified as eligible for inclusion.[x]

Data-extraction from the PSDs focused on the information contained in the sections in those documents labelled background/history, indication, comparator, the clinical claim, consumer comments (for PSDs after 2014), the economic model, pre-modelling

[ix] The term "meta-health effect" used to describe the concepts investigated in this thesis is not yet in common usage.

[x] For example, golimumab for RA was cost-minimised to adalimumab and etanercept but at a slightly higher price to reflect a difference in the resources required for administration (due to differences in dosing frequency and use of physician assistance). While such differences in treatment administration are potentially a source of value, there is no such suggestion within these PSDs that the intent was for these cost-offsets to reflect the value to the patient of the convenience associated with reduced dosing frequency.

studies (where included in the PSD) and the PBAC recommendation. Information extracted from the PSDs included:

a. the drug name (and indications if there were more than one);
b. the meeting(s) at which it was considered;
c. the basis on which the PBAC made its recommendation, classified as: comparison based on health outcomes only (superiority claim); cost-minimisation analysis, no differentiation; clinical only;
d. the specific terms related to the meta-health effect, where relevant;
e. for PSDs classified as including a meta-health effect claim, the comparator and basis of the claim made by the sponsor (clinical/safety superiority/non-inferiority/inferiority) was also extracted and
f. the PBAC recommendation itself: rejected, positive, deferral, partial. A partial recommendation covered those instances where a submission contained multiple requests, some of which resulted in a positive recommendation and others a rejection.

2.2.3 Results

A total of 407 drug-indication pairings and their accompanying PSDs arising from PBAC deliberations between March 2008 and July 2014 (inclusive) were catalogued for review. These covered at least 624 PBAC deliberations (this number is an underestimate as a number of PSDs referred to prior submissions but did not provide the date or number).

Of the 407 drug-indication pairings reviewed, 273 (67.1%) were recommended for listing by the PBAC; 118 (29.0%) were rejected. Through the use of the criteria for consideration of meta-health effects, 41 applications were initially identified for further review. More detailed review of those PSDs for this research resulted in the exclusion of six due to the inclusion of indirect effects (2), the use of cost-offsets that did not reflect patient values (3), or no direct reference to meta-health effects (1). The remaining 35 drug-indication pairings that included a claim referring to a meta-health

effect, or mentioned a meta-health effect in differentiating between products, are presented in Table 2. Among those 35 applications, 26 (74.3%) were recommended for listing by the PBAC.

The 35 drug-indication pairings identified were classified into those in which it could be discerned from the PSD that the sponsor had included a claim based on a meta-health effect (n=20), or those where in making its recommendation, the PBAC had subsequently noted a potential claim based on a meta-health effect, either directly or in reference to input received via the consumer comment process on submissions being considered by the PBAC (n=15).

In over half of those in which the sponsor made a claim based on a meta-health effect, the source of difference referred to treatment administration, notably injection schedules or the difference between oral and injectable therapies. The sources of difference in the remaining applications, and those classified as noted by the PBAC, were more diverse (including convenience, patient satisfaction/acceptability, coping, freedom from transfusion, reduced anxiety, comfort and dignity).

2.2.4 Discussion and limitations

This is the first review of PSDs of applications to the PBAC for evidence of whether meta-health effects are being considered in decision-making by the PBAC in Australia. The results show that the PBAC is being presented with, and separately considering, meta-health effects in its decisions. Not surprisingly, the bulk of the applications that included such effects claimed an improvement in the mode of administration as the key differentiator. This might suggest that convenience (improved administration) is the meta-health effect most often incorporated into superiority claims made by sponsors of submissions to the PBAC. Changes in mode of administration might also be associated with other meta-health effects such as reassurance (associated with reduced needle phobia), but this could not be discerned from the PSDs. Overall, the inclusion of meta-health effects did not appear to result in a difference in the resulting

recommendations from the PBAC, with the proportion of positive recommendations largely mirroring that of applications that did not include a meta-health effect claim.

This research has relied on the PSDs explicitly referencing all sources of value upon which the claims in the applications to the PBAC relied. However, the information included in PSDs is based on agreement between the sponsor of the application and the Department of Health. Accordingly, there are inconsistencies between similar applications that mean that it was possible that some applications in which meta-health effects were included were not identified because those terms were not referenced in the PSD. For example, lenalidomide for -5q myelodysplastic syndrome (MDS) was included on the basis of using freedom from transfusion as an endpoint of relevance. However, azacitidine for MDS also used freedom from transfusion as the endpoint of relevance and source of utilities (personal knowledge), but this was not noted in the PSD for that drug. Similarly, the initial application for paliperidone long acting injections included a TTO study to value the difference from other more frequently administered anti-psychotics (personal knowledge). While the underlying TTO has been published[63], it was not referenced in the relevant PSD. While inclusion of these additional drug-indication pairings in the review would provide more examples where meta-health effects have been included (adding to this dataset), it violates the research method insofar as it deviates from a consistent approach to allocating relevant drug-indication pairings for consideration.

PSDs are also not required for minor applications (those which do not require assessment by an external evaluation group prior to consideration by the PBAC). The result is that in classifying the outcomes from the PBAC, some information might not be available if it pertains to decisions reached as the result of a subsequent minor application not referenced in the available PSDs. Thus, in using the information from the available PSDs to report on meeting outcomes (positive recommendations, or rejections), the outcomes from minor applications might not be available if not

appended to the preceding meeting's PSD. This means that the assessment of the number of positive recommendations in this review is likely to be an underestimate.

Table 2: Appearance of meta-health effects in PSDs

Product	PBAC Meetings	Comparator	Main Claims	Terms used related to meta-health effects	Outcome
Drug-Indication pairings in which the sponsor included meta-health effects/process utility in the clinical claim or the economic evaluation					
1. Aflibercept - AMD	Mar-12	Ranibizumab	Clinical non-inferior, reduced treatment burden.	**The submission also claimed that less frequent injections required for aflibercept will help to reduce the burden on patients, caregivers, physicians and the health care system.**	Positive
2. Botox - lower limb post stroke	Jul-08	Placebo	Clinical superiority, equivalent safety.	Utility adjusted for injection schedule (*not clear what this adjustment was; was there disutility due to injection, or did they assume that a change in the schedule resulted in different QoL because of different utility gain?*).	Rejected
3. Botox – idiopathic overactive bladder	Nov-13	BSC	Clinical superiority, inferior safety.	The model used a disutility score of 0.25 for one day for the administration of botulinum toxin – these effects were independent of AEs.	Positive
4. Budesonide foam enema – ulcerative colitis	Jul-13	Prednisolone enema	Clinical non-inferiority, superior safety.	Submission's claim of greater patient preference to the foam preparation was adequately supported.	Positive
5. Cladribine – Multiple sclerosis	Mar-11	Natalizumab and interferon-beta	Clinical non-inferiority (natalizumab), superiority (interferon-beta).	TTO to value oral vs injectable therapies (intramuscular and subcutaneous).	Rejected
6. Dimethyl fumarate – Multiple sclerosis	Jul-13	ABCR therapies	Clinical superiority, non-inferior safety.	More convenient route of administration than interferon based therapies.	Positive
7. Exenatide - Diabetes	Nov-13; Jul-13; Jul-11	Exenatide b.d.	Clinical superiority, non-inferior safety.	The model was sensitive to disutilities associated with differing administration schedule (weekly versus daily).	Rejected

Product	PBAC Meetings	Comparator	Main Claims	Terms used related to meta-health effects	Outcome
8. Fingolimod – Multiple sclerosis	Mar-11	Interferon-beta and natalizumab	Clinical superiority, non-inferior safety.	Value of oral compared with different injectable therapies (including potential benefits, and whether there is disutility associated with intra-muscular injections).	Positive
9. Gefitinib – EGFR NSCLC	Jul-13; Nov-12; Nov-10	Carboplatin plus paclitaxel	Clinical superiority and safety.	**The additional QALYs gained in the proposed scenario rest on utility gains from delaying progression, deferring utility decrements in later health states, and the improved quality of life (QoL) associated with gefitinib treatment (oral administration, less serious toxicity), as opposed to doublet chemotherapy.**	Positive
10. Icatibant - C1-esterase inhibitor deficiency	Jul-11; Jul-10	Placebo	Clinical superiority, inferior safety.	**The incremental benefits of icatibant over best supportive care are almost entirely due to elements not directly relating to an attack (i.e. convenience and reduction of anxiety in the asymptomatic phase), rather than the incremental gain in clinical effectiveness.**	Rejected
11. Lenalidomide - -5q myelodysplastic syndrome	Mar-13; Jul-11; Mar-11	BSC	Clinical superiority, inferior safety.	Transfusion independence - mixed outcome - reflects health outcome and process.	Positive
12. Liraglutide - Diabetes	Mar-13	Exenatide	Clinical and safety non-inferiority.	Flexibility of injections, needles avoided, flexibility of diet	Positive
13. Methylnaltrexone – constipation	Jul-09; Mar-09	Placebo	Clinical superiority, non-inferior safety.	**Providing comfort and dignity in the end of life situation**	Positive
14. Mifepristone – medical termination	Mar-13	Surgical termination	Clinical and safety non-inferiority.	**Proportion of women with complications or adverse events related to the termination, (including duration and extent of haemorrhage, distress or depression); Acceptability (i.e. proportion of women who would opt for that procedure again if necessary).**	Positive
15. Panitumumab – Kras wild type CRC	Mar-13; Nov-08	Cetuximab	Clinical and safety non-inferiority.	**The submission also noted that panitumumab was associated with a much lower incidence of infusion reactions and requires less frequent administration, which provides an**	Partial

Product	PBAC Meetings	Comparator	Main Claims	Terms used related to meta-health effects	Outcome
				imperative to list.	
16. Pregabalin – neuropathic pain	Mar-12; Mar-11	Amitriptyline	Clinical and safety non-inferiority.	No treatment - satisfied/not satisfied (with pain control).	Positive
17. Ribavirin and PegIntron – hepatitis C	Jul-08	Usual care.	Clinical superiority, inferior safety.	Disutility of treatment (*not clear if this is AE or treatment itself*).	Positive
18. Tobramycin – PA infection in CF	Mar-13	Tobramycin solution	Clinical and safety non-inferiority; superior satisfaction.	Patient satisfaction, treatment administration, WTP, TTO to value mode and duration of treatment administration.	Rejected
19. Trastuzumab – breast cancer	Jul-12	Adjuvant chemotherapy.	Clinical and safety non-inferiority.	Patient preferences taken into account (MOGA submissions).	Positive
20. Zoledronic acid - osteoporosis	Nov-08; Jul-08	Alendronate sodium	Clinical and safety non-inferiority.	TTO to value difference in treatment (injection) administration frequency.	Positive
Drug-indication pairings in which the PBAC noted meta-health effects/process utility was a potential source of value (but this was not apparent in the sponsor's claim)					
21. Abiraterone	Nov, Mar, Jul-12; Nov-11	BSC	Clinical superiority, non-inferior safety.	More convenient to administer (sponsor had included difference in administration costs, PBAC noted this and appears to include convenience as a source of difference).	Positive
22. Alemtuzumab - MS	Jul-2014	Fingolimod and natalizumab	Clinical superiority (long-term).	**The comments described a range of perceived benefits of treatment with alemtuzumab including its use in patients with aggressive forms of multiple sclerosis, and the convenience of yearly administration of alemtuzumab.**	Positive
23. Capecitabine - CRC	Mar-11	Oxaliplatin	Clinical and safety non-inferiority.	**The PBAC also acknowledged the option of an oral treatment for patients in this treatment setting was valuable.**	Positive
24. Dabigatran – Atrial fibrillation patients	Mar-13; Jul-12;	Warfarin and aspirin	Clinical and safety	PBAC discussion refers to potential meta-health effects and process benefits regarding patients not keen to take warfarin.	Positive

Product	PBAC Meetings	Comparator	Main Claims	Terms used related to meta-health effects	Outcome
	Mar-11		superiority.		
25. Dolutegravir - HIV	Nov-13	Raltegravir	Clinical and safety non-inferiority.	**Reduced dosing interval and reduction in pill burden for the majority of dolutegravir patients, potentially leading to an increase in compliance.**	Positive
26. Ferric carboxymaltose - anaemia	Nov-13, Mar-13	IV iron polymaltose	Clinical and safety non-inferiority.	**The comments described a range of benefits of treatment with ferric carboxymaltose including reduced infusion time, greater convenience, use of fewer hospital resources, fewer side effects and improved quality of life.**	Positive
27. Fluticasone & Vilanterol FDC - asthma	Mar-14	Fluticasone proprionate/ salmeterol FDC	Clinical and safety non-inferiority.	**The convenience of a once daily formulation may promote greater use and increase substitution of single agent ICS preparations by fixed dose combinations of ICS/LABA.**	Positive
28. Fluticasone & Vilanterol FDC - COPD	Jul-14, Mar-14	Fluticasone propionate and salmeterol FDC	Clinical and safety non-inferiority.	**The PBAC noted that as FF/VI was dosing once daily compared with twice daily for FP/SAL, patient preference for a once daily dose may drive utilisation. The PBAC considered however that once daily compared with twice daily dosing was unlikely to yield an adherence benefit.**	Positive
29. Goserelin – Breast cancer	Nov-09	Chemotherapy	Clinical non-inferiority, superior safety and QoL.	**PBAC also noted the significant advantage of goserelin for some women in that it avoided the need to cope with infertility contemporaneously with a diagnosis of breast cancer.**	Positive
30. Ingenol – solar keratosis	Jul-14; Nov-13; Nov-12	Placebo	Clinical superiority.	**It was also recognised that patients may prefer treatment with three days of ingenol compared to comparators such as 5-FU cream or cryotherapy, and that this preference could drive utilisation.**	Rejected
31. Ivermectin – ATSI scabies	Mar-14	Placebo	Clinical superiority, non-inferior safety.	**The PBAC considered that there is a high risk of leakage to first-line treatment in patients with typical scabies and household contacts, given the inconvenience and poor adherence to topical treatment, the convenience of taking a tablet, and the possibility of other untreated patients with infestations or household contacts sharing unused tablets remaining in a pack.**	Positive

Product	PBAC Meetings	Comparator	Main Claims	Terms used related to meta-health effects	Outcome
32. Oxycodone - pain	Nov-10, Jul-10	Oxycodone controlled release	Clinical superiority, non-inferior safety.	Noted abuse potential of the comparator (not apparent in oxycodone).	Positive
33. Pomalidomide - MM	Mar -14	High dose dexamethasone	Clinical superiority, inferior safety.	**The comments described a range of benefits of treatment with pomalidomide including giving patients more time until the next advancement or cure, the convenience of an oral dose form, need for affordability of the drug, and an unmet need in patients with relapsed/refractory myeloma following failure of lenalidomide and bortezomib.**	Rejected
34. Plerixafor – stem cell mobilisation	Nov-13; Jul-12; Nov-11; Nov-10	GCSF plus chemotherapy	Clinical superiority	The comments highlighted a number of perceived benefits of use of plerixafor including enabling more patients to proceed to transplantation, decreased pain and discomfort, and reducing anxiety associated with failing mobilisations. (Nov 13) Many benefits not accounted for, but difficult to address in PBS (Jul 12)	Positive
35. Sofosbuvir – hepatitis C	July-14	Standard of care dependent on genotype.	Clinical and safety superiority.	**Sofosbuvir is the first oral direct acting antiviral agent that can be used to treat HCV genotypes 1-6 and provides patients with the first interferon-free treatment option. These benefits of new treatments were highlighted by the large number of comments from patients, health care professionals and organisations.**	Rejected

Note: Text in bold has been copied directly from the relevant product PSD.
Clinical/safety superiority/ non-inferiority/inferiority requires the proposed therapy to be superior/non-inferior/inferior to the proposed comparator as nominated by the application sponsor.

Abbreviations: AE, adverse event; ABCR therapies, intramuscular interferon beta-1a, subcutaneous interferon beta-1a, interferon beta-1b and glatiramer acetate; AMD, age-related macular degeneration; ATSI, Aboriginal and Torres Strait Islander; BSC, best supportive care; CF, cystic fibrosis; COPD, chronic obstructive pulmonary disease; CRC, colorectal cancer; EGFR, epidermal growth factor receptor; 5-FU, 5 fluorouracil; FDC, fixed dose combination; FF, fluticasone; GCSF, growth colony stimulating factor; HCV, hepatitis C virus; ICS, inhaled cortico-steroid; K-RAS, Kirsten rat sarcoma; LABA, long-acting beta agonist; MOGA, Medical Oncology Group of Australia; MM, multiple myeloma; MS, multiple sclerosis; NSCLC, non-small cell lung cancer; PA, pseudomonas aeruginosa; PBAC, Pharmaceutical Benefits Advisory Committee; PBS, Pharmaceutical Benefits Scheme; QoL, quality of life; SAL, salbuterol; TTO, time-trade-off; VI, vilanterol; WTP, willingness to pay.

2.3 Valuing Meta-Health Effects for Use in Economic Evaluations: A Review of the Published Literature

2.3.1 Introduction

The intention of this systematic literature review is to consider how meta-health effects have been measured previously, whether those methods produced values that could be used in economic evaluations and, if so, the extent to which patterns could be discerned in the methods used or resulting values between meta-health effects. The focus is on values for use in economic evaluations as these are often a key source of information used to assist in decisions about resource allocation in health care.

The literature review proceeds with a description of the search terms and scope. The criteria used to categorise meta-health effects, in deciding whether resulting values are appropriate for use in an economic evaluation, and to assess the comparability across studies of the manner in which values for meta-health effects were derived are subsequently described. The literature search results, along with a summary of the findings in terms of which meta-health effects were evaluated and an assessment of the values elicited for those meta-health effects, are subsequently presented. The literature review concludes with a discussion of the implications of these findings for the general literature and this thesis.

2.3.2 Methods

2.3.2.1 Literature searches

Searches were conducted of the Embase, Medline, Econlit, Cinahl, SocIndex and Psychinfo Databases for publications reporting on the valuation of meta-health effects using a combination of the terms listed below. Given the emergence of DCEs in this field, search terms synonymous with that field were included specifically to allow for the possibility that those studies might not have been indexed under other study types. Studies were not limited by publication type or year; searches were conducted for the period from database inception up to April 2016.

The search terms used were based on a combination (using the term "AND") of the following subject and measurement terms:

- Subject terms:
    - non-health outcome/effect/impact
    - beyond health/extra health effects
    - process effect/utility
    - patient preference (psychology)
- Measurement terms:
    - utility assessment/valuation/measurement/weight/value
    - quality of life
    - patient/societal preference
    - valuation or value assessment
    - willingness-to-pay or contingent valuation
    - conjoint analysis or discrete choice or stated preference

Since the term 'meta-health effect' has not been used previously to describe the concepts of interest for this research, it was not used as a term in the literature search. However, it was part of the exclusion criteria applied in reviewing citations retrieved via that search, to wit:

- A direct meta-health effect was not included.
- The study reported an economic evaluation or DCE without a meta-health effect.
- The citation was for a letter/commentary/news item/or editorial that lacked sufficient information/detail for further analysis.
- The study was a clinical/preference study that did not include the valuation of meta-health effects (e.g. a simple rating study or statement of preference for one intervention over another).
- The publication was subsequently superseded by a later update; or reproduced the results of a previous publication or study.

Only papers reporting on primary research were included in the systematic review. Systematic reviews identified in the search were initially retained for review of their references lists to ensure all relevant studies were included, and for consideration of those systematic reviews in the discussion of the results of the current review.

For each included study, the data extracted were those referring to how it was conducted and its subsequent outcomes. The following data were extracted into a Word table: authors; year of publication; the intervention being investigated; the meta-health effect; whether a value was reported that could be used in an economic evaluation; the method of assessing value; the measure of value (utility or WTP) reported; the source of funding for the study; and the purpose of the research. All studies were reviewed and data extraction conducted by the same researcher (the PhD candidate). The following sub-sections describe the classifications of these data and their analysis.

2.3.2.2 Categorisation of meta-health effects

Meta-health effects reported within studies deemed relevant for inclusion were categorised using an approach similar to that of Haas (2002)[98] in her PhD investigating non-health outcomes arising from GP services. The approach was as follows:

1. Where a study specifically categorised an outcome as a meta-health effect (e.g. elation or regret in Salkeld et al. (2004)[99]) those categorisations were retained.
2. Meta-health effects that appeared in a health state description, such as in a TTO, SG or CV task, or as an attribute in a DCE, were categorised according to the definitions in Table 3.

Table 3: Categorisation of meta-health effects

Meta-Health Effect	Product or Health State Element
Convenience	Method of treatment/service administration where delivery or accessing the service relies predominantly on the patient. Includes concepts of acceptability of the product/service in terms of physical effects (discomfort etc.), visual impacts (aesthetics), and disruption to normal routine (the method of administration, duration of administration and frequency of administration).
Process of Care	Method of health care intervention (where delivery or access relies solely on third parties)
Information	Information about the product, a service, a health care condition (including future risks).
Hope	Expectations about positive outcomes for the future.
Anxiety/fear	Discomfort/stress about a current situation or about prospects for the future.
Access	The ability to obtain care.
Reassurance	Reductions in anxiety, or having expectations for the future affirmed.
Autonomy	Self-determination or a share of voice in the process of care.

Note: The definitions for hope, anxiety/fear and reassurance were in part informed by those used by Gandjour (2001), Koszegi (2003) and Salkeld et al. (2004) respectively.[6,99,100]

3. Where a duration of treatment attribute or description was included in a health state description or DCE (such as in Bunge et al. (2013)[101]), it was not listed as contributing to a meta-health effect if the intent of the attribute was to reflect the time to treatment effect rather than the time required for the administration/consumption of care.

2.3.2.3 Categorisation of methods of valuation

During the process of data extraction, each study was classified based on the method used to value preferences:

1. WTP via a CV approach;
2. Direct utility elicitation via a TTO/SG or an approach based on modifications of those methods (such as the waiting time-trade-off);
3. A stated preference DCE or conjoint analysis (CA); and
4. Multiple method studies that used a combination of two or more of these approaches within the one study.

Classifying studies in this way allowed an examination of the association between meta-health effects and the study valuation method.

2.3.2.4 Categorisation of interventions studies

Similarly, the link between the meta-health effect under investigation and the overall intervention or programme considered in each study was examined. This was facilitated by categorising studies according to the intervention/programme type investigated: device, diagnostic service, drug, health promotion, multiple types of services, obstetric services, population health, primary care, screening, surgery, telemedicine and vaccines. An assessment of the link between the year of publication and the meta-health effects investigated was also conducted for the most commonly cited meta-health effects.

2.3.2.5 Classification of sources of funding

For each study reviewed, information was extracted on the source of funding for the research and the reason for the study. Funding was classified as being from commercial sources, research grants, public funding (being funding from a public sector organisation that was not specifically a research grant), or not discernible. The primary reason for each study as could be ascertained from the publications, were categorised as follows:

1. Inform policy – the publication included a specific statement that the authors intended the results to inform policy or decision-making, or included a discussion of policy implications, without fulfilling one of the other criteria in this category.
2. Innovation – the study specifically mentioned the word innovation in the research purpose or discussed innovations in health care delivery.
3. Research focus methods – focus on the use of the particular method of preference elicitation (not the actual values derived).
4. Research focus preferences – focus on the preference values derived as the main motivation for the study.

5. Trial focus – reporting of a preference study, or comparison of an intervention, conducted alongside a clinical trial.

2.3.2.6 Assessing the relative value of meta-health effects

Each study was classified in terms of whether or not it provided values that could be used within an economic evaluation. This was based on whether the study reported mean WTP values in the event that it was a CV or DCE study, or utility values bounded by 0 and 1 and anchored by a survival/time metric as is required for use in estimating QALYs in the case of TTO/SG or DCE studies. The latter criterion was included to distinguish between the use of the term 'utility' value in the DCE literature to refer to the coefficients produced in regressing the resulting stated preferences (the choices) on the underlying explanatory variables (typically the attributes used within the DCE task) and those derived from TTO/SG tasks. Typically, the former are not anchored in survival unless a survival attribute has been included in the DCE task and the utility values have been expressed as marginal rates of substitution of the attribute of interest with respect to the survival attribute (see Norman et al. (2013)[41] for an example of the use of a DCE task to produce utility values on the 0-1 scale). In the absence of the inclusion of a survival or risk attribute, the re-anchoring of DCE coefficients onto a 0-1 scale for interpretation as utility values for QALY estimation has been conducted (see Ratcliffe et al. (2009)[102] for an example). DCE studies that produced utility values through rescaling could be included in this review.

For those studies that included a meta-health effect, and were classified as providing information that could be used for an economic evaluation, information was subsequently extracted on the reported mean effect size (WTP or utility value) and the effect size for the meta-health effect.

Using this information, a standard measure of the relative value of the meta-health effect was derived based on the ratio of the reported value (utility or WTP) for the meta-health effect itself to the overall mean value for the product or intervention under

consideration. This standard measure was developed in order to compare meta-health effect values across studies, allowing for differences in the concept being valued, as well as the study methods. In this was way the comparison is of the relative contribution of the meta-health effect to the overall reported value, rather than comparing across different types of measures in their base units (utilities or WTP). Relative effect sizes were only estimated for those studies which reported a separate and statistically significant value for the meta-health effect.

To estimate the relative effect of each meta-health effect, the overall mean effect estimate (WTP/utility value) was used as the denominator and the increment attributed to the meta-health effect as the numerator. The exception was where the value for the meta-health effect was derived via the difference between two health states. For example, in Osborne et al. (2007)[64] the value for convenience is given by the difference between the oral administration health state and the subcutaneous administration health state for iron chelation therapy. In this and similar instances, the health state with the higher value was used as the numerator to estimate the contribution of the meta-health effect to that value. The resulting relative values were reviewed to assess whether there were meta-health effects that might be considered outliers (either consistently higher or lower) compared with the others, or some valuation methods that produced higher or lower relative values than others.

2.3.2.7 Study comparisons and tests of association

The quality of included studies was not assessed formally using a quality appraisal tool. Where studies were compared, or were noted to be outliers (reported the lowest or highest values for a given meta-health effect, controlling for valuation method) in terms of values reported for their meta-health effects, consideration was given to the quality of the study conduct and its reporting.

Tests for possible associations between meta-health effects and study characteristics (such as year of publication, intervention investigated, funding source, etc.) were

included to explore which study related factors might influence whether or not meta-health effects were valued. These tests were carried out using the Pearson χ^2, with Cramer's V (allowing for differences in the number of elements compared in each table) shown as a subsequent measure of the strength of association. All statements regarding significance of association are based on the results of the χ^2.

2.3.3 Results

2.3.3.1 Findings in the literature

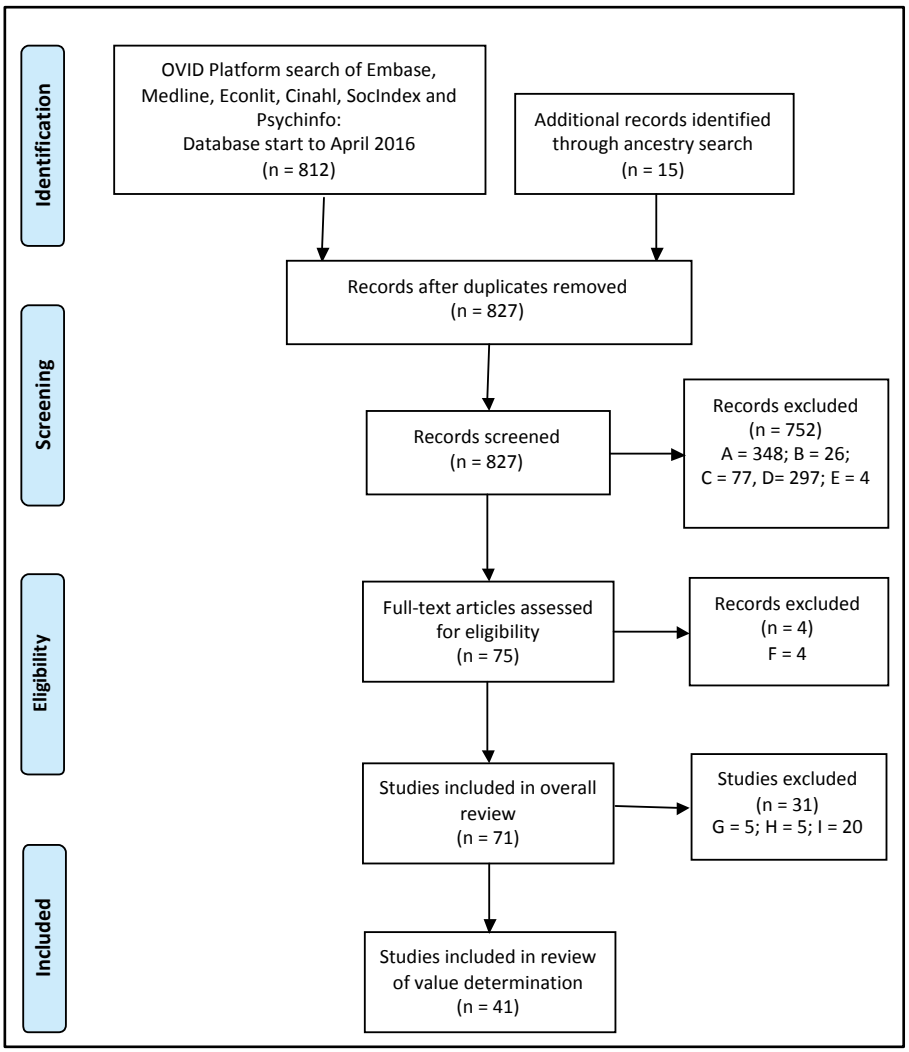
A total of 812 unique citations were retrieved and imported into EndNote for review, as outlined in Figure 3. The review of citations revealed 60 studies deemed relevant for inclusion on the basis that they reported on the valuation of a meta-health effect. Of these, four were systematic reviews: Opmeer et al. (2010), Sadique et al. (2012), Brennan et al. (2013) and Higgins et al. (2014).[5,39,93,103]

Opmeer et al. (2010)[5] conducted a systematic review of preferences for meta-health effects using 567 studies. However, the authors did not provide a list of the studies reviewed (so that it was not possible to cross-check with those included for the purposes of the current review). Attempts to contact the lead author did not prove fruitful. Sadique et al. (2012)[103] conducted a systematic review specifically focused on preference elicitation in the management of cervical abnormalities. In their review, Brennan et al. (2013)[39], considered studies that used utility based assessments of meta-health effects, reviewing 15 papers. In their systematic review, Higgins et al. (2014)[93] reviewed 27 studies that provided an assessment of the value of convenience for use in economic evaluations. This review became available in the course of preparing this PhD, and while it is concerned with aspects of the same topic area presented in this chapter, it focuses only on the valuation of convenience.

Excluding the systematic reviews reduced the count of included studies, many of which were also identified in the excluded systematic reviews, to 56 primary studies

considered relevant for inclusion. An ancestry search encompassing the reference lists of the aforementioned reviews as well as those of other included lead to the inclusion of an additional 14 papers.[38,62-66,91,104-112] This increased the number of included studies to 70. Neither the primary literature search nor the ancestry search identified a previously known paper by Prosser et al. (2003)[113] which utilised SG to assess the impact of the mode of treatment administration on quality of life in patients with multiple sclerosis (MS). This paper was therefore retrieved manually for inclusion, taking the total number of papers reviewed to 71. The full list of 71 studies and a summary of the data extracted is provided in Appendix 1.

Figure 3: Selection of studies for the systematic review



Note: The following exclusion criteria were applied to citations retrieved in the search:
A = Publication did not include a direct meta-health effect.
B = Economic evaluation or DCE without a meta-health effect.

C = Letter/commentary/news item/or editorial that lacked sufficient information/detail for further analysis.
D = Clinical or 'preferences' study that did not include valuation of meta-health effects (e.g. a simple rating study or statement of preference for one intervention over another).
E = Publication superseded by a later update; or reproduced the results of a previous publication.
F = Systematic reviews. Excluded from the primary analysis but included in the discussion and reference lists were searched for relevant references as part of the ancestry search.
G = Study did not estimate values for use in an economic evaluation despite the studies underlying method (e.g. cost was included as an attribute in the DCE but WTP was not estimated).
H = Publication did not report values in a way that could be used in an economic evaluation despite underlying method.
I = DCE analysis that did not include a cost, survival or risk attribute.

2.3.3.1.1 Classification by evaluation and intervention

The use of the various evaluation methods by intervention investigated across the studies is reported in Table 4. More than half (55%) of the included studies used DCEs or CA to investigate the value of meta-health effects. In reality, the majority of these can be considered to be DCEs, as the term CA is often used to describe what are actually stated preference DCEs.[114] Utility based methods (TTO and SG) represent the next most widely used method, occurring in 15 (21%) studies. Meta-health effects were most widely investigated in relation to drug interventions (35% of studies), with screening and studies with multiple interventions occurring the next most frequently. The link between the intervention investigated and the type of meta-health effect is discussed below.

Table 4: Studies by evaluation and intervention

Intervention	**DCE** (n=28)	**CA** (n=11)	**CV** (n=10)	**TTO/SG** (n=15)	**MAUI** (n=1)	**Multiple** (n=6)
Drug Interventions (n=25)	10[90,115-123]	3[110,124,125]	2[73,126]	7[62-64,91,107,113,127]	1[38]	2[109,128]
Screening (genetic, health) (n=10)	4[129-132]		3[72,105,133]	3[65,104,106]		
Multiple Types of Service Interventions (n=10)	5[134-138]	3[139-141]	1[20]			1[142]
Surgery (n=5)		3[143-145]	1[146]	1[108]		
Obstetrics (n=5)	2[147,148]	1[149]				2[150,151]
Diagnostic Services (n=6)	2[112,152]		1[153]	3[66-68]		
Health Promotion (n=2)			1[71]	1[99]		
Population Health (n=3)	2[154,155]		1[70]			
Devices (n=2)	2[101,156]					
Primary care (n=1)	1[157]					
Telemedicine (n=1)		1[158]				
Vaccine (n=1)						1[111]

Note: Numbers in superscript are the citations for each study.
Abbreviations: CA, conjoint analysis; CV, contingent valuation; DCE, discrete choice experiment; MAUI, multi-attribute utility instrument; TTO/SG, time-trade-off or standard gamble.

2.3.3.1.2 Categorisation of meta-health effects

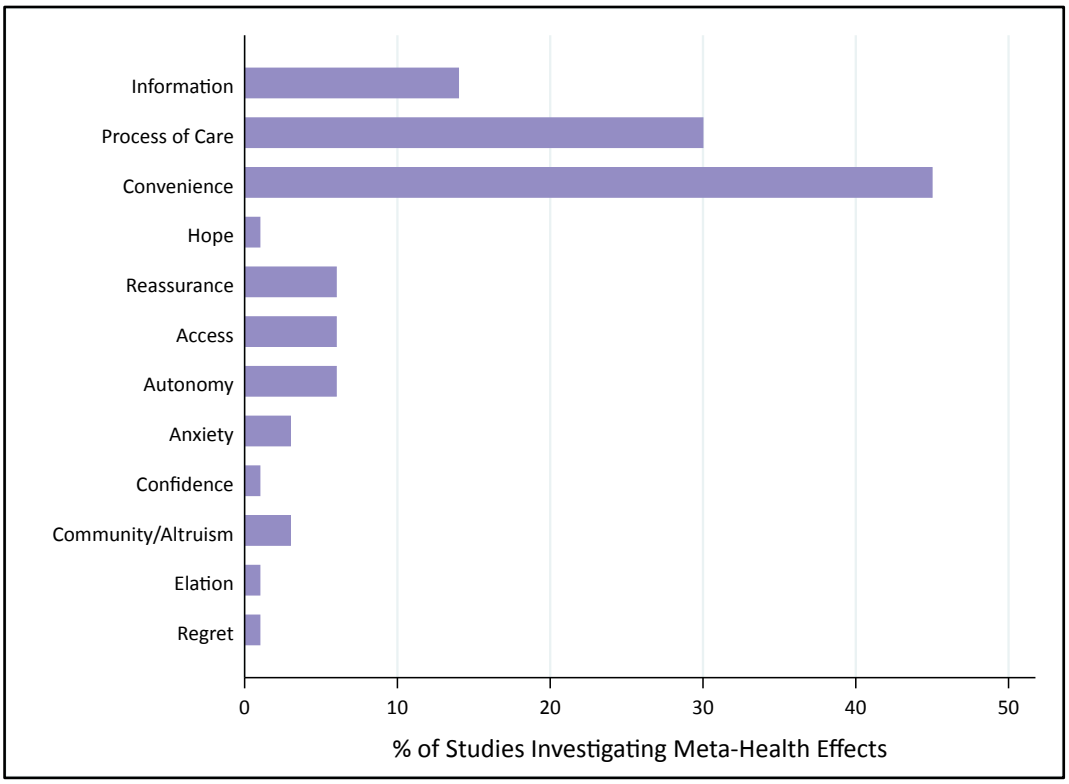
The classification of meta-health effects is presented in Table 5. For a given meta-health effect, the frequencies within this table are interpreted as unique counts per study, rather than counts of the number of times the term appears within studies. As such, they represent the number of studies in which those meta-health effects appear as catalogued in this review. One study might contain more than one meta-health effect such that the total number of meta-health effects cited (n=94) is greater than the number of included studies (n=57).

Table 5: Frequency of meta-health effects

Concept	Number of Studies
Meta-health effects as classified by this review	
Information	14
Process of care	28
Convenience/Acceptability	43
Hope	1
Reassurance/assurance	6
Access	6
Autonomy	6
Anxiety	2
Meta-health effects specifically classified within their source paper (and not allocated to a category within this review)	
Confidence	1
Community/Altruism	3
Elation	1
Regret	1
Total occurrences of meta-health effects	**112**

The most commonly included meta-health effect was convenience (n=43), followed by process of care (n=28), as shown in Figure 4. It is likely that this reflects the fact that these two concepts can be most readily researched and investigated. In addition, changes in mode of administration and delivery systems in drug treatment are cited as offering potential advantages to patients, largely due to convenience gains.[62-64,90-92,159] It is not surprising therefore that the number of studies dealing with the assessment of convenience is the highest (24 of the 25 studies dealing with drug therapies contained convenience as the meta-health effect of interest).

Figure 4: Occurrences of meta-health effects



Pragmatically, the six remaining meta-health effects are perhaps the most relevant, dominated by convenience and process of care. Arguably, process of care is a broad meta-health effect that is likely to encompass other meta-health effects. For example, elation and regret were meta-health effects included by Salkeld et al. (2004)[99] in their assessment of the value of hip protectors in preventing fractures in women at risk of falls. Within their analysis, these separate concepts of elation and regret were seen as part of the broader overall concept of process utility.[99] Therefore, some meta-health effects might be better included under a broader heading of process of care lest classifying them into increasingly smaller categories makes them appear to be of lesser value.

The classification of the meta-health effects includes four concepts not outlined in the initial methods but categorised as meta-health effects by the study authors: confidence, community/altruism, elation and regret. Confidence has been included as

a potential meta-health effect arising from the health promotion intervention evaluated by Borghi and Jan (2008).[71] Within that intervention, women and their partners participated in group-education sessions designed to increase awareness and improve pre-natal health, disseminate information and promote confidence. Participation in those groups and the provision of that information has been classified as affording a meta-health effect in the form of confidence (as well as information and altruism). In a similar health promotion study by Alayli-Goebbels et al. (2013)[10], confidence was also identified as a meta-health effect. However, within that study, changes in confidence (measured by the authors as changes in body image) relied on a change in health state and were a direct result of participating in the intervention itself. Accordingly, it violated the definition of a direct meta-health effect and was excluded from further review. Elation and regret were both concepts separately identified by Salkeld et al. (2004), while community/altruism were included by Borghi and Jan (2008), and Dixon and Shackley (1999).[70,71,99]

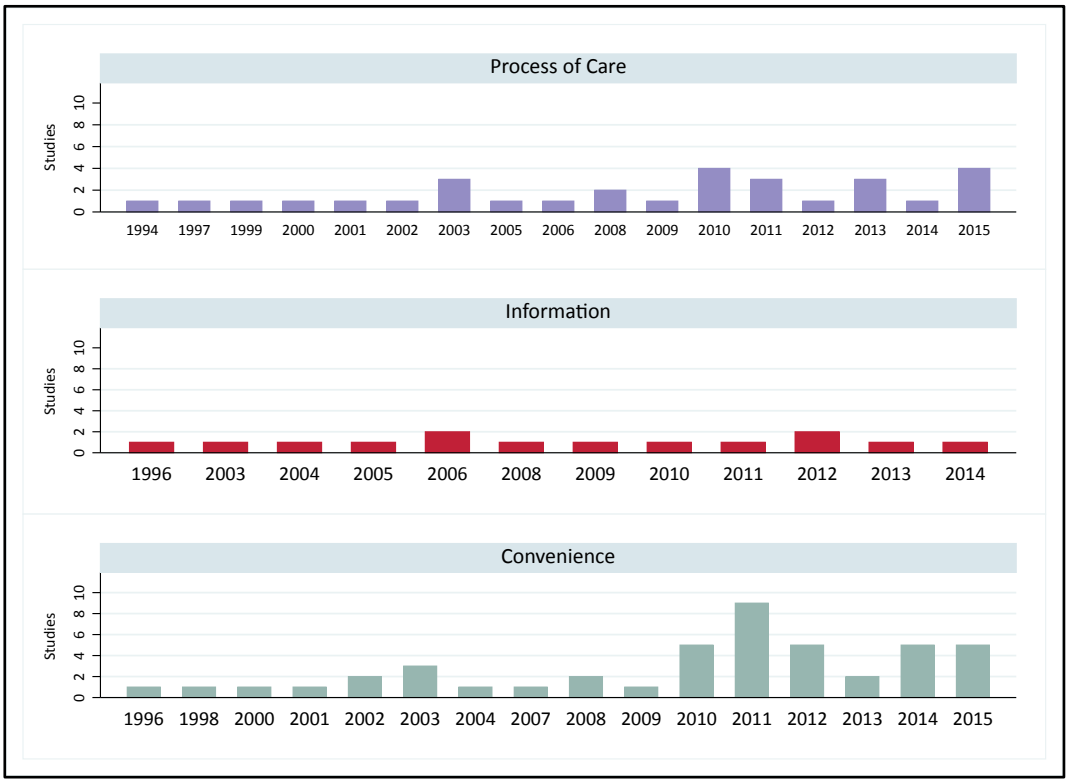
For the remaining analyses in this section, the association of the meta-health effect investigated with the year of publication, evaluation type, intervention, basis for the study and source of funds was assessed only for the three most commonly investigated meta-health effects: convenience, information and process of care. The incidence of the remaining meta-health effects was too low as to be informative for further investigation.

Meta-Health Effects by Year

The number of studies by year in which convenience, information and process of care were investigated is presented in Figure 5. These results show that publication of studies investigating convenience was skewed toward more recent years peaking in 2011 (n=9), with 31 of the 45 studies investigating convenience being published after 2010. The 10 publications in 2011 were spread over a range of interventions and evaluation methods used (although six were DCEs).[91,115-118,128,131,142,158] More recently, there appears to have been a resurgence in interest in meta-health effects with six

publications in 2015, encompassing both convenience and process of care.[121,127,138,151,153,156] While there were fewer publications that included process of care as a meta-health effect, they were more evenly distributed (peaking with four in 2010, one more than in each of 2003 and 2011). Publications including information were fairly evenly distributed over the years examined.

Figure 5: Meta-health effects by publication year



Results of χ^2 tests of association between year of publication and whether the meta-health effects were included (as a binary outcome) were 26.84 ($p = 0.14$; Cramer's V = 0.61) for convenience, 20.85 ($p = 0.41$; Cramer's V = 0.54) for process of care, and 16.10 ($p = 0.71$; Cramer's V = 0.48) for information.[xi] Despite the apparent growth in the

[xi] Significance for Cramer's V is the same as that for χ^2. Interpretation of Cramer's V depends on the d.f., smallest of one minus the number of rows or columns. For d.f.=1, thresholds for association are V=0.1 is weak association, 0.3 is medium, 0.5 is strong. For d.f. = 2, V=0.07 is weak association, 0.21 is medium, 0.35 is strong. For d.f.=3, V=0.06 is weak association, 0.17 is medium, 0.29 is strong.[160]

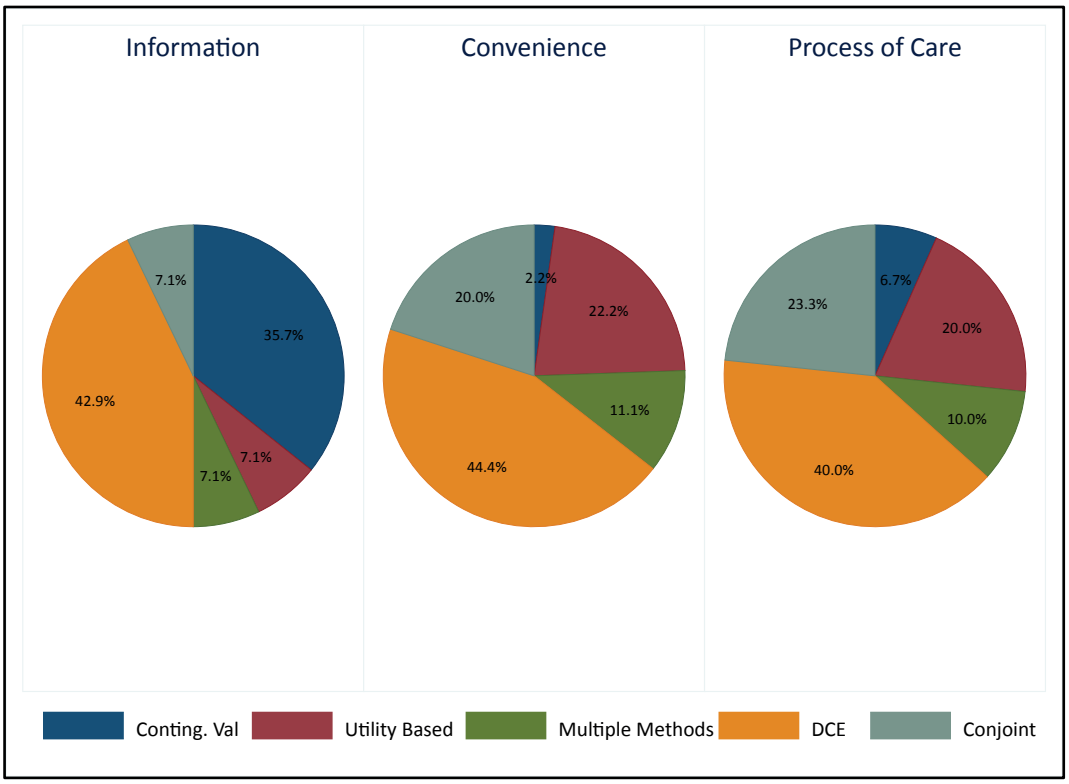
number of publications investigating meta-health effects, particularly convenience, after 2010, the results show no statistically significant association between year of publication and the meta-health effect investigated. The same results were observed collapsing the years of publication into four eras: pre-1994, 1994-2000, 2000-2010, post-2010.

Meta-Health Effects by Evaluation Type

The proportion of studies accounted for by the respective evaluation types is depicted in Figure 6. DCEs were the most widely used method when considering convenience and process, and jointly (with CV) for process of care as a meta-health effect. Interestingly, CV was used least with respect to convenience or process of care, but the most with respect to information as a meta-health effect. The results show a significant and strong association (indicated by the Cramer's V) between convenience and the type of evaluation $\chi^2 = 14.93$ ($p < 0.01$; Cramer's V = 0.46). However, there was no association shown for process of care ($\chi^2 = 4.13$, $p = 0.39$; Cramer's V = 0.24) or information ($\chi^2 = 8.37$, $p = 0.08$; Cramer's V = 0.34). This result is surprising given the apparent variability in evaluation types used across those meta-health effects.

As noted above, there appear be more studies published post-2010 that investigated meta-health effects, particularly convenience. This might reflect the timing of new health care products becoming available, or the more widespread use of evaluation methods, such as DCEs, that are better suited to the assessment of meta-health effects. Publications of DCEs peak in the years after 2010. The χ^2 test of association between evaluation type and year of publication (classified into pre-2000, 2000-09, or 2010 and later) was 19.86 ($p = 0.01$; Cramer's V = 0.37), indicating some association.

Figure 6: Meta-health effects by evaluation type



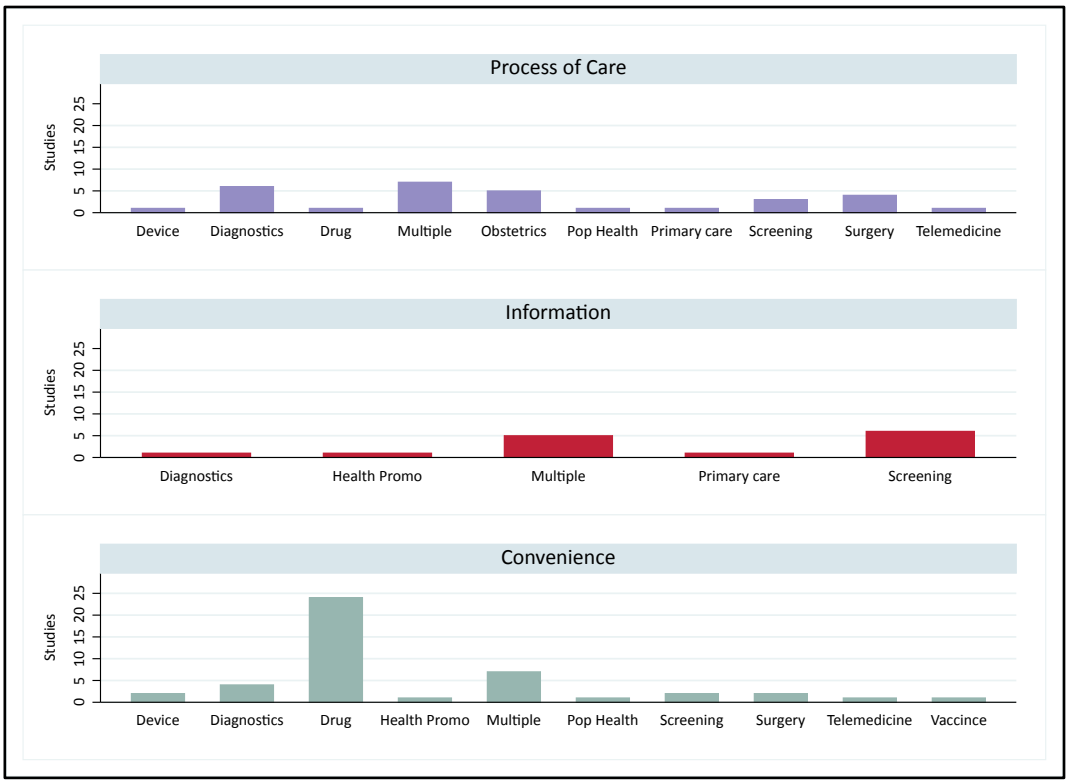
Abbreviation: DCE, discrete choice experiment.

Meta-Health Effects by Intervention

The count by intervention for each meta-health effect is depicted in Figure 7. The most striking feature of these data is the number of studies investigating convenience as a meta-health effect that were focused on drug interventions; of the 45 studies that considered convenience 24 were drug studies. As might be expected, screening interventions, for which the primary outcome is the detection of disease or disease pre-cursors, were most often associated with the investigation of information as a meta-health effect.

Results of χ^2 tests of association between intervention and whether the meta-health effects were included (as a binary outcome) were 33.53 ($p < 0.001$; Cramer's V = 0.69) for convenience, 38.69 ($p < 0.001$; Cramer's V = 0.74) for process of care, and 34.00 ($p < 0.001$; Cramer's V = 0.69) for information. These results indicate a strong association between the intervention investigated and the meta-health effect assessed.

Figure 7: Meta-health effects by intervention



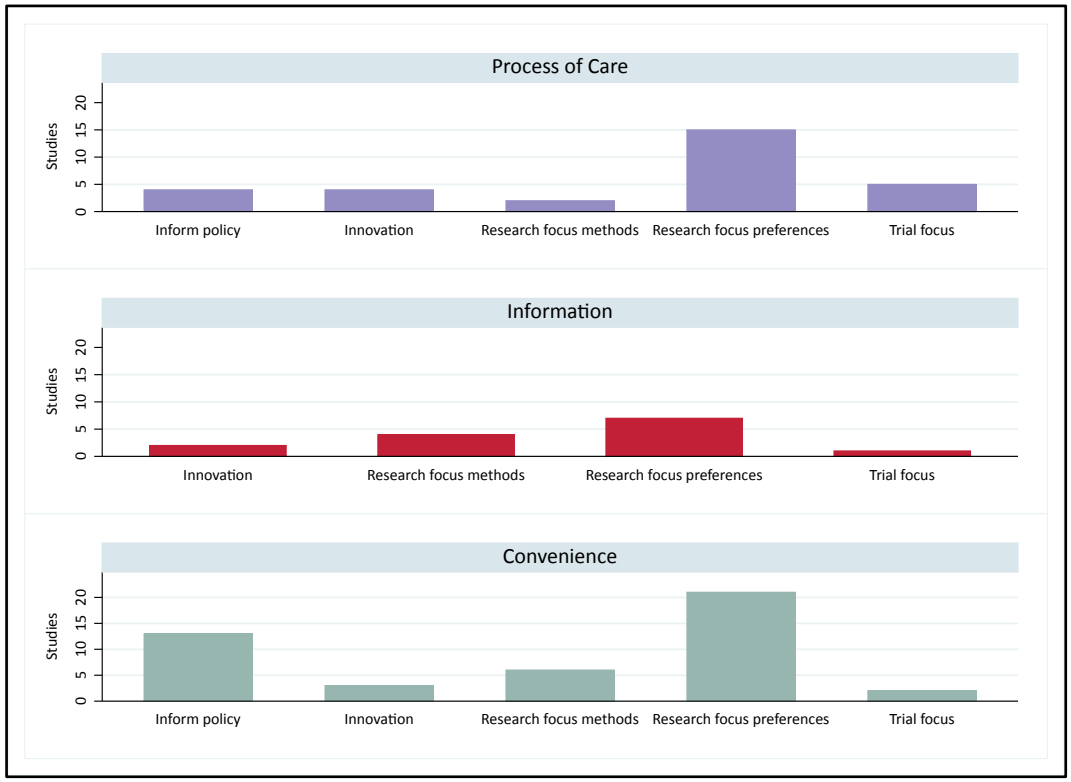
Basis for the Study and Source of Funds

In the majority of studies, the reason for the study was a research focus – preferences (45%), followed by informing policy (21%). Reporting of the reason for the study by the three most common meta-health effects is provided in Figure 8. These results show that for all three meta-health effects, the most common reason for the study comparison was a research focus on the elicitation of preferences. For studies in which convenience was investigated, informing policy (including estimating values for use in reimbursement applications) was the next most common reason for conducting the study.

However, there does not appear to be an association between the meta-health effects and the basis for the comparison in these studies. Results of tests of association between the basis of the comparison in each study and whether the meta-health effects were included (as a binary outcome) were significant for convenience (χ^2=9.45; p = 0.05;

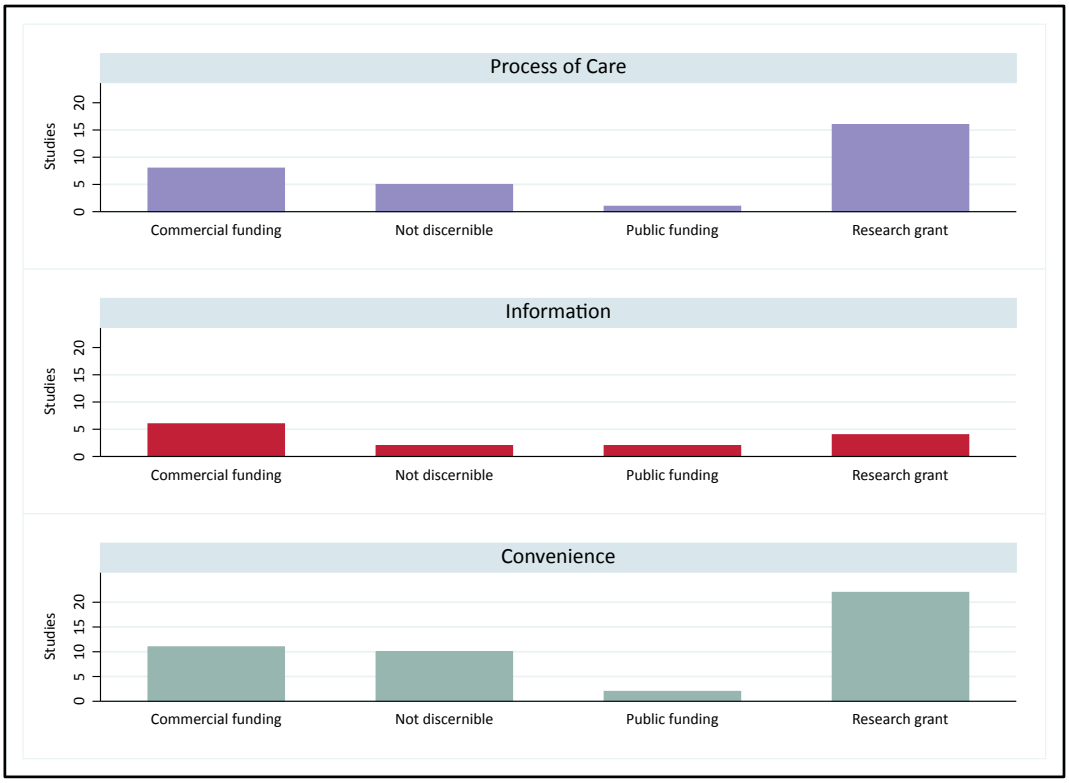
Cramer's V = 0.37), but not for process of care (χ^2=7.42; p = 0.12; Cramer's V = 0.32), nor information (χ^2=7.33; p = 0.12; Cramer's V = 0.32). This perhaps reflects the prevalence of drug studies being associated with investigating convenience, and the use of such studies to attempt to influence product reimbursement.

Figure 8: Reasons to investigate meta-health effects



A breakdown of the source of funding by the three most common meta-health effects investigated is provided in Figure 9. The majority of studies were funded by research grants (46%), with commercial funding being the next most common (28%). Research grants dominate for process of care and convenience, whereas commercial funding was the most common source of funds where information was investigated as a meta-health effect. Results of χ^2 tests of association between the source of funds and whether the meta-health effects were included (as a binary outcome) were not significant: 0.84 (p = 0.84; Cramer's V = 0.11) for convenience, 1.15 (p = 0.76; Cramer's V = 0.13) for process of care, and 7.10 (p = 0.07; Cramer's V = 0.32) for information.

Figure 9: Source of funding



2.3.3.2 *What values were assigned to meta-health effects?*

Of the 71 included papers, 41 reported values that could be used in an economic evaluation.[20,62-68,70-73,90,91,99,104-109,111-113,115,120,122,126-129,131,133,134,142,146,149,151,152,156,158] The remaining 30 papers included meta-health effects but either did not produce, or did not report, values that could be incorporated into an economic evaluation, despite using a valuation method appropriate for that purpose (e.g. DCEs that included a cost attribute).[38,101,110,116-119,121,123-125,130,132,135-141,143-145,147,148,150,153-155,157]

Results for each valuation exercise, and the implications therein for the assessment of meta-health effects, were extracted from the 41 studies that reported values that could be used in an economic evaluation. This information, including a breakdown of studies by assessment method in terms of whether or not meta-health effects were found to be significant, is summarised in Table 6. Among the 41 studies, the contribution of the meta-health effects to overall value was reported separately in 32.

Within those 32 studies, the value reported for the meta-health effect was found to be statistically significant in 27 (84%), and not significant in five (16%). In the remaining nine of the 41 studies, a separate value for the meta-health effect could not be discerned from the overall value reported (either the overall WTP or utility value).

It is worth considering whether there is something particular to the design or implementation of those five studies in which the meta-health effect was found to be not statistically significant:

- Yasunaga et al. (2006)[133] – a CV study examining WTP for prostate cancer screening, and for which the meta-health effect of interest was information. In the analysis, the level of information provided was tested as a covariate in a regression of WTP and was found to not be significant. In this case, respondents did not evaluate information directly, but its impact was subsequently tested in the analysis.
- Palumbo et al. (2011)[142] - this study included a CV and DCE component assessing preferences for assisted reproductive technologies, for which the meta-health effects of interest were information, process of care and convenience. Within the CV element of the study, approximately 2/3 of respondents reported that they were not willing to pay anything for convenience factors (comfort and tolerance) associated with the treatment. This is in stark contrast to high WTP amounts reported for efficacy dimensions, and suggests that efficacy dominates convenience in the context of assisted reproductive technologies.
- Donaldson and Shackley (1997)[146] – a CV study assessing surgical techniques for cholelithiasis (disease of the gall bladder) for which process of care was the meta-health effect of interest. The scenario that included process effects produced a lower WTP than that with only the outcome effects. The authors were surprised by this finding and suggest that rather than being evidence that respondents do not value process of care effects, it is perhaps a consequence of the manner in which the process was described.

- Prosser et al. (2003)[113] – a SG in which patients and members of the community valued the mode and frequency of administering treatment for MS. While the authors did not report the between frequency differences for utility, this was estimated using the available information (s.d. and sample size) and compared on the basis of confidence intervals. All confidence intervals for the means overlapped (despite the non-normal nature of these data, the extent of overlap is such that the values produced are not different).
- Yee et al. (2015)[151] – incorporated both a SG and TTO to the valuation of women's preferred approach to obstetric delivery. Women assessed 14 health states that were all planned, but differed in terms of complication and birth outcome. TTO utility estimates for comparable health states (e.g. planned vaginal birth compared on the basis of the reported mean and s.d.. These estimates did not differ.

Of the five studies without a significant value for meta-health effects, three were CV studies, the other two utility studies. This might suggest that some study types are more likely to be associated with significant values for meta-health effects than others. Consider the simple frequencies reported in Table 6. As previously noted, 84% of studies in which meta-health effects were valued separately found them to be significant. The corresponding proportions by study type were 63% for CV, 50% for multiple methods, 91% for TTO/SG and 100% for DCE[xii] and CA studies. When considering all studies of a given type (including those for which the value of the meta-health effect could not be discerned), these percentages were 45% for CV, 50% for multiple methods, 63% for TTO/SG, 90% for DCE and 100% for CA.

[xii] In classifying whether or not a DCE delivered a significant value for a meta-health effect, it was sufficient that one significant meta-health effect attribute was identified even if others were not significant.

Table 6: Reporting of values for meta-health effects

	CV	Direct Elicitation: TTO/SG	Multiple Methods	CA	DCE	Total
Meta-health Effects valued separately: significant[20,62-68,71-73,91,99,105,111,112,115,120,122,127,129,131,134,149,152,156,158]	5	10	1	2	9	**27**
Meta-health Effects valued separately: not significant[113,133,142,146,151]	2	1	2	-	-	**5**
Meta-health Effects subsumed in overall value only[70,90,104,106-109,126,128]	2	4	2	-	1	**9**
Total	**9**	**15**	**5**	**2**	**10**	**41**

Abbreviations: CA, conjoint analysis; CV, contingent valuation; DCE, discrete choice experiment; SG, standard gamble; TTO, time-trade-off.

2.3.3.3 Exploring differences in meta-health effect values.

2.3.3.3.1 Non-significant values for meta-health effects

It is worth exploring the reasons why Prosser et al. (2003)[113], Yee et al. (2015)[151], Yasunaga et al. (2006)[133], Donaldson and Shackley (1997)[146], and Palumbo et al. (2011)[142] reported a separate but non-significant value for their meta-health effects. The absence of difference in Prosser et al. (2003) [113], is potentially due to the small number of respondents per treatment state (approximately 20), the inclusion of patients accustomed to treatment management and the purposive sampling frame (all community respondents were recruited from within the same apartment building). The suggestion that the inclusion of patients in the study might diminish the impact of the meta-health effects might support claims that such effects induce a 'shock' response when valued by members of the public. Granted, Prosser et al. (2003)[113] did not observe statistically significant differences between patient and community member values within health states. Within Yee et al. (2015)[151], the absence of a significant value for process of care effects might arise due to the way in which the valuation tasks were framed; within the TTO women traded off against their life expectancy given their preferred mode of obstetric delivery, rather than full health. This would be anticipated to vary between women, potentially giving rise to the large observed

standard deviations in reported values. In addition, review of the underlying vignettes for comparable health states, such as planned vaginal delivery without complication and planned Caesarean delivery without complication, showed that they did not describe the same aspects of care. For example, the need for post-delivery analgesia was mentioned for Caesarean delivery but was not included in vaginal delivery.[151] The presence of such differences might have influenced how respondents framed the elicitation task, and thus the comparability of the health states within that study.

While not explicitly stated as such, Yasunaga et al. (2006)[133] and Donaldson and Shackley (1997)[146] appear to have tested framing effects in terms of how information was presented rather than testing the value of the meta-health effects themselves. Yasunaga et al. (2006)[133] conducted a CV study examining WTP for prostate cancer screening in which the meta-health effect of interest was information. In the analysis, the level of information provided was tested as a covariate in a regression of WTP and was found not to be significant. In this case, respondents did not evaluate information directly, but information effects were tested in a subsequent analysis.[133] Similarly, Donaldson and Shackley (1997)[146] conducted a CV study assessing surgical techniques for cholelithiasis in which process of care was the meta-health effect of interest. The scenario that included process of care effects produced a lower WTP than that which included only health outcome effects, but the authors hypothesise that this was due to how the information was described.[146] Finally, the analysis by Palumbo et al. (2011)[142] suggests that efficacy dominates convenience in the context of assisted reproductive technologies. This study included a CV and DCE component assessing preferences for assisted reproductive technologies in which the meta-health effects of interest were information, process of care and convenience. Within the CV element of the study, approximately 2/3 of respondents reported that they were not willing to pay anything for convenience factors (comfort and tolerance) associated with the treatment, but the WTP values for efficacy were very high (values not reported).[142]

2.3.3.3.2 Significant values for meta-health effects

The 27 studies for which a statistically significant value for a meta-health effect was reported separately were investigated further for potential insights into whether some meta-health effects produced higher values than others, and what might have produced different values for similar meta-health effects. This investigation was tempered by a number of factors:

1. The meta-health effects explored were largely classified into the categories used for the purposes of this review. The researchers responsible for the original studies might not accept that the meta-health effect categories to which they have been assigned capture the same aspects of patient utility or QoL they had intended.
2. The studies vary in their country and health care settings. This means that the context in which the meta-health effects are investigated, and the units in which some (expressed as WTP) are reported are not comparable as absolutes, but potentially only relative to a base value.
3. The methods varied across studies such that in some, the primary focus of the valuation exercise was meta-health effects (excluding other health outcomes from the valuation), whereas in others they were part of an overall suite of factors valued. This is likely to have impacted on the relative values obtained within and across studies.

In these 27 studies, it was not always possible to assess the relative contribution of the meta-health effect to total value. This arose in the context of WTP estimates (or in DCEs where value was assessed using WTP) for which an overall mean WTP was not reported. Thus while WTP might have been reported for the meta-health effect, or stratified by particular patient characteristics, it was not possible to compare the reported WTP for a meta-health effect to a base mean WTP (hence a relative value for the meta-health effect could not be formed). This occurred for 11 studies[73,111,112,115,120,122,131,149,152,156,158], leaving 16 studies for which the values for the meta-health effect could be compared with a base value.

The relative measures for the meta-health effects in the 16 studies for which they could be formed are presented in Table 7. In each instance, the measure of interest is the contribution of the meta-health effect to the reported measure of value (either utility or WTP) expressed as a percentage. For example, Birch et al. (2003)[65] used SG to assess the utility associated with the process of care for cervical cancer lesions. For that study, the overall utility associated with the health state in which women underwent cryotherapy was 0.947; the process of care itself was attributed with a utility value of 0.021 in absolute terms, producing a relative value of 2.22% (0.021/0.947).[65]

Table 7: Relative contributions of meta-health effects

Study	Meta-health Effects	Base value	Meta-health Effects Contribution	
			Absolute	Relative (%)
WTP Measures				
Borghi and Jan (2008)[71]	Information, community, confidence: male respondents.	NPR449	262	58.35
	Information, community, confidence: females not at group.	NPR449	177	39.42
	Information, community, confidence: females at group.	NPR449	106	23.61
Protiere et al (2004)[20]	Information (heart disease)	FRF339	158	46.61
	Information (helicopter ambulance)	FRF251	97	38.65
	Information (breast cancer)	FRF306	98	32.03
Philips et al (2006)[105]	Reassurance (negative information)	GBP239	76	31.80
	Reassurance (positive information)	GBP239	70	29.29
Neumann et al (2012)[72]	Information (10% disease risk, Alzheimer's)	USD479	70	14.61
	Information (25% disease risk, prostate cancer)	USD622	76	12.22
Naik-Panvelkar et al (2012)[134]	Information (Advice)	n.a.	22.8	24.04
	Process (Privacy)	n.a.	18.38	19.38
	Process (Access)	AUD94.86	9.18	9.68
Chan et al (2009)[129]	Information	GBP854.74	267.82	31.33
	Anxiety	GBP 854.74	14.39	1.68

Study	Meta-health Effects	Base value	Meta-health Effects Contribution	
			Absolute	Relative (%)
Utility Measures				
Salkeld et al (2004)[99]	Regret	0.31	-0.21	-67.74
	Elation/ Reassurance	0.63	0.32	50.79
	Elation	0.71	0.07	9.86
Johnson et al (1996)[62]	Convenience	0.732	0.223	30.36
Osborne et al (2007)[64]	Convenience	0.85	0.23	27.06
Osborne et al (2012)[63]	Convenience (2 wk-3 mth)	0.7	0.098	14.00
	Convenience (2-4 wk)	0.65	0.047	7.23
Swan et al (2003)[68]	Process & Anxiety (Quality adjusted days)	7	5	71.43
Swan et al (2000)[66]	Process & Anxiety (Quality adjusted days)	15.2	9.4	61.84
Swan et al (2006)[67]	Process & Anxiety (Quality adjusted days)	5.5	2.4	43.63
Matza et al (2015)[127]	Convenience***	0.77	-0.07	-9.10
Birch et al (2003)[65]	Process (Immediate care – no pathology)	0.958	-0.031	-3.24
	Process (Immediate care – cryotherapy)	0.947	0.021	2.22
	Process (Immediate care – cone biopsy)	0.922	0.017	1.84
Boye et al (2011)[91]	Convenience (frequency)*	0.73	0.024	3.29
	Convenience (flexibility)**	0.89	0.008	0.90

Notes: *Based on difference between daily and weekly dosing in the presence of injection site reactions with flexible dosing.
**Based on difference between flexible and inflexible dosing without injection site reactions and once per week dosing.
***Matza et al (2015) evaluated multiple possible administration regimens. Utility difference shown is on seven tablets per day plus a weekly injection, compared with 18 tablets per day plus a weekly injection.
Studies are sorted from highest to lowest by the absolute magnitude of the relative value (based on the highest relative value and lowest relative value within any given study).

Abbreviations: AUD, Australian dollars; FRF, French francs; GBP, Great Britain pounds; NPR, Nepalese rupees; mth, month; USD, United States dollars; wk, week.

In general, higher relative values were produced for meta-health effects assessed via WTP, as compared with TTO/SG based methods. Of the six studies that produced WTP values, four used CV while two (Chan et al. (2009), Naik-Panvelkar et al. (2012)) used DCEs.[20,71,72,105,129,134] Within those WTP studies, the relative value derived for a given meta-health effect (such as information) appears to be context specific, varying across studies, and within studies across the health conditions considered. For example, within Protiere et al. (2004)[20] the value of information was assessed using

WTP for three different health care programmes: helicopter ambulance, heart disease and breast cancer. The relative value of information was 38.65%, 46.61% and 32.03% respectively across these three programmes. In contrast, the relative value of information with respect to prostate cancer screening in Neumann et al. (2012)[72] was 12.22% of the mean WTP. These studies are described more fully below.

Within the studies that applied a utility valuation method, the highest relative values were produced by Salkeld et al. (2004)[99] in assessing regret and elation/reassurance in valuing hip protectors in a falls prevention study. This study produced relative values of 67.74% (a disutility) for regret and 50.79% for elation/reassurance. The other relative values of note because they were comparatively high were those achieved in the studies by Johnson et al. (1996)[62] and Osborne et al. (2007)[64] for convenience of 30.46% and 27.06% respectively. Both assessed the value in moving to an oral therapy from an intravenous (IV) therapy; the former for ongoing maintenance treatment for cytomegalovirus (CMV) in patients with AIDS, the latter in paediatric patients requiring daily oral iron chelation therapy. The lowest relative value for convenience (and any meta-health effect) was 0.9% from the study by Boye et al. (2011)[91] comparing drug administration modalities in patients with diabetes.

There was variation noted in the manner in which values were derived across the utility studies. In particular, all three Swan et al. studies assessed the value of the meta-health effects (process and anxiety) using the waiting time approach, and reported it as disutility adjusted days rather than disutility values specifically.[66-68] The gains reported (and hence values for meta-health effects) therefore include the difference in waiting time (for results of the relevant test) associated with the procedures investigated. The relative values reported for these studies might therefore be inflated depending on the combination of the effect of the difference in waiting time (days) and associated utility values. In addition, while the authors propose that the studies assess a value for the testing procedures[66-68], the method used (in terms of

trading off waiting time for access to information for diagnostic tests) might actually reflect the value individuals place on the results produced by those tests.

2.3.3.3.3 What were the potential causes of differences in relative values?

Potential sources of differences for relative values for meta-health were assessed using a cross-study comparison, controlling for the valuation metric (WTP or utility value) and meta-health effect (information and convenience). Study design, analysis methods and the approach to the assessment of the meta-health effect were compared for two studies, each for information or convenience using WTP or utility values respectively:

1. Protiere et al. (2004)[20] and Neumann et al. (2012)[72] were compared as the studies with the highest and lowest relative values respectively for information using WTP as the valuation method.
2. Johnson et al. (1996)[62] and Boye et al. (2011)[91] were compared as the studies with the highest and lowest relative values respectively for convenience using a utility measure.

The exploration of potential sources of difference between the two studies for each of the meta-health effects (convenience and information) is presented in Table 8.

Table 8: Sources of difference in relative values of meta-health effects

	Protiere et al (2004)[20]	Neumann et al (2012)[72]	Johnson et al (1996)[62]	Boye et al (2011)[91]
Measurement type	WTP	WTP	Utility	Utility
Relative value ranking	High Case	Low Case	High Case	Low Case
Meta-health Effect assessed	Information	Information	Convenience	Convenience
Relative value	Range from 32.03% (breast cancer) to 46.61% (heart disease).	Range from 12.2% (prostate cancer) to 14.61% (Alzheimer's disease).	30.46%	Range from 0.90% (flexibility) to 3.29% (frequency).
Evaluation method	Contingent valuation	Contingent valuation	TTO	SG
Efficacy/safety mentioned in vignette?	Within each programme such information on efficacy (or the disease risk) was provided. The most comprehensive was for cardiac surgery (for all information groups).	Within each scenario, disease risk was stated. In addition, "information" was couched in terms of test accuracy.	Efficacy was described as the time until the need for a 21 day IV treatment (on average: 62 days for IV, 57 for oral). Vision effects the same. Chance of IV line infection resulting in hospitalisation included (shown as positive for oral). Adverse effects (resulting in hospitalisation) were included (see text for discussion).	Described diabetes symptom control as part of the base health state. Injection site reactions were noted in some of the vignettes (where relevant).
Was the meta-health effect the only variable factor?	Information varied for the heart disease health state only (three levels).	Disease risk varied between scenarios (10% and 25%), as did the level of test accuracy (perfect of imperfect).	No. Adverse effects, and time to treatment due to CMV worsening.	No. Injection site reactions also varied in the vignettes.
Analysis methods	Base WTP assessed using mean statistics. Relationship of WTP and explanatory variables described using seemingly unrelated regressions.	WTP assessed using survival analysis, based on double bounded bids. Logistic regressions and maximum likelihood regression used to assess mean and median values among those with positive WTP.	Utility values expressed as means and medians, and tested with non-parametric methods.	Utility values expressed as means, and differences between health states assessed with paired t-tests.
Condition of interest	Acute care retrieval; heart disease and breast cancer.	Multiple conditions (Alzheimer's/Arthritis/Prostate Cancer/Breast Cancer).	CMV (maintenance therapy).	Diabetes.

	Protiere et al (2004)[20]	Neumann et al (2012)[72]	Johnson et al (1996)[62]	Boye et al (2011)[91]
Intervention of interest	Helicopter ambulance; cardiac surgery; breast cancer treatment.	Diagnostic screening (predictive test).	Ganciclovir: four tablets t.i.d. compared with 90 minute IV administration per day.	Injections; comparing daily (once/twice) with weekly, and with and without time of day (dietary) restrictions.
Respondent group	Adult community sample (France).	Adult community sample (USA).	Adult sample (Australia) who were HIV^+, but had not developed CMV.	Diabetes patients (Scotland).

Abbreviations: CMV, cytomegalovirus; HIV, human immunodeficiency virus; IV, intravenous; SG, standard gamble; t.i.d., three times daily; TTO, time-trade-off; WTP, willingness to pay.

Within Protiere et al. (2004)[20], participants valued three different health care programs – one for breast cancer, cardiac surgery, and a helicopter ambulance. Participants were assigned into one of three groups based on the level of information they saw for cardiac surgery, which the authors describe as being increasingly positive. Despite information only differing across the groups for cardiac surgery, the authors found a positive value for information across all three interventions.[20] The relative values obtained for Neumann et al. (2012)[72] were lower than those in Protiere et al. (2004).[20] This is surprising given that the latter explored WTP in the field of screening in which information is considered to be an important source of value.[13] However, despite the differences in the underlying conditions and interventions being valued, there is a key distinction between these studies that might explain the apparent difference in the observed relative values for information; the manner in which information itself is presented within the respective vignettes.

Protiere et al. (2004)[20] test the value of information by changing how much and the type of information that is shown for cardiac surgery across all three groups (described above). Differences in the resulting WTP values between the groups are therefore taken as an indication of the value of information. In this sense, the information is exogenous to the intervention/condition under evaluation. In the case of Neumann et al. (2012)[72], the information is made operational in terms of the accuracy of the screening tests themselves (perfect or imperfect test accuracy); it is therefore endogenous to the intervention and has direct implications for the resulting health status of the individual. Therefore, considerations of those health impacts (in terms of the implications of the accuracy of the testing) might have carried greater impact in determining WTP than the intrinsic value of the information itself. In addition, the relative values achieved for information in Neumann et al. (2012) appear to be insensitive to whether or not the disease risk was 10% or 25% (i.e. the relative value for Alzheimer's disease was 14.6% with a disease risk of 10% and 14.4% with a disease risk of 25%; and a similar pattern was observed for the other conditions investigated).

Within the utility based studies, the largest relative value when assessing convenience was observed in the study by Johnson et al. (1996)[62]; a difference of 30.46%. In contrast, the largest difference observed by Boye et al. (2011)[91], who also assessed convenience, was 3.29%. These studies differ in their valuation method; the former is a TTO, the latter a SG, and the two methods have been observed to produce different estimates of utility for the same health states (typically higher for SG).[161] However, in this instance, it is likely that the source of the difference is due to the conditions under investigation, and magnitude of the convenience gain, to wit:

1. Both studies used existing patients: Johnson et al. (1996)[62] used HIV patients who did not have AIDS, and Boye et al. (2011)[91] used patients with diabetes.[62,91] Participants in Boye et al. (2011)[91] might have reflected coping/adaptation in their responses depending on their prior exposure to similar treatments. While this might also have been the case for patients in Johnson et al. (1996)[62], the health states described a potential progression in their disease state which might have served to compound their interpretation of differences in the convenience gains presented (thereby conflating shock with the impact of the IV treatment compared with the oral).
2. In the health states described in Johnson et al. (1996)[62], the IV treatment is described as being administered for 90 minutes each day through a permanent IV line implanted in the chest, while the oral (four tablets three times a day) requires about 15 minutes per day. In contrast, in Boye et al. (2011)[91] the diabetes injections are described as administered once or twice per day (immediate) daily, or weekly; and either one hour prior to meals or at any time. In both cases, the interventions under investigation are life-long.

With the exception of differences in the valuation method (TTO compared with SG) and participant group sampled, the construction of the health states in Boye et al. (2011)[91] and Johnson et al. (1996)[62] appears to be consistent. The driver in this case would appear to be the magnitude of the gain in convenience. Shifting from 90 minutes a day to 15 minutes a day for CMV maintenance therapy in AIDS patients is

likely to constitute a significant convenience advantage, relative to moving from twice daily to once daily or weekly injections in the context of diabetes care.[62,91]

It is possible that part of the difference in the measured utility value within Johnson et al. (1996)[62] is due to the manner in which adverse effects have been framed. These were shown as the chance of an infection developing in the IV line requiring hospitalisation, and therefore occurred only for the IV administration route, and were presented as a positive for the oral "because no IV line is used…there is no chance it would become infected".[62] This positive framing of the side effects for the oral formulation might have biased the resulting utility values (although it is noted that a higher incidence of adverse events for the IV compared to the oral formulation was significant only for anaemia and in-line infections[162]).

The gain in utility associated with convenience estimated by Johnson et al. (1996)[62] also seems high given that the health state used in the study describes patients in both treatment groups (oral and IV) requiring upfront treatment for CMV using a 21 day intensive IV regimen of ganciclovir. This would be anticipated to diminish some of the convenience gain of the oral compared with the IV administration health states. Caution is therefore required in interpreting the utility value shown since it incorporates values for a period of no difference between the oral and IV formulations.

2.3.4 Discussion and Limitations

2.3.4.1 Discussion

The focus of the literature review has been on how meta-health effects have been assessed within existing publications that have reported values that can be used within an economic evaluation. A distinction has been made between studies that explored strength of preference and those that attempted to enumerate those preferences on a utility scale that can be used to construct QALYs, or in monetary terms for use in CBA.

The systematic review showed that values for convenience and the process of care are investigated most commonly. It is likely that this reflects the fact that these two meta-health effects can be most easily conceptualised and described for the purposes of research. It could be argued that the number of studies in which each meta-health effect appears could be used as a proxy of their potential importance. For example, meta-health effects such as anxiety, elation, regret, confidence, community/altruism and hope were cited only once or twice in the studies reviewed, suggesting that they are less important as a source of value. However, it might also be the case that the low instance of these concepts being investigated reflects the difficulty of doing so, rather than their inherent lack of value.

The dominance of convenience as the most commonly measured meta-health effect was highlighted in the systematic review by Higgins et al. (2014)[93] who identified a similar number of studies that investigated convenience as in the present literature review. As with the present review, they find that while a number of studies report values for convenience that are significant and can be used as inputs to an economic evaluation, there is a lack of consistency between studies in the methods used to derive and report those values. This complicates the inclusion of such values in economic evaluations and therefore their consideration in reimbursement decisions.[93]

This review has a broader focus from that of Higgins et al. (2014).[93] It has included meta-health effects other than convenience. Within their review, Opmeer et al. (2010)[5] also apply a broad definition of meta-health effects, including: features of the intervention (invasive); process (discomfort, duration, health care professionals); and being subjective or having a perceived impact on patients. The definition is perhaps broader than that used in the current review which focuses on direct meta-health effects only (those which do not require a change in the patient's health status to take effect). As with the current review, Opmeer et al. (2010)[5] focus on the methods of preference elicitation, the five used most often being: binary choice (A vs B) – 49%;

judgement of separate components/attributes of an intervention/programme (this technique was not explained) – 15%; trade-off based methods – 13%; DCEs – 8%; rankings using Likert scales – 7%. Within the current review, the overall number of studies (71) was almost evenly distributed between DCE/CA methods at 39 studies and those which incorporated an element of trading (including CV and TTO studies) at 31 studies. This reflects both the underlying literature search terms and the criteria used to include citations for further review. The latter called for the exclusion of studies that compared programmes/interventions without the valuation of meta-health effects. Such studies are likely to have been included in the review by Opmeer et al. (2010).

Similarly, a broader range of valuation methods has been included in the current review compared with Brennan et al. (2013)[39] which focused only on the use of utility based measures for the assessment of meta-health effects. The current review goes beyond Brennan et al. (2013)[39] in its focus in that it has sought to further explore the links between meta-health effects and: the type of evaluations conducted; the health care interventions; and sources of differences in resulting values.

In expanding the scope of the evaluation methods covered, the current review allows inferences to be drawn on whether the method of evaluation can be linked to the magnitude of the effect demonstrated for a meta-health effect. It would appear that, all else being equal, the largest relative effects for a meta-health effect are produced using CV methods and the smallest effects are produced using utility measures. This result is tempered by the number of studies for which relative values for meta-health effects were available, and the ability to compare them, both across meta-health effects and for a given meta-health effect across studies.

It was possible within this review to compare studies valuing information and convenience, restricting those comparisons by their respective valuation methods – WTP and utility measures. Even when controlling for the valuation method and meta-health effect in this way, other differences in the context (such as the health condition)

and approaches to valuing meta-health effects make it difficult to compare the results. Overall, the use of DCEs appeared to be more closely associated with demonstrating value for meta-health effects. This is not surprising given that DCEs are designed to describe a health state/intervention in terms of its attributes: meta-health effects can be readily conceived of within such a structure. Similarly, the use of TTO/SG to demonstrate value for meta-health effects is not surprising since health scenarios can be constructed to express the domain/element of interest (in this case meta-health effects).

A large proportion of studies valuing convenience as a meta-health effect were also investigating drug interventions/treatments. This reflects the ability to describe drug interventions as attributes amenable to investigation as meta-health effects. In addition, it is likely that meta-health effects – particularly convenience - are becoming a more important means of differentiating and valuing new drug interventions. This is to some extent corroborated by the results of the review of decisions made by the PBAC in which the key meta-health effect included in submissions that contained such a claim was associated with improved convenience. Information was closely linked to screening interventions. However, there is no reason to expect that methods used in a particular intervention to investigate a meta-health effect cannot be applied to other interventions or other meta-health effects.

The results from this review indicate that different valuations arise depending on the magnitude of the convenience gain under investigation (witness the difference in Johnson et al. (1996)[62] and Boye et al. (2011)[91]). However, as previously noted such cross-study comparisons are limited by differences in the conditions under investigation and the study methods. Testing whether the magnitude of the convenience gain, and potentially the population to which it accrues influences the value derived for convenience is empirically possible.

Exploring the reported values for information as a meta-health effect also revealed that perhaps individuals value receiving information irrespective of its content; the results in Neumann et al. (2012)[72] showed that perfect or imperfect information had the same relative value. The assessment conducted in this review appears to suggest that individuals place a value on information for information's sake, regardless of the quality of that information. This is consistent with work by Shani et al. (2008)[163] from psychological economics which suggests that individuals would rather know, even if in knowing they reveal a negative outcome.

Potentially further evidence of this effect is provided by Donaldson and Shackley (1997)[146], who find that if the process itself is described (through the provision of more information) it produces a lower WTP than if just the label for the surgery (alternative versus traditional) is used. Within that study the inclusion of additional information was negatively received, resulting in a reduction in WTP, and revealed the process individuals had inferred when not supplied with that information.[146] A similar information effect is apparent in Osborne et al. (2007)[64] where there is a difference in values between the anchor and treated states resulting from explaining the process.[64] Thus, while information as a meta-health effect can be valued, so too the amount of information provided in health states/scenarios designed to elicit meta-health effect values also impacts on their value.[20]

While a number of the studies considered the type and quality of information provided to respondents about the meta-health effects, none specifically addressed the balance between information provided on the meta-health effects *vis a vis* the health outcomes. In addition to the amount of information provided regarding a meta-health effect, it is worth exploring the manner in which meta-health effects are framed in relation to overall health and treatment effects. As observed for the valuation of information, the link between the meta-health effect and efficacy appears to influence the magnitude of the value derived. Similarly, within Johnson et al. (1996)[62], the manner in which adverse effects are described compared to the efficacy parameters, suggests a possible

avenue for the large difference seen in that study. These framing effects can be tested further. That is, for a given meta-health effect of interest, it would be possible to test the impact on individuals' preferences of varying the degree of information provided with regard to the health outcomes that are associated with the programme that delivers that meta-health effect. The purpose of such an exploration would be to examine the extent to which valuations of meta-health effects are dependent on the degree of information provided on the associated health outcomes.

2.3.4.2 Limitations

One of the key limitations of this review was the lack of consistent keywords to use in the search for articles containing meta-health effects. This has been previously noted by Opmeer et al. (2010).[5] There are no specific Medical Subject Headings (MESH) or index terms which pertain to meta-health effects or non-health outcomes. Published studies containing meta-health effects might therefore be indexed under a number of different terms not covered by those used in this review.

The classification of meta-health effects across the studies has been conducted on the basis of what is reported in the publications. It is possible that in some instances, the authors were not intending there to be a distinction between health and meta-health effects associated with a product or service under review (and therefore might not have designed their study with meta-health effects in mind).

Comparisons across meta-health effects have been limited due to the number of studies available for a given meta-health effect which were conducted using the same valuation method. Even where available, these comparisons are all tempered by the fact that it was not possible to compare studies within a given health care context. While it might appear that a meta-health effect is a meta-health effect, framing and contextual effects are likely to mean that assessment of a given meta-health effect (e.g. convenience) will vary depending on the health care setting in which it is measured. This was well illustrated by the difference observed in the values for convenience

observed between the studies by Osborne et al. (2007)[64] and Boye et al. (2011).[91] Overall, the ability to confidently compare values across studies for meta-health effects will remain limited until such time as they are consistently and repeatedly assessed within a given health care context.

Confidence in the tests of association between which meta-health effects were investigated and the underlying study methods was limited by the number of studies for any given meta-health effect. The associations shown should be considered exploratory given the number of studies available per meta-health effect. Finally, the classification of meta-health effects and all data extraction was conducted by one researcher. While this ensures consistency in approach, it also means that there was no external validation of the data extraction process.

2.3.5 Concluding Comments

This review demonstrates there is considerable variability in how meta-health effects have been investigated and valued. Of those studies that produced a value for a meta-health effect, over half reported it using a WTP measure. Analysis of the magnitude of the resulting valuations suggests WTP based values were larger than those garnered using utility based assessments. This has implications for the interpretation of meta-health effects values derived for the same intervention through different means and potentially bears further exploration.

In terms of specific meta-health effects, convenience was investigated most often, and largely in studies focused on drug interventions. To the extent that it was possible to compare across studies, it would appear that the magnitude of the gain in convenience resulted in different measures of relative value. Whether such a finding can be generalised is not clear, but is testable.

There are considerations regarding the measurement and valuation of meta-health effects that have not been addressed in this review. Most importantly perhaps is the

extent to which elicited values should be adjusted to reflect to whom that value accrues relative to who bears the burden for health care provision. While testable, the implications of incorporating such an adjustment are broad (including for the manner in which clinical effects and health outcomes are treated) and are investigated as part of the RA Therapy study (described in Chapter 1 and presented in Chapter 5 of this thesis).

2.4 Conclusion Regarding Existing Evidence

The evidence presented in this chapter indicates that meta-health effects are being incorporated into reimbursement claims made to the PBAC in Australia. The review of the PSDs showed that convenience was the most commonly cited meta-health effect. This is in accordance with the evidence from the published literature that also found that the majority of studies investigating the value of meta-health effects focused on convenience gains. The importance of convenience is thus reflected in the empirical work undertaken in this thesis in which potential convenience gains are tested in three different health care contexts, the choice of patient-GP relationship, treatment for RA and management of breast cancer recurrence risk.

Of the published studies that produced a value for a meta-health effect, over half were expressed using a WTP measure. Analysis of the magnitude of the resulting values indicated that, all else being equal, those measured using WTP were larger than those obtained using utility based assessments. The review also showed that framing effects are likely to influence the values derived for meta-health effects. The influence of framing on WTP valuations of meta-health effects using DCEs is investigated in both the RA Therapy and Mastectomy studies in this thesis.

That different valuation methods place different weights on meta-health effects is potentially problematic. This is particularly the case where values from different studies, using different methods might be used to make decisions regarding the subsidy of new interventions. In those instances, it is important that decision-makers,

like the PBAC, have confidence in the consistency of the methods used to value the meta-health effects, if not in the resulting values themselves.

3 Meta-Health Effects and the Factors Influencing Patient-GP Loyalty Relationships

Chapter Summary

Previous studies of patient-GP relationships have focused on their socio-demographic determinants. The research in this chapter extends those approaches by including the meta-health effects of choice (e.g. ability to choose a GP) and convenience (e.g. the practice location) as determinants of the patient-GP loyalty relationship.

Data were from a survey of 2,303 Australian respondents with at least one GP visit in the past 12 months. The patient-GP loyalty relationship was modelled using a bivariate probit (separate decisions about remaining loyal or multiple practice use) or multinomial probit (contemporaneous decision about remaining loyal or multiple practice use) specification. The choice of patient-GP loyalty relationship was dependent on a group of socio-demographic variables, as well as the importance of meta-health effects and price sensitivity. Estimated coefficients, marginal effects and probabilities for the two models were compared.

The results indicate that choice influenced patient-GP loyalty relationships; respondents who felt that having a choice of GP was important had a higher probability of being loyal. Convenience did not influence the patient-GP loyalty relationship. Being older and retired, and living outside of a major city, were associated with a higher probability of having a usual GP practice, or being loyal. Those who were younger had a lower probability of being loyal and a higher probability of multiple practice use. Price sensitivity also influenced the patient-GP loyalty relationship; those who felt access to bulk-billing was important had a higher probability of multiple practice use. While the results from the two probit models were largely consistent, those from the bivariate specification were more readily estimable and consistent with the underlying data.

The results from this research expand on the previously published literature by demonstrating the influence of meta-health effects on patient-GP loyalty relationships. The finding that choice

influences patient-GP loyalty relationships helps to inform potential policy changes, such as the introduction of capitation payment systems, that might impact on patient choice.

3.1 Introduction

One of the stated goals of the National Primary Care Strategic Framework for Australia is to "*Promote health system models that facilitate long term relationships between consumers and general practices*".[164] Understanding the factors that influence how those relationships are formed is important in order to assess policies designed to achieve such a goal.

The role of choice in determining relationships between patients and their general practitioner (GP) is particularly relevant in the context of the Australian health care system in which patients are free to move between providers without prior enrolment or registration. The impact of policies that might affect the freedom to choose providers, such as programmes to encourage greater continuity through patient enrolment[165,166], will be influenced by the nature of relationships individuals form with their GPs. Whether patients consider themselves to be loyal to a single GP or practice, to have a usual GP but visit multiple practices, or do not claim an affiliation with a particular practice or practitioner, might influence how they will respond to enrolment or capitation arrangements.

In this research the phrase 'patient-GP loyalty relationship' is used to encompass the different relationships an individual might have with their GP practice. Patient-GP loyalty relationships not only reflect individuals' underlying socio-demographic characteristics, but potentially the value and importance they place on the experience of care.[166-168] The review of the evidence presented in Chapter 2 identified choice and convenience as important meta-health effects that influence how individuals value the experience of care. With this in mind, the importance individuals place on choice and convenience might also influence the relationships they form with GPs. In this chapter

the influence of socio-demographic factors and meta-health effects on the relationships individuals choose to form with their GPs is explored.

3.1.1 Exploring the patient-GP loyalty relationship

Previous studies of the determinants of patient-GP loyalty relationships have focused on sociodemographic factors.[169-172] In a multivariate analysis of survey data from 1,146 Australians, McRae et al. (2011)[170] find that individuals who are older, retired, have lower incomes, live in major cities and are satisfied with their GP were more likely to classify themselves as being affiliated with a GP. A univariate analysis of the same data revealed that smokers and those without private health insurance were less likely to be affiliated with a single GP.[170] Similar findings were reported for New Zealand[169] and the United States[172], although the influence of gender differs across these two studies; in the United States males are less likely to have a usual source of care than females, while the opposite has been observed for New Zealand.[169,172]

In addition to socio-demographic factors, there is also the potential for patient-GP loyalty relationships to be influenced by meta-health effects, such as convenience (e.g. location of the practice, or the availability of appointment times) and choice (of a particular GP, or the option of choosing from multiple practitioners). Fotaki (2013)[167] poses the question of whether such choice should be considered a means to an end, resulting in more efficient and equitable health care systems, or an end in and of itself. However, the evidence for gains in efficiency and equity is not clear. There is some evidence that enhanced continuity, in the form of ongoing patient-GP relationships, is associated with reduced hospital use (emergency department) and lower health care costs.[173,174] Whether such relationships are the result of choice (patients choosing to remain with their preferred provider)[174], or of reduced choice associated with GP practice enrolment is not clear.[173] Results of an experiment in Sweden to introduce choice of provider into primary care relationships showed that this led to an increase in GP visits overall, and that use of the chosen provider increased most among those with

higher income levels and older age groups (65-84 year olds); perhaps neither efficient nor equitable.[175]

Whether there is an intrinsic value to choice might depend on whether individuals are exercising their autonomy to choose between GPs within a practice, or between different practices. In a qualitative analysis of such choice among patients in Southampton UK, Freeman and Richards (1993)[176] find that while patients value the ability to choose their own practitioner, and have a preference for the same practitioner, they might waive that choice where it results in a delay in receiving care, choosing instead to opt for another practitioner in the same practice. Preferences for choice in the primary care setting have also been investigated using DCEs, including in terms of choice of practitioner and involvement in decision-making[95,157,177-183], and whether there was an aspect of continuity of care available in the GP service (in terms of registration, or access to a known clinician or medical records).[157,179,184-189] Beyond choice, the role of convenience in understanding patient preferences for types of GP practices/services has also been explored in terms of the location of a practice, time of service, or choice of appointment time.[94-96,179,181,184,185,187-193]

3.1.2 Exploring the role of meta-health effects in patient-GP loyalty relationship

The factors influencing patient-GP loyalty relationships are explored in this chapter using the results of a survey of Australians regarding their use of and attitudes toward GP services. Survey respondents' health and socio-demographic characteristics are explored as possible factors influencing the patient-GP loyalty relationship. The role of meta-health effects in influencing the patient-GP loyalty relationship is explored using responses to a number of questions regarding respondents' opinions about what they consider to be important when choosing between GP practices.

The analysis in this chapter provides two main points of difference from previous examinations of patient-GP loyalty relationships: first it investigates the role of meta-health effects in affecting the patient-GP loyalty relationship; second, as foreshadowed

in Chapter 1, the role of meta-health effects is examined using attitudes rather than preferences. The use of attitude ratings from a survey of individual experiences with GP services as a measure of meta-health effects provides a direct assessment of the role of those effects in forming patient-GP relationships. It does so without considering trade-offs within hypothetical scenarios as would typically occur with preference measures. In this manner it provides an insight into the value of attitude data in exploring the role of meta-health effects when deciding about the type of patient-GP loyalty relationship to form.

In this analysis, the possible patient-GP loyalty relationships are classified as; loyal, sometimes loyal, multiple practice user, and multiple single practice use (individuals who do not identify with a usual practice, but use one practice multiple times). The derivation and specification of these relationships is described in more detail below. Methodologically, the factors influencing the choice of patient-GP loyalty relationship are explored using two probit specifications; a bivariate analysis that preserves the sequential nature of the survey questions pertaining to GP-practice loyalty; and an alternative multinomial probit (MNP) specification in which the responses to the survey questions are analysed as joint outcomes. The bivariate analysis thus provides some insight into whether decisions regarding patient-GP loyalty occur in two steps (decide to be loyal to a GP, then decide whether or not to see multiple practitioners), while the MNP models this decision as one single step.

The results of the two specification methods are compared in terms of the estimates they provide, and importantly their ability to describe the available data on the patient-GP loyalty relationship. The influence of meta-health effects compared with health and socio-demographic factors on the patient-GP loyalty relationship is investigated by assessing their respective impacts on the marginal effects and probability of membership for individuals' in each of the patient-GP loyalty categories.

3.2 Methods

3.2.1 Overview

The role of meta-health effects in influencing the patient-GP loyalty relationship is investigated in this research by assuming that individuals derive utility from the relationship they form with their GP. The nature of that relationship and how it is influenced by factors such as socio-demographic characteristics, attitudes towards meta-health effects and price sensitivity, is investigated using data from a survey of the general community in Australia.

3.2.2 The survey data

3.2.2.1 The online survey

The data used in this chapter are drawn from an online demographics, attitudinal and health care experiences survey developed and implemented by the Centre for Health Economics Research and Evaluation (CHERE); herein the CHERE Survey.[194] While the survey was completed as part of a broader project, this researcher had a key role leading the design and implementation of that survey, including the development of the specific questions used in this thesis.

The focus of the survey was individuals' most recent primary health care experiences, with the majority of questions being about GP usage. Questions on use related to individuals' most recent GP visits and those in the 12 months prior to completing the survey. To allow for the analysis of patterns of use, questions were also included about current health and socio-demographic characteristics. To allow for comparability, a number of the questions[xiii] included in the CHERE Survey drew on existing questionnaires such as the Australian Bureau of Statistics Patient Experiences in Australia Survey[195] and the Commonwealth Fund Survey.[196] The survey also included questions specific to the CHERE Survey that were about GP practice

[xiii] This included questions such as health status, socio-demographic factors and what occurred during GP consultations.

structure, and what individuals considered to be important in choosing between providers of primary care. The latter was assessed based on respondents' ratings of the importance of 36 attributes affecting the choice of GP.[xiv] Ratings of attribute importance in the survey were anchored on a scale from 1 to 5, being 'not at all important' to 'extremely important'.[197] A copy of the survey is provided at Appendix 2.

3.2.2.2 The survey sample

The survey was administered in July 2013 to a sample of individuals aged 16 years or older on the Pureprofile online panel using the Qualtrics platform.[194,198,199] This included all individuals regardless of the number of GP visits they had in the previous 12 months. However, as the focus of the research in this chapter is on assessing the patient-GP loyalty relationship, the main analyses were conducted for those respondents who reported at least one GP visit in the previous 12 months. However, some of those respondents might have only had one visit, and thus not had an opportunity to choose between single and multiple practice use. Thus, a sensitivity analysis is conducted in which respondents are required to have had at least two GP visits in the previous 12 months.

3.2.3 The analysis variables

This research explores the factors influencing patient-GP loyalty relationships, and how individuals choose those relationships. One possibility is that individuals might choose whether or not to remain loyal to a GP practice, are they 'loyal'; then choose whether or not to use the same or multiple practices when they actually visit the GP. This choice possibility is explored using a bivariate probit model. The prospect that individuals don't first characterise their patient-GP relationship, but rather choose between the four types of relationship when choosing a GP to visit, is explored using a

[xiv] These attributes were identified in a review of the DCE literature relating to primary care conducted during the development of the CHERE Survey.

multinomial probit (MNP) model. The variables used in investigating these models are described in this section, with the specification of these models described below.

3.2.3.1 Constructing the dependent variable

The dependent variable for use in the two probit specifications was drawn from two questions in the CHERE Survey: "*Do you usually go to the same practice each time you need to see a GP?"* and *"In the last 12 months, have you been to more than one practice?".*

The construction of the dependent variables using these two questions is presented in Table 9. Dependent variables for bivariate probit analyses are binary in nature.[200] Accordingly, respondents were classified as being 'loyal' or not (1,0) for the first choice in that relationship (this decision is labelled 'Usual practice'), then as users of multiple practices or not (1,0) for the second choice (this decision is labelled as 'Multiple practice'). For the MNP, the responses were combined using cross-tabulations to construct a four level dependent variable. Frequencies of response for each of the dependent variable outcomes are provided in the response section of this chapter.

Table 9: Constructing the patient-GP loyalty variable

	Survey Question Responses		**Coding for Model**		
	Do you usually go to the same practice each time you need to see a GP?	*In the last 12 months, have you been to more than one practice?*	*Bivariate Probit: Usual practice*	*Multiple practice*	*MNP*
Relationship					
Loyal	Yes	No	1	0	1
Sometimes loyal	Yes	Yes	1	1	2
Multiple practice user	No	Yes	0	1	3
Multiple single practice user	No	No (with more than one visit)	0	0	4

Note: Multiple single practice users reported using a single practice multiple times without having noted they usually went to the same GP practice.

It is possible that the dependent variable could also have been expressed as three levels only, collapsing multiple practice use into one level. The implications of this for the regression results are tested in a sensitivity analysis.

3.2.3.2 The explanatory variables

3.2.3.2.1 Defining variables to capture meta-health effects

Respondents to the CHERE Survey who used more than one practice were asked to report the main reasons for that use from a pre-specified list of options. These options included doctor availability, visit costs, being away from home, the location of the practice, practice opening hours, the services available, the ability to be bulk-billed and other reasons (directly provided by respondent). Survey respondents were able to choose more than one option from this list for why they chose to use multiple practices. Thus, these could not be used readily as variables to explain the patient-GP loyalty relationship.

Instead, the reasons for multiple practice use have been used to guide which of the 36 attitude rating variables would be used to gauge the influence of meta-health effects on patient-GP loyalty relationships. The reasons most often cited as the basis for using multiple practices corresponded to the meta-health effects of choice and convenience, and are presented in Table 10. As previously described, these attitude variables are respondents' ratings (1, unimportant to 5 extremely important) of the importance of those factors to their decision-making when choosing between GPs.[197] For the research in this thesis, those ratings were collapsed into a binary classification of importance; important (levels 4 and 5), or unimportant (1-3).

Table 10: Meta-health effects and attitude variables

Meta-Health Effect	Attitude Variable	Label
Convenience	Location of the GP Practice	*Location*
	Practice Opening Hours	*Hours*
Choice	Choice of GP	*GPChoice*
	GP is part of a Group	*GPGroup*

Abbreviation: GP, general practitioner.

Two attitude variables were included to reflect 'choice'; Choice of GP (*GPChoice)*, and GP is part of a Group (*GPGroup*).[xv] Accordingly, both variables are included in the analysis, but are anticipated to measure the same underlying meta-health effect; choice. *Location* and *Hours* are included to capture the potential effect of convenience factors in mediating the patient-GP loyalty relationship. The importance of the availability of bulk-billing (*BBilling)* was also included to assess the influence of individuals' price sensitivity on the patient-GP loyalty relationship.

3.2.3.2.2 Socio-demographic and practice specific variables

Following previous Australian and New Zealand studies of patient-GP relationships[169,170,201], the following socio-demographic variables were included: self-reported health (*Health)*; age (*Age*); place of birth (*Origin*); education (*Education*); residence (*Area*); employment status (*Employment*); income (*Income*); private health insurance status (*PInsurance*); and, smoking status (*Smoking*). As well as potentially indicating health status, *PInsurance* and *Smoking* are included in this instance as potential markers of attitudes to risk; having private health insurance indicating risk aversion and hence the potential to more actively engage with care[202], while smoking could indicate risky behaviour and therefore a reduced potential to engage with care.[203]

Two other sociodemographic variables have been included: *Gender* and *Concession* (concession card status). The influence of gender on the patient-GP loyalty relationship has varied across studies,[169,170,201] so its effect is uncertain in this instance. While McRae et al. (2011)[170] found that *Concession* was not significant, it is included in this analysis given the public nature of funding for access to GP care in Australia, and to control for potential differences in the underlying survey population.

Although an individual's satisfaction with a GP visit has been shown to influence their likelihood of ongoing affiliation with that GP, there is the potential for endogeneity in

[xv] GPGroup captures the possibility that individuals derive value from being able to choose a GP from within a group if their preferred GP is not available.

that relationship.[170,201] That is, it is likely that an individual who claims to be loyal to a practice is also likely to report ongoing satisfaction with a visit. Such a relationship was shown by Nutting et al. (2003)[204]; satisfaction with GP visits was higher among patients who valued continuity and had a regular GP. Similarly, low satisfaction with GP services has been found to be a leading cause of patients changing practice loyalties (by changing their enrolment status) in the UK.[205] Visit satisfaction is therefore excluded on the basis that it is potentially endogenous.

The CHERE Survey also asked respondents to provide information relating to their GP practice (such as the number of practitioners), or GP visits (duration, or number of visits over the past 12 months). However, the number of GPs and visit duration related only to the last visit and is unlikely to reveal information about the possible range of characteristics applied to other practices from which respondents might have chosen. Rundle-Thiele et al. (2010)[171] report that the frequency of previous GP visits influenced the strength of ongoing association between a patient and their GP. However, it is possible that the frequency and number of GP visits is influenced by the patient-GP loyalty relationship; individuals who are loyal to a GP have more visits. Thus, given this potential endogeneity, and the use of the number of GP visits to define the relevant sample for analysis, the number of GP visits is excluded as an explanatory variable from the analyses within this chapter. The full list of explanatory variables is provided in Table 11.

3.2.4 The modelling framework

The modeling approach in the analysis is motivated by assuming that individuals will choose the type of patient-GP loyalty relationship from which they derive the greatest utility, which can be expressed as:

$$U_{ij} = \propto_j + x_i'\beta_j + \varepsilon_{ij} \qquad (1)$$

where U represents the unobserved utility, and will be captured by estimating the latent variable U^* which is the observed choice corresponding to the patient-GP relationship, x is the vector of explanatory variables (including demographics, visit numbers and attitudes to practice choice), $\propto$ is the alternative specific constant for the jth option, and ε the vector of error terms (assumed to be normally distributed).

3.2.4.1 The multinomial probit

Within the MNP analysis, it is assumed that the patient-GP loyalty relationship is the product of a contemporaneous choice decision in which the possible outcomes from equation (1) are *j=1,2,3*,4: 'loyal', 'sometimes loyal', 'multiple practice user', and 'multiple single practice user' respectively. These are specified in equations (2)-(5). For estimation purposes, the base case is *j*=3 (the factors influencing the choice of all other relationship types is relative to the choice to be a 'multiple practice user').

$$U_{i1} = \propto_1 + x_i'\beta_1 + \varepsilon_{i1} \quad (2)$$

$$U_{i2} = \propto_2 + x_i'\beta_2 + \varepsilon_{i2} \quad (3)$$

$$U_{i3} = 0 \quad (4)$$

$$U_{i4} = \propto_4 + x_i'\beta_4 + \varepsilon_{i4} \quad (5)$$

The vector of ε is normally distributed, mean=0, and taking into account the normalisation (using option 3 as the base), has the following covariance matrix (allowing for correlation across the options)[200,206]:

$$\tilde{\Omega}_1 = \begin{bmatrix} \theta_{11} & & \\ \theta_{12} & \theta_{22} & \\ \theta_{14} & \theta_{24} & \theta_{44} \end{bmatrix}$$

and:

$$\theta_{11} = \sigma_{11} + \sigma_{33} - 2\sigma_{13}$$

$$\theta_{22} = \sigma_{22} + \sigma_{33} - 2\sigma_{23}$$

$$\theta_{44} = \sigma_{44} + \sigma_{33} - 2\sigma_{43}$$

$$\theta_{12} = \sigma_{12} + \sigma_{33} - \sigma_{13} - \sigma_{23}$$

$$\theta_{14} = \sigma_{14} + \sigma_{33} - \sigma_{14} - \sigma_{43}$$

$\theta_{24} = \sigma_{24} + \sigma_{33} - \sigma_{23} - \sigma_{24}$

where each element can be described as the variances (e.g. σ_{11} the variance with respect to choice option 1 ('loyal')) and covariances (e.g. σ_{23} the covariance of choice option 2 ('sometimes loyal') and 3 ('multiple practice user')). An alternative model structure (a mixed logit), assuming the independence of irrelevant alternatives (IIA) and an identity covariance matrix for the vector ε, is tested in a sensitivity analysis.

3.2.4.2 The bivariate probit

Within equations (2)-(5), the influence on choice of each of the dependent variables, x, depends on individual specific factors only and does not vary with the choices themselves. That is, the explanatory variables do not vary with the relationship options themselves. Keane (1992) suggests that in such circumstances, the use of MNP might lead to estimates of the underlying relationship that are not stable.[207] In addition, the structure of the MNP is such that it assumes the choices are being made contemporaneously, which might not always be the case. Individuals might choose whether or not to remain loyal to a GP practice and separately decide whether to use multiple practices. For these two reasons (the potential instability of the MNP, and the potential that relationship choices are not contemporaneous) an alternative modeling structure was investigated using a bivariate probit. This describes the patient-GP loyalty relationship in two parts:

$U_{ij}^{L} = \propto_j + x_i'\beta_j + \varepsilon_{ij}$ (6)

$U_{ij}^{M} = \propto_j + x_i'\beta_j + \varepsilon_{ij}$ (7)

Here, U_{ij}^{L} is the choice by the individual as to whether to remain loyal to the GP practice. An individual might choose to remain loyal, j=1 loyal, or not, j=0. U_{ij}^{M} is the separate choice to use multiple practices (j=1) or not (j=0). Combined, these two decisions give rise to four possible outcomes: 'loyal' (1,0) or 'sometimes loyal' (1,1) for

the first decision; and 'multiple practice use' (0,1) or 'multiple single practice user' (0,0) for the second decision.

The existence of 'sometimes loyal' as an outcome indicates that outcomes from (6) and (7) are not necessarily mutually exclusive; it is possible for a 'loyal' patient, with a usual practice, to use more than one practice (e.g. when they are away from home). This differs from the use of bivariate estimation to describe censored or selection relationships.[xvi]

Nonetheless, the two possible patient-GP loyalty choices ('loyal' or 'multiple practice use') are interdependent; the factors that influence one are likely to influence the other, and the outcomes arise from decisions made by the same individual. This interdependence is explored using the bivariate probit specification for which it is anticipated that there would be some correlation in the factors influencing the choices to the first decision (to remain loyal) and the second decision (to use multiple practices). The bivariate probit assumes that the error terms are distributed as normal $var\left(\varepsilon_{ij}\right) = 1$ with $cov(\varepsilon_{i1}, \varepsilon_{i2}) = \rho$. The degree of correlation between the choice decisions is thus described by ρ.

3.2.5 Modelling approach

The analyses proceeded by first regressing the multinomial dependent variable for loyalty (the four level outcome variable as defined for the MNP analysis) on the socio-demographic variables using ordinary least squares (OLS) to examine the data structure for multicollinearity between the variables. Following any changes required to address observed multicollinearity, the MNP and bivariate probit analyses were initially estimated including only the socio-demographic factors as explanatory variables. Subsequently, price sensitivity and attitude variables testing the influence of meta-health effects on the patient-GP loyalty relationship were included. In addition,

[xvi] Such as screening behaviour conditional on awareness of test availability, or continuing university education based on completion of first year).[208,209]

the variable for self-reported health status was substituted with a variable for whether respondents reported the presence of any ongoing chronic health issues. This change was made because ongoing chronic health issues may be a stronger determinant of the patient-GP loyalty relationship than self-reported health status, which relates to the health rating at the time of completing the survey. Indeed, Jatrana et al. (2009)[169] found that individuals in New Zealand with more chronic health issues were more likely to report going to the same provider if they needed to see a doctor.

3.2.5.1 Model specifications

The model specifications for the bivariate probit analysis, substituting the explanatory variables for x in (6) and (7), are presented in equations (8) and (9). The same variables were used in modeling the MNP (equations (2-5)) but are excluded for parsimony.

$$U_{ij}^{L} = \propto_j + \beta_{1j}Smoking_i + \beta_{2j}Employment_i + \beta_{3j}Area_i + \beta_{4j}Income_i + \beta_{5j}Gender_i + \beta_{6j}PInsurance_i + \beta_{7j}Origin_i + \beta_{8j}Education_i + \beta_{9j}Health_i + \beta_{10j}Age_i + \beta_{11j}Concession_i + \beta_{12j}Location_i + \beta_{13j}GPGroup_i + \beta_{14j}BBilling_i + \beta_{15j}Hours_i + \beta_{16j}GPChoice_i + \varepsilon_{ij} \quad (8)$$

$$U_{ij}^{M} = \propto_j + \beta_{1j}Smoking_i + \beta_{2j}Employment_i + \beta_{3j}Area_i + \beta_{4j}Income_i + \beta_{5j}Gender_i + \beta_{6j}PInsurance_i + \beta_{7j}Origin_i + \beta_{8j}Education_i + \beta_{9j}Health_i + \beta_{10j}Age_i + \beta_{11j}Concession_i + \beta_{12j}Location_i + \beta_{13j}GPGroup_i + \beta_{14j}BBilling_i + \beta_{15j}Hours_i + \beta_{16j}GPChoice_i + \varepsilon_{ij} \quad (9)$$

A summary of the labels and brief description of each of the dependent and explanatory variables is provided in Table 11, with summary statistics provided in the results section (Table 12).

Table 11: Variable labels and description

Variable	Label
Dependent variables	
Bivariate probit (1) - Usual GP practice	Usual Practice
Bivariate probit (2) – Multiple practice user	Multiple Practices
MNP – Multinomial response outcome	Loyal, sometimes loyal, multiple practice user, multiple single practice use
Explanatory variables	
Smoking status	*Smoking*
Employment status	*Employment*
Respondent area of residence	*Area*
Household income	*Income*
Respondent gender	*Gender*
Private health insurance status	*PInsurance*
Respondent place of birth	*Origin*
Education attained	*Education*
Self-reported health; Chronic health issues	*Health*
Respondent age	*Age*
Concession card status	*Concession*
GP practice location	*Location*
GP works as part of a group practice	*GPGroup*
Bulk-billing is available	*BBilling*
After hours visits are available	*Hours*
Choice of GP is possible	*GPChoice*

Notes: Variable names are shown in italics and correspond to the vector x in the equations defining the modeling approach.
Two separate questions were used to populate *Health* as described in the text.
Abbreviations: GP, general practitioner; MNP, multinomial probit.

3.2.5.2 Treatment of missing observations

There were missing observations for a number of the dependent variables. A cut-off of 1% missing observations was deemed to be acceptable for missing data without requiring modification of the data. Variables that did not meet this cut off (*Smoking, Area, Income, PInsurance and Employment)* were recoded to include a separate category (labelled as *Unknown*) for missing observations (these variables are all categorical, allowing the addition of another category for missing data). The effect of this method of coding missing observations on the modelled results is tested in a sensitivity analysis by recoding all missing observations to either the upper or lower category of each relevant variable (the "bounds" method), and the use of a complete case analysis which excludes respondents with missing observations in the independent variables.[210,211]

3.2.5.3 Model comparisons and analysis

To facilitate comparison of the model outcomes, marginal effects and estimated probabilities are reported for both models. The results from the two model specifications (MNP and bivariate) are presented separately and then contrasted based on the coefficient estimates, margins, probabilities and fit, similar to the process used for model comparisons in Dow et al. (2004).[212] Model fit was assessed using the pseudo-R^2, Akaike information criterion (AIC) and Bayesian information criterion (BIC). All analyses were conducted using STATA version 12, and copies of the STATA code are available upon request.

The online survey company, Pureprofile uses a general sampling technique where all panel members are invited to participate and essentially self-select into the survey. This means that it is not possible to take into account the sampling frame used when conducting analyses using survey data from Pureprofile. That is, it is not possible to adjust for the estimation procedure for the participation rate in the survey since the denominator, those who actually view the survey invitation notice, is not known. Therefore, all regression analyses within STATA have been conducted using robust standard errors to account for the survey nature of the data.[213]

3.3 Results

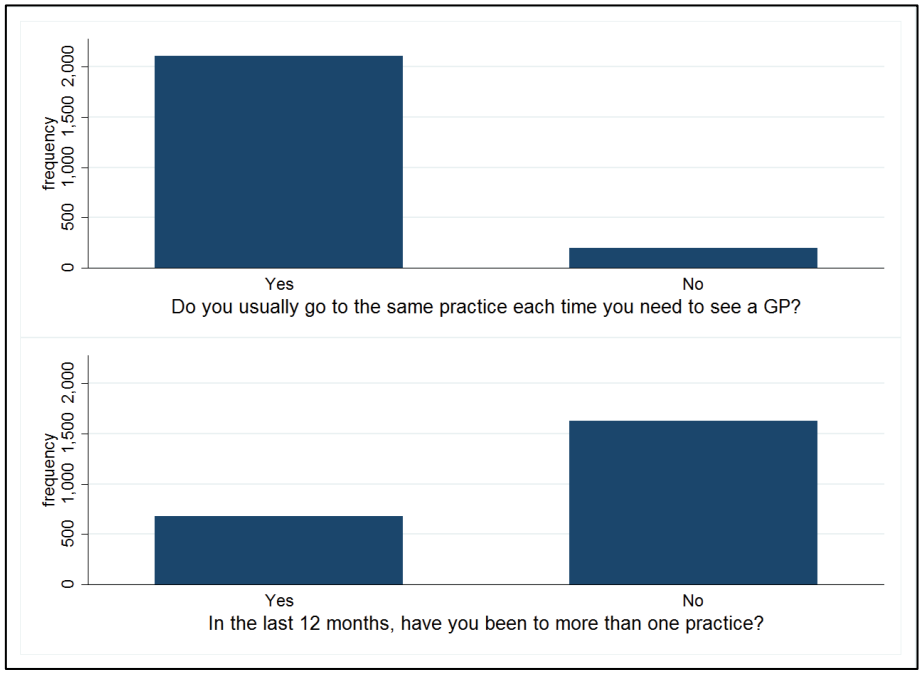
3.3.1 The data

A total of 2,477 respondents completed the CHERE Survey during July 2013. Of those, 2,303 had visited a GP practice at least once in the previous 12 months. Demographic characteristics of the participants were comparable with those of the Australian adult population with respect to gender and income (median weekly household income for Australia is $1,234, and the median category reported for CHERE Survey participants was $1,150-$1,529).[214] The youngest and oldest age groups were under-represented in the CHERE Survey compared with the Australian population.[214] Similarly, the

proportion of patients in major city areas was higher compared with those in the National Health Survey.[215] Fewer respondents reported having private health insurance compared with those in the Australian Bureau of Statistics Patient Experiences Survey.[216]

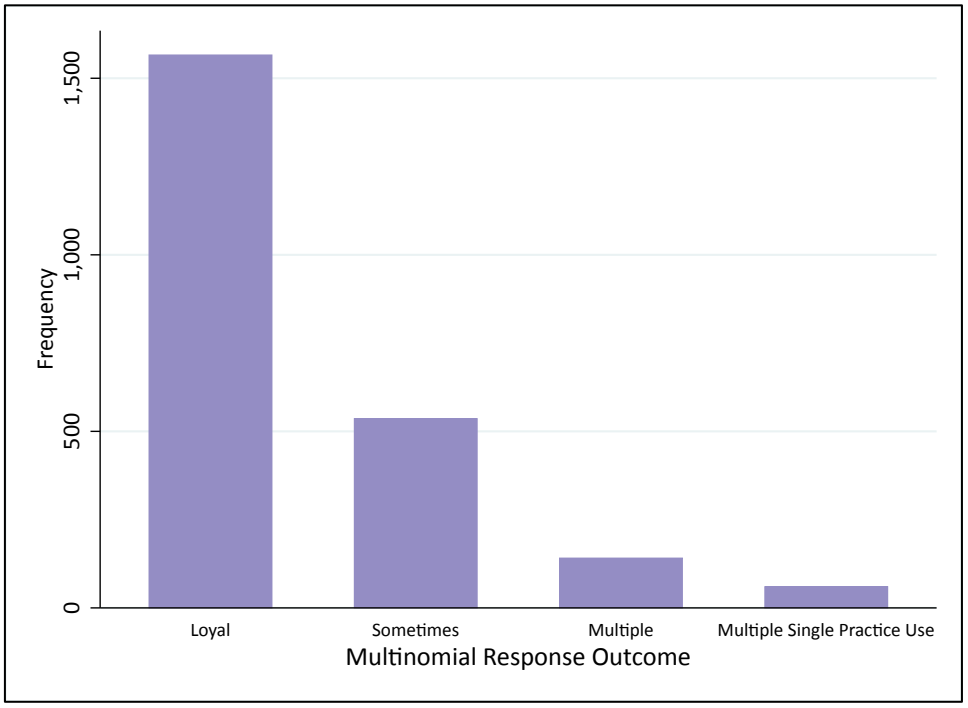
The summary data for the variables included in the analysis are provided in Table 12, with variable levels and frequencies presented in Table A 1 in Appendix 3. The pattern of GP visits in the previous 12 months for the overall sample is presented in Table A 2 in Appendix 3. The frequency of responses to the questions used to form the dependent variables is reported in Figure 10, and expressed as a contemporaneous choice in Figure 11. From these data it can be observed that the majority of respondents were 'loyal' (n=1,566; 68%) or 'sometimes loyal' (n=536, 23%), with 141 'multiple practice users' (6%) and 60 respondents (3%) in the 'multiple single practice use' category. Overall, 677 respondents reported the use of multiple practices within the last 12 months (some of which occurred among those who also report having a usual practice).

Figure 10: GP practice use – consecutive questions



Note: These questions were consecutive, but were not conditional.
Abbreviation: GP, general practitioner.

Figure 11: GP practice use – multinomial outcome



Abbreviation: GP, general practitioner.

Table 12: Variable description and data observations

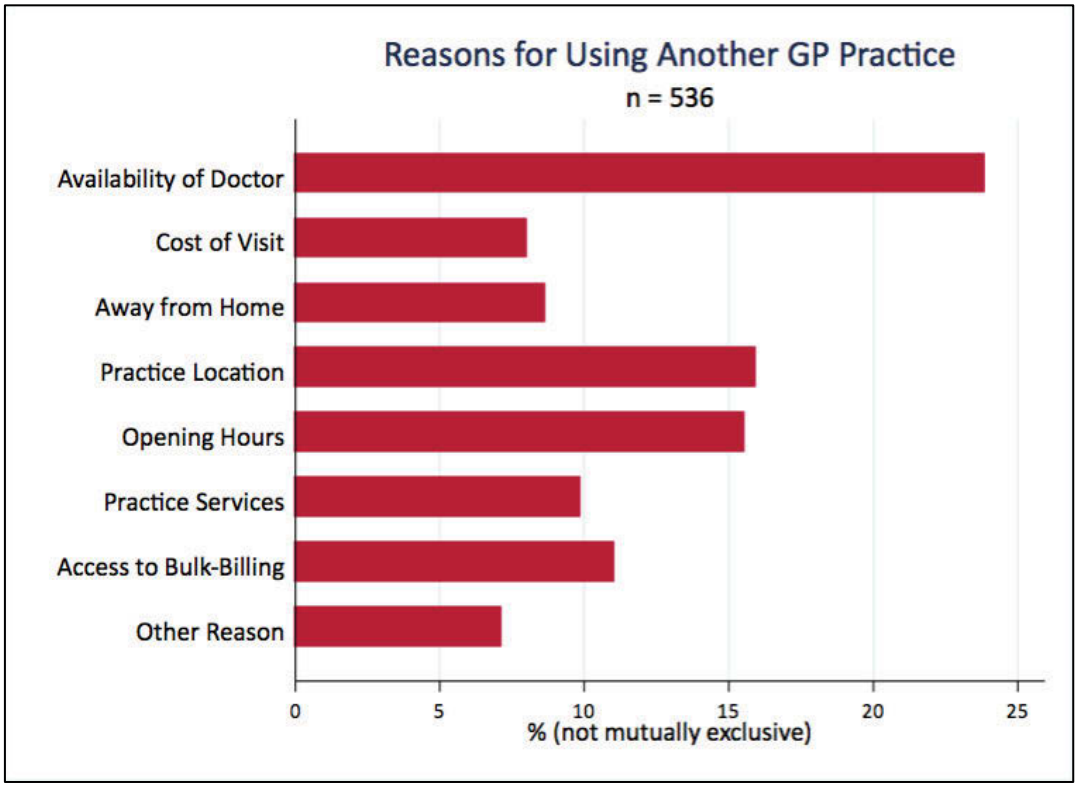
Variable	Obs.	Missing Obs.	Value Range
Bivariate probit – Usual GP Practice	2,303	0	0,1
Bivariate probit – Multiple Practice User	2,303	0	0,1
MNP – Multinomial Response Outcome	2,303	0	1,2,3,4
Explanatory Variables: Survey Derived			
Smoking: Smoking status	2,101	202	1,2,3
Employment: Employment status	2,219	84	1,2,3,4
Area: Respondent area of residence	2,198	105	1,2,3,4
Income: Household income	1,976	327	1,2,3,4
Gender: Respondent gender	2,295	8	1,2
PInsurance: Private health insurance status	2,101	202	0,1
Origin: Respondent place of birth	2,297	6	0,1
Education: Education attained	2,285	18	1,2,3
Health: Self-reported health	2,303	0	1,2,3,4,5
Chronic health issues	2,303	0	0,1
Age: Respondent age	2,298	5	1,2,3
Concession: Concession card status	2,299	4	0,1
Attitude Variables: Ratings of importance in choice			
Location: GP practice location	2,303	0	1,2
GPGroup: GP works as part of a group practice	2,303	0	1,2
BBilling: Bulk-billing is available	2,303	0	1,2
Hours: After hours visits are available	2,303	0	1,2
GPChoice: Choice of GP is possible	2,303	0	1,2

Notes: Variable names are shown in italics and correspond to those shown in equations 1-6.
Two separate questions were used to populate *Health* as described in the text.
Abbreviations: GP, general practitioner; MNP, multinomial probit; s.d., standard deviation.

The reasons for using multiple practices in the previous 12 months are reported in Figure 12. The reasons most commonly reported as explaining the use of a practice other than the usual GP practice were doctor availability (23.9%), location of practice (15.9%), opening hours (15.5%) and ability to access bulk-billing (11.0%). Overall, 56.0% of respondents gave only one reason explaining the use of another practice, 21.6% gave two reasons, and 22.5% gave three reasons or more. As previously described, responses to this question were used to measure attitude importance (important/unimportant) for the variables of practice location (*Location)*, availability of bulk-billing (*BBilling)*, availability of appointments out-of-hours *(Hours)*, being able to see the GP of choice (*GPChoice)*, and the practice being part of a larger medical group (*GPGroup)*.

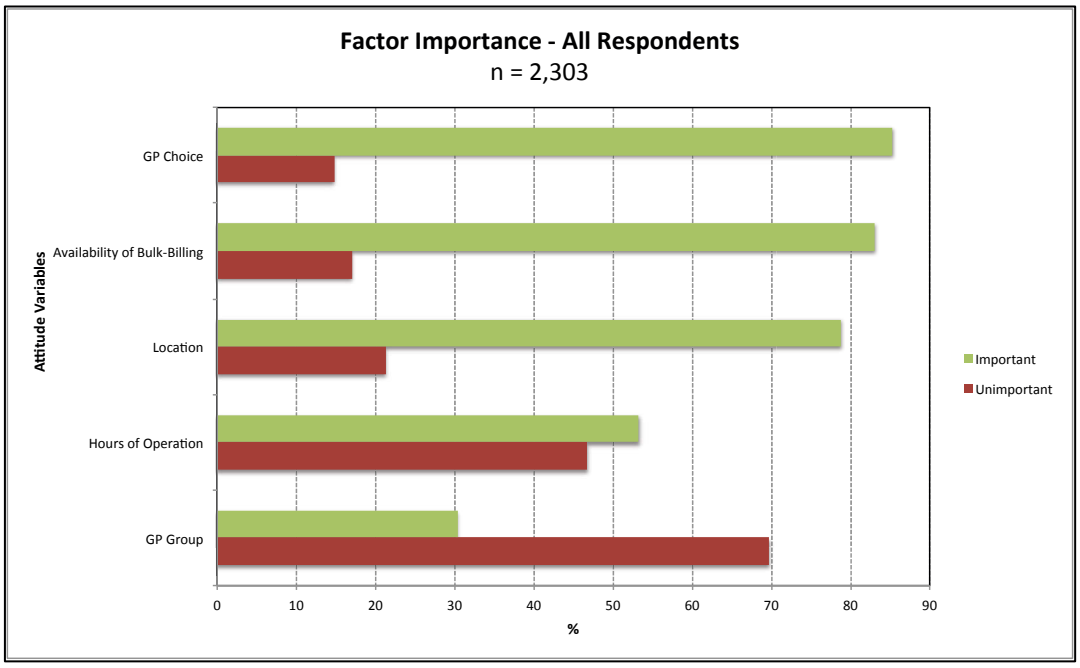
The results of the attitude ratings for the five variables of interest (*Location*, *Hours*, *GPChoice*, *GPGroup*, and *BBilling*) are shown in Figure 13. The results are shown for all survey respondents who had at least one GP visit in the 12 months prior to the survey, regardless of whether or not they had attended multiple practices. From these results it can be seen that the majority of respondents considered GP choice to be important (85.2%), followed by the availability of bulk-billing (83.0%), and practice location (79.0%).

Figure 12: Reasons for using more than one GP practice



Abbreviation: GP denotes general practitioner.

Figure 13: Attitudes ratings - importance of most frequent factors



Abbreviations: GP, general practitioner

These results indicate that attitudes associated with meta-health effects for choice and convenience are considered important in choosing between GP practices; as is price sensitivity (indicated by the attitude to bulk-billing). Accordingly, it is expected that these factors will influence the choice of patient-GP loyalty relationship. Attitudes to GP opening hours and whether the GP was part of a group were either more evenly distributed (53.3% considered opening hours to be important) or were not important; 70.0% did not consider the GP being part of a group to be an important factor in choosing between GP practices. An association for these two factors in the modelled patient-GP relationship is therefore not expected.

3.3.2 Regression analysis

3.3.2.1 Initial analyses

The results of the OLS regression revealed *PInsurance* to be perfectly collinear with *Smoking*, and the presence of multicollinearity for *Education* and *Age* (variance inflation factors (VIF) scores ranging from 1.43 to 23.03 for *Education*; and 1.68 to 6.54 for *Age*). In general, VIF scores of 10 or higher are considered to indicate the presence of multicollinearity.[183] Within these data the presence of collinearity has been addressed by reducing the number of categories describing each of the variables with a VIF above 5. For *PInsurance*, there were only three categories so this variable could not be recoded meaningfully to address collinearity. Thus it was subsequently omitted from further regressions. *Smoking*, with which *PInsurance* was collinear, was retained since smoking status was shown in univariate analyses by McCrae et al. (2011)[170] to be associated with GP affiliation. To reduce the collinearity in the remaining variables, *Education* was recoded from six into three categories (mirroring those of McCrae et al. (2011)[170]), reducing the VIF to an average of 1.51, and *Age* was recoded from five into three categories, reducing the VIF to an average of 2.01.

The results of the initial MNP and bivariate estimations using the socio-demographic variables are provided in Appendix 4. These are qualitatively consistent with previous analyses; age, gender, residential area, health status and being retired were significant

in influencing the patient-GP loyalty relationship.[169,170] From the results of both estimation methods it can be inferred that individuals were more likely to have a usual practice (be 'loyal') rather than use multiple practices if they were retired rather than employed, or living in an inner regional area rather than a major city.[xvii] Based on the MNP results, individuals were less likely to be 'multiple single practice users' compared with 'multiple practice users' if they lived in a remote area compared with a major city; or if they reported being in very good health compared with poor health. Results from the bivariate analysis indicate that individuals in the inner and outer regional areas were less likely to be 'multiple practice users' than those in major cities, and health status was not a significant determinant of the patient-GP loyalty relationship. Similarly, females were more likely to be 'multiple practice users' in the bivariate analysis; but gender was not significant in the MNP analysis. Age was a significant factor in both estimation methods; both analyses showed that increasing age resulted in an increased likelihood of being 'loyal' or a 'multiple single practice user' (in the MNP estimation method), or a reduced likelihood of 'multiple practice use' (the bivariate probit).

In contrast to a similar analysis by McRae et al. (2011)[170], income was not a significant explanatory variable in the current analysis. This could be due to differences in the coding of the income variable across the two analyses; McRae et al. (2011)[170] include only three income categories, the highest of which is $50,000 and above, while income in this analysis was classified into five categories, with the top category being $150,000 or above. This potentially dilutes the effects that might have been observed with greater clustering of the individuals into lower income categories as occurred in McRae et al. (2011).[170]

Other differences noted include that *Smoking* (in the MNP only) and *Education* (in the bivariate probit only) are both significant in the current analysis. Being a smoker

[xvii] The positive influence of living in an inner regional area on having a usual practice, or be less likely to use multiple practices, might reflect a lower density of GPs[217] in such areas.

reduced the probability of being a 'multiple single practice user' relative to being a 'multiple practice user'; while being in the unknown smoker group contributed positively to the probability of being a 'multiple practice user' relative to all other categories.[xviii] Those who completed vocational and other studies (relative to primary/secondary school only) had a higher probability of being a 'multiple practice user' relative to not.

Overall, the explanatory power of the included variables was relatively low (as indicated by the pseudo-R^2), but the estimation of both the MNP and bivariate results can be considered significant in terms of the Wald test. The inter-equation correlation for the bivariate probit was significant (ρ = -0.515; p<0.001), indicating correlation between the two choices modelled (*Usual practice* and *Multiple practices*). The negative sign is as expected, indicating that factors which act positively on the choice to have a *Usual practice*, act negatively on the choice to use *Multiple practices*.

3.3.2.2 Inclusion of attitude variables: bivariate probit

The estimation results for the bivariate probit are presented in Table 13. As previously described, a variable for the presence of one or more chronic health issues (yes, no) was substituted for current health status. This change, together with the addition of the five attitude variables, resulted in improved overall model fit with improved log-likelihoods and pseudo-R^2 values compared with the analyses based on the socio-demographic variables alone. The correlation across equations increased, with ρ = -0.535 (p<0.0001).

Within this specification, the presence of one or more chronic health issues increased the probability of being 'loyal' to a GP practice compared to not being loyal. The inclusion of the attitude variables, accounting for meta-health effects and price

[xviii] The importance of the unknown smokers category is tested in the sensitivity analysis in which a complete case analysis is conducted (essentially omitting those individuals for which data for the variable of interest are not available).

sensitivity, did not alter the signs or significance of the variables previously identified as influencing the patient-GP loyalty relationship (*Area, Employment, Education, Gender* and *Age*).

Of the attitude variables, price sensitivity as assessed by *BBilling* was significant; rating access to bulk-billing as important contributed positively to the probability of being a 'multiple practice user' (or negatively to being 'loyal'). *GPChoice* was also significant but showed a negative contribution to the probability of being a 'multiple practice user'. Thus, individuals for whom the choice of GP is important have a higher probability of being classified as 'loyal'/'sometimes loyal'. This finding indicates that the meta-health effect of choice is likely to affect the patient-GP loyalty relationship. That *Location* and *Hours* were not significant indicates that, on the basis of these data, convenience is not a factor affecting the patient-GP loyalty relationship.

Table 13: Bivariate probit coefficient estimates

	Bivariate Probit Results	
	Usual practice	*Multiple practices*
Smoking: *Non-Smoker*		
Sometimes	-0.035	0.297
	(0.204)	(0.152)
Smokers	0.009	0.126
	(0.107)	(0.078)
Unknown	-0.209	0.094
	(0.134)	(0.107)
Employment: *Employed*		
Not Employed	0.018	-0.009
	(0.106)	(0.081)
Retired	0.389*	-0.155
	(0.178)	(0.112)
Unknown	-0.127	-0.164
	(0.197)	(0.158)
Area: *Major City*		
Inner Regional	0.222	-0.392***
	(0.126)	(0.092)
Outer Regional	0.120	-0.288*
	(0.167)	(0.129)
Remote	-0.033	0.170
	(0.530)	(0.381)
Unknown	-0.164	0.057
	(0.161)	(0.134)
Income: *Low*		
Medium	-0.141	0.020
	(0.126)	(0.089)
High	-0.088	0.082
	(0.131)	(0.096)
Very High	-0.162	0.043
	(0.168)	(0.126)
Unknown	-0.063	-0.122
	(0.140)	(0.103)
Gender: *Male*		
Female	-0.036	0.155*
	(0.080)	(0.060)
Origin: *Overseas*		
Australia	0.024	-0.052
	(0.088)	(0.067)
Health: *None*		
One or More	0.260**	0.118
	(0.080)	(0.064)
Education: *School Only*		
University	0.034	0.101
	(0.101)	(0.074)
Vocational & Other	0.070	0.203**
	(0.099)	(0.076)
Age: *16-34*		
35-54	0.278**	-0.399***
	(0.088)	(0.070)
>55	0.474***	-0.658***
	(0.125)	(0.095)
Concession: *No*		

	Bivariate Probit Results	
	Usual practice	*Multiple practices*
Yes	0.005	-0.003
	(0.096)	(0.071)
Location: *Unimportant*		
Important	0.078	-0.034
	(0.101)	(0.075)
GPGroup: *Unimportant*		
Important	0.118	0.121
	(0.089)	(0.064)
BBilling: *Unimportant*		
Important	-0.052	0.203*
	(0.106)	(0.082)
Hours: *Unimportant*		
Important	-0.070	0.043
	(0.085)	(0.062)
GPChoice: *Unimportant*		
Important	0.204	-0.198*
	(0.107)	(0.085)
Constant	0.810***	-0.418**
	(0.205)	(0.159)
ρ	-0.535***	
	(0.040)	
n	2,264	
LLH	-1831.381	
Chi-2	251.500	
p-values	0.000	
Pseudo-R^2	0.083 (333.410)	
AIC	3776.762	
BIC	4103.08	

Notes: * p<0.05, ** p<0.01, *** p<0.0001
Standard errors are shown in parentheses.
The omitted (base) level for each independent variable is shown in italics alongside the category name.
Pseudo-R^2 estimated as 1-(L1/L0), where L1 is the log-likelihood of the fitted regression, and L0 is the log-likelihood of the corresponding constant only regression. Values in parentheses for the Pseudo-R^2 are the LR test statistic estimated as 2*(L1-L0).

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; GP, general practitioner; LLH, log-likelihood; MNP, multinomial probit.

The marginal effects for the significant variables in Table 13 are presented in Table 14 (the full set of marginal effects is in Table A 4, Appendix 4). The marginal effects represent the contribution to the overall choice probability of each variable category. From the marginal effects it can be seen that being in the group that considers *GPChoice* to be important adds 7.5% to the probability of being 'loyal', and reduces the probability of being a 'multiple practice user' by 2.5% (in contrast, only the coefficient estimate for *GPChoice* in *Multiple practices* was significant). Of the remaining measures

of meta-health effects, only the marginal effects related to *GPGroup* were statistically significant, making a positive contribution of approximately 5% to the probability of being 'sometimes loyal', and reducing the probability of being a 'multiple single practice user' by 1% in the bivariate analysis. Price sensitivity, as observed by the marginal effects on *BBilling*, remained statistically significant; thinking that bulk-billing is important reduces the probability of being 'loyal' by 6.2%, but increases the probability of being 'sometimes loyal' by 5.5%.

Table 14: Marginal effects: bivariate probit

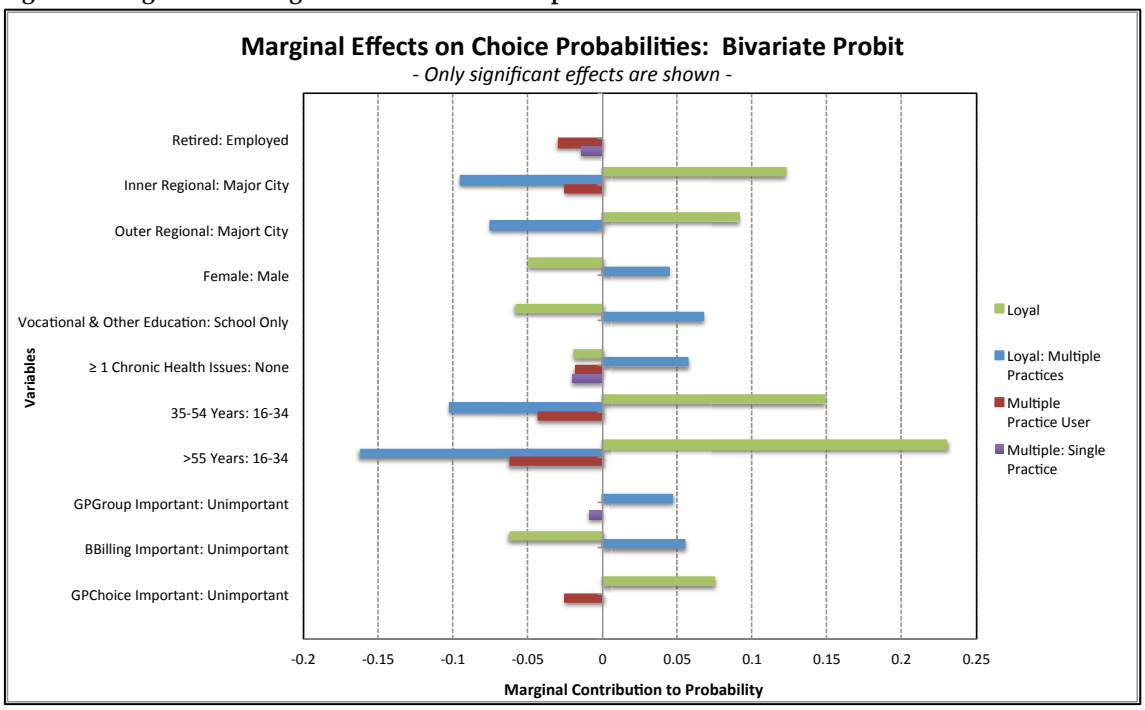
	Loyal	*Sometimes Loyal*	*Multiple Practice User*	*Multiple Single Practice User*
Employment: *Employed*				
Retired	0.065 (0.035)	-0.021 (0.032)	-0.03** (0.011)	-0.014* (0.006)
Area: *Major City*				
Inner Regional	0.122*** (0.026)	-0.095*** (0.022)	-0.025** (0.008)	-0.002 (0.006)
Outer Regional	0.091** (0.039)	-0.075* (0.032)	-0.017 (0.012)	0.001 (0.009)
Remote	-0.058 (0.143)	0.053 (0.123)	0.008 (0.059)	-0.003 (0.024)
Gender: *Male*				
Female	-0.049* (0.02)	0.044* (0.018)	0.007 (0.007)	-0.002 (0.004)
Education: *School Only*				
Vocational & Other	-0.058* (0.025)	0.067** (0.022)	0 (0.009)	-0.01 (0.006)
Health: *None*				
One or More	-0.019 (0.021)	0.057** (0.018)	-0.018* (0.008)	-0.02*** (0.006)
Age: *16-34*				
35-54	0.148*** (0.025)	-0.102*** (0.023)	-0.043*** (0.011)	-0.003 (0.005)
>55	0.23*** (0.031)	-0.162*** (0.028)	-0.062*** (0.013)	-0.007 (0.007)
GPGroup: *Unimportant*				
Important	-0.032 (0.022)	0.047* (0.02)	-0.006 (0.008)	-0.009* (0.004)
BBilling: *Unimportant*				
Important	-0.062* (0.025)	0.055* (0.022)	0.01 (0.008)	-0.003 (0.006)
GPChoice: *Unimportant*				
Important	0.075* (0.03)	-0.044 (0.027)	-0.025* (0.012)	-0.006 (0.007)

Notes: * p<0.05, ** p<0.01, *** p<0.0001
Standard errors are shown in parentheses.
Margins are the discrete change from the base levels, and are evaluated at the means for those variables that achieved statistical significance.
The omitted (base) level for each independent variable is shown in italics alongside the category name.

Abbreviation: GP, general practitioner.

Overall, the marginal effects show that *Age* and *Area* have the greatest influence on the patient-GP loyalty relationship. This can be observed from the graphical representation of those marginal effects that demonstrated significance at a minimum level of 5%; these are presented in Figure 14. From these results it can also be observed that sociodemographic factors have a greater influence in contributing to the probability of individuals belonging to specific patient-GP loyalty relationship categories than do meta-health effects, as modelled by these data. For example, the results indicate that being between 35 and 54 years of age increases the probability of being 'loyal' to a GP practice by 14.8%. This is almost double the effect associated with considering the choice of GP to be an important factor when choosing between practices; the marginal effect in this instance was 7.5%.

Figure 14: Significant marginal effects – bivariate probit



Abbreviation: GP, general practitioner.

3.3.2.3 Inclusion of attitude variables: MNP

The estimation results for the MNP are presented in Table 15. The same changes with respect to the inclusion of health status were made for the MNP as noted for the bivariate probit. This showed that individuals with more than one chronic health issue were less likely to be classified as 'multiple single practice users', relative to 'multiple practice user'. The remaining variables retained the same signs and significance as demonstrated in the initial analysis; *Smoking, Area, Employment,* and *Age* all influenced the GP-loyalty relationship. Rating access to bulk-billing as important contributed positively to the probability of being a 'multiple practice user', while considering *GPGroup* to be important significantly increased the probability of being 'sometimes loyal'. As with the bivariate probit, *Location* and *Hours* were not statistically significant (compared with the results of the initial MNP regression, addition of the five attitude variables resulted in an improved model fit in terms of the log-likelihood and pseudo-R^2 value).

Table 15: MNP coefficient estimates

	Multinomial Probit: Four Outcomes		
	Loyal *vs* *Multiple practices*	*Sometimes* *vs* *Multiple practices*	*Multiple single practice use* *vs* *Multiple practices*
Smoking: *Non-Smoker*			
Sometimes	-0.276	0.128	-0.259
	(0.295)	(0.305)	(0.476)
Smokers	-0.225	-0.103	-0.604*
	(0.155)	(0.163)	(0.266)
Unknown	-0.515**	-0.585**	-0.785*
	(0.189)	(0.203)	(0.309)
Employment: *Employed*			
Not Employed	0.037	0.029	-0.004
	(0.160)	(0.167)	(0.238)
Retired	0.531*	0.358	0.029
	(0.268)	(0.284)	(0.446)
Unknown	0.212	0.086	0.787*
	(0.297)	(0.326)	(0.399)
Area: *Major City*			
Inner Regional	0.511*	-0.019	0.213
	(0.209)	(0.219)	(0.284)
Outer Regional	0.213	-0.241	-0.242
	(0.248)	(0.271)	(0.419)
Remote	-0.289	-0.122	-8.630***
	(0.752)	(0.799)	(0.621)

	Multinomial Probit: Four Outcomes		
	Loyal *vs* *Multiple practices*	*Sometimes* *vs* *Multiple practices*	*Multiple single practice use* *vs* *Multiple practices*
Unknown	0.004	0.196	0.544
	(0.270)	(0.283)	(0.339)
Income: *Low*			
Medium	-0.148	-0.132	0.095
	(0.181)	(0.192)	(0.291)
High	-0.105	0.017	0.123
	(0.193)	(0.204)	(0.303)
Very High	-0.179	-0.133	0.069
	(0.244)	(0.256)	(0.369)
Unknown	0.116	-0.025	0.371
	(0.215)	(0.228)	(0.316)
Gender: *Male*			
Female	-0.191	0.011	-0.244
	(0.117)	(0.124)	(0.181)
Origin: *Overseas*			
Australia	0.050	-0.038	-0.002
	(0.134)	(0.141)	(0.194)
Health: None			
One or More	0.034	0.190	-0.631***
	(0.124)	(0.130)	(0.184)
Education: *School Only*			
University	0.063	0.258	0.184
	(0.151)	(0.161)	(0.223)
Vocational & Other	-0.027	0.302	-0.033
	(0.149)	(0.158)	(0.227)
Age: *16-34*			
35-54	0.661***	0.171	0.416*
	(0.128)	(0.135)	(0.200)
>55	0.984***	0.127	0.410
	(0.190)	(0.202)	(0.295)
Concession: *No*			
Yes	0.094	0.146	0.303
	(0.140)	(0.147)	(0.209)
Location: Unimportant			
Important	0.127	0.094	0.000
	(0.153)	(0.161)	(0.214)
GPGroup: *Unimportant*			
Important	0.051	0.269*	-0.019
	(0.130)	(0.136)	(0.198)
BBilling: *Unimportant*			
Important	-0.388*	-0.203	-0.627**
	(0.178)	(0.184)	(0.230)
Hours: *Unimportant*			
Important	-0.082	-0.020	0.056
	(0.125)	(0.131)	(0.190)
GPChoice: *Unimportant*			
Important	0.311	0.032	-0.007
	(0.169)	(0.174)	(0.224)
Constant	1.265***	0.528	-0.032
	(0.308)	(0.326)	(0.475)

	Multinomial Probit: Four Outcomes		
	Loyal *vs* *Multiple practices*	*Sometimes* *vs* *Multiple practices*	*Multiple single practice use* *vs* *Multiple practices*
n	2,264		
LLH	-1814.439		
Chi-2	3850.210		
p-values	0.000		
Pseudo-R^2	0.092 (367.295)		
AIC	3796.877		
BIC	4277.768		

Notes: * p<0.05, ** p<0.01, *** p<0.0001
Standard errors are shown in parentheses.
The omitted (base) level for each independent variable is shown in italics alongside the category name.
Pseudo-R^2 estimated as 1-(L1/L0), where L1 is the log-likelihood of the fitted regression, and L0 is the log-likelihood of the corresponding constant only regression. Values in parentheses for the Pseudo-R^2 are the LR test statistic estimated as 2*(L1-L0).

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; GP, general practitioner; LLH, log-likelihood; MNP, multinomial probit.

The significant marginal effects for the MNP regression results are presented in Table 16. For a number of variables – *Smoking*, *Area (Inner Regional)*, *Education (Vocational & Other)*, *Health* and *GPChoice* - the pattern of significance for the marginal effects differs from that observed in the coefficient estimates. For example, the *Smoking Unknown* group makes a positive 5.6% contribution to the probability of being a 'multiple practice user' – but does not make a significant contribution to any of the other choice outcomes. This indicates that the negative and significant coefficient estimates observed for *Smoking Unknown* in the MNP regression results in Table 15 reflect the fact that those outcomes were due to the influence of that variable on the other possible choice outcomes relative to being a 'multiple practice user'. In general, the effects of *Age* and *Area* continue to be the most pronounced, as observed in Figure 15.

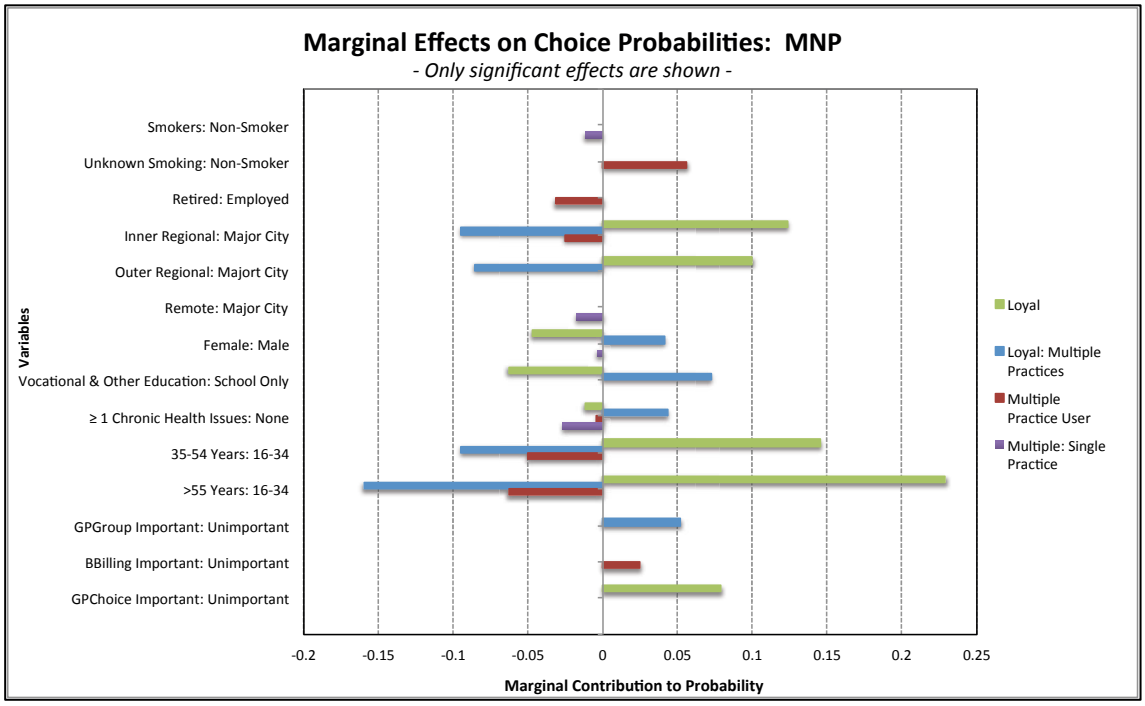
Table 16: Marginal effects - MNP

	Loyal	*Sometimes Loyal*	*Multiple Practice User*	*Multiple Single Practice User*
Smoking: *Non-Smoker*				
Smokers	-0.03 (0.027)	0.025 (0.025)	0.017 (0.013)	-0.012* (0.005)
Unknown	-0.019 (0.037)	-0.028 (0.031)	0.056* (0.023)	-0.009 (0.007)
Employment: *Employed*				
Retired	0.065 (0.035)	-0.023 (0.034)	-0.032* (0.014)	-0.01 (0.007)
Area: *Major City*				
Inner Regional	0.123*** (0.026)	-0.095*** (0.023)	-0.025* (0.012)	-0.003 (0.007)
Outer Regional	0.1** (0.038)	-0.086* (0.034)	-0.006 (0.019)	-0.008 (0.008)
Remote	-0.044 (0.142)	0.036 (0.134)	0.026 (0.085)	-0.018*** (0.004)
Gender: *Male*				
Female	-0.047* (0.02)	0.041* (0.019)	0.011 (0.009)	-0.004* (0.005)
Education: *School Only*				
Vocational & Other	-0.063* (0.026)	0.072** (0.024)	-0.006 (0.012)	-0.003 (0.006)
Health: *None*				
One or More	-0.012 (0.022)	0.043* (0.019)	-0.004 (0.01)	-0.027*** (0.007)
Age: *16-34*				
35-54	0.145*** (0.026)	-0.095*** (0.025)	-0.05*** (0.014)	0 (0.006)
>55	0.229*** (0.031)	-0.16*** (0.029)	-0.063*** (0.016)	-0.006 (0.007)
GPGroup: *Unimportant*				
Important	-0.038 (0.022)	0.052* (0.021)	-0.009 (0.01)	-0.004 (0.005)
BBilling: *Unimportant*				
Important	-0.046 (0.026)	0.035 (0.024)	0.025* (0.01)	-0.013 (0.008)
GPChoice: *Unimportant*				
Important	0.079* (0.031)	-0.053 (0.028)	-0.019 (0.016)	-0.007 (0.007)

Notes: * p<0.05, ** p<0.01, *** p<0.0001
Standard errors are shown in parentheses.
Margins are the discrete change from the base levels, and are evaluated at the means for those variables that achieved statistical significance.
The omitted (base) level for each independent variable is shown in italics alongside the category name.

Abbreviation: GP, general practitioner.

Figure 15: Significant marginal effects - MNP



Abbreviations: GP, general practitioner; MNP, multinomial probit.

Of the measures of meta-health effects, it can be seen that being in the group that considers *GPChoice* to be important increases the probability of being 'loyal', and reduces the probability of being a 'multiple practice user'. Similarly, considering *GPGroup* to be important adds 5% to the probability of being 'sometimes loyal'. Price sensitivity, as observed in the significance of *BBilling*, continues to be observed.

3.3.2.4 Factors influencing patient-GP loyalty: bivariate probit vs MNP

Both modelling specifications identified a role for meta-health effects in the patient-GP loyalty relationship. This was expressed in the form of the ability to choose practitioners, either directly (*GPChoice* was significant in the bivariate probit) or indirectly through the knowledge that the GP operated as part of a group (*GPGroup* was significant in the MP). None of the attitude variables reflecting the influence of convenience factors on the patient-GP loyalty relationship were significant.

However, there were three key differences in terms of the sociodemographic variables of influence identified by the two modelling approaches: *Gender* and *Education* were significant in the bivariate probit, but not the MNP; and *Smoking* was significant in the MNP but not the bivariate probit. It is reasonable to expect that gender and education level might influence patient-GP loyalty, so the absence of those variables as influencers from the MNP specification suggests it might not be the best means of representing the data structure.

3.3.3 Predicted probabilities

The predicted probabilities of belonging to the various patient-GP loyalty relationship categories arising from both estimation methods are presented in Table 17 and are largely consistent with one another. This is to be expected given the construction of the data; that is the four level dependent variable used in the MNP has been generated by collapsing the data used to inform the two independent variables used in the bivariate probit analysis. The remaining analysis of the estimated probabilities is therefore restricted to those produced by the bivariate probit analysis, given that the latter directly reflects the questions derived from the CHERE Survey.

Table 17: Predicted probabilities – category memberships

Loyalty Categories	MNP	Bivariate	Survey Observed Frequencies (%)
Loyal	0.682 (0.139)	0.682 (0.137)	68.00
Sometimes Loyal	0.231 (0.101)	0.230 (0.099)	23.27
Multiple Practice User	0.061 (0.045)	0.062 (0.045)	6.12
Multiple Single Practice User	0.026 (0.031)	0.026 (0.019)	2.61

Note: Values in parentheses are standard deviations.
Abbreviation: MNP denotes multinomial probit.

The influence of the explanatory variables on the patient-GP loyalty relationship was explored by estimating the mean predicted probabilities from the bivariate probit estimation for each of the variables with a significant marginal effect. This provides the probability of being classified into one of the four possible loyalty categories for the subgroups defined by each of those variables, such as age. The mean probabilities of

each subgroup being classified to the various patient-GP relationship categories e.g. 'loyal', are presented in Table 18. Those with the highest probability of being in the 'loyal' category are the retired and those over 55 years of age (showing the influence of *Age*), while those least likely to be in the 'loyal' category are those in the youngest age group (16 to 34 years of age).

Contrary to what might be expected based on the marginal effects, those who felt that *GPChoice* was important have a probability of being in the 'loyal' category slightly lower than the average (suggesting possible interactions with factors that reduce the probability of being 'loyal'). These results support the findings that the only meta-health effect that is influential in determining the patient-GP loyalty relationship is choice. Rather, it is the socio-demographic factors that have primacy in influencing those relationships as can be observed from the graphical presentation of probabilities by sub-group (see Figure 16).

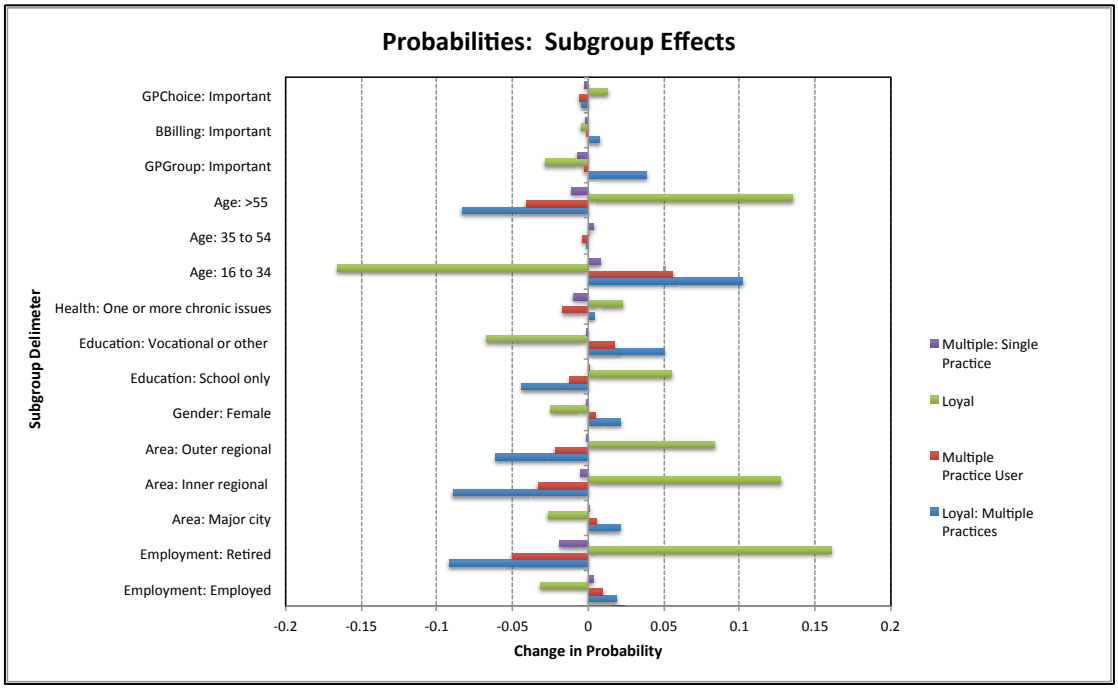
Table 18: Probabilities by sub-group

	Loyal *Mean (s.d.)*	**Sometimes Loyal** *Mean (s.d.)*	**Multiple Practice User** *Mean (s.d.)*	**Multiple Single Practice Use** *Mean (s.d.)*
16 to 34 years of age	0.516 (0.083)	0.332 (0.083)	0.118 (0.038)	0.034 (0.021)
Vocation or other education	0.615 (0.132)	0.281 (0.097)	0.079 (0.047)	0.026 (0.016)
Employed	0.65 (0.126)	0.249 (0.098)	0.071 (0.042)	0.029 (0.017)
GP as part of a group practice important	0.653 (0.14)	0.269 (0.105)	0.059 (0.043)	0.019 (0.014)
Major city resident	0.655 (0.128)	0.251 (0.094)	0.067 (0.044)	0.026 (0.018)
Female	0.657 (0.137)	0.252 (0.101)	0.066 (0.046)	0.025 (0.017)
Choice of GP is important	0.677 (0.138)	0.238 (0.1)	0.061 (0.045)	0.024 (0.018)
35 to 54 years of age	0.682 (0.077)	0.23 (0.073)	0.058 (0.024)	0.029 (0.018)
Overall Group Mean	0.682 (0.137)	0.23 (0.099)	0.062 (0.045)	0.026 (0.019)
Bulk-billing is important	0.695 (0.133)	0.226 (0.099)	0.056 (0.041)	0.024 (0.017)
One or more chronic health issues	0.705 (0.135)	0.234 (0.104)	0.045 (0.034)	0.016 (0.01)
School education only	0.737 (0.12)	0.186 (0.083)	0.05 (0.041)	0.027 (0.022)
Outer regional resident	0.766 (0.103)	0.169 (0.073)	0.04 (0.033)	0.025 (0.019)
Inner regional resident	0.809 (0.091)	0.141 (0.068)	0.029 (0.023)	0.021 (0.016)
55 years or older	0.817 (0.062)	0.147 (0.053)	0.021 (0.014)	0.015 (0.013)
Retired	0.843 (0.055)	0.139 (0.05)	0.012 (0.007)	0.007 (0.004)

Note: Data are presented in ascending rank order for the mean probabilities of each group belonging to the category 'loyal'.

Abbreviation: GP, general practitioner; MNP, multinomial probit; s.d., standard deviation.

Figure 16: Subgroup probability contributions



3.3.4 Model comparisons and sensitivity analyses

Qualitatively, the results for the bivariate probit and MNP specifications are similar. Despite this consistency, the bivariate model appears to be the more favourable of the two on the basis of the AIC and BIC (as reported in Table 13 and Table 15); even in light of the higher Pseudo-R^2 values for the MNP. In addition, the results from the bivariate analysis are more amenable to interpretation because the dependent variable used is derived directly from the source data within the CHERE Survey without added manipulation (compared with the dependent variable for the MNP that is derived from the two questions in the CHERE Survey used to populate the bivariate probit analysis). While it is reasonable to infer that the four outcomes modelled in the MNP represent the same patient-GP loyalty relationships as in the bivariate probit model, it is possible that individuals would have characterised their loyalty relationships differently if they had been asked to choose between four options (as have been constructed for the MNP) within a single question.

Moreover, the MNP results presented have not been constrained as suggested by Keane (1992)[207] and are therefore potentially fragile (noted by their large s.e.). This was apparent in that the initial specification using the restricted socio-demographic variables would not converge unless a starting matrix of coefficient values was input for the maximisation procedure. The values used were those obtained from the initial estimation sequence from the bivariate probit analysis, which was executed without specifying a convergence error. Thus while the MNP presents an intuitively attractive option for the modelling of the patient-GP loyalty relationship, the bivariate relationship appears more stable.

3.3.4.1 Sensitivity analyses

Four sensitivity analyses were conducted to test the robustness of the results to different modelling approaches: (1) that the dependent variable could be expressed as a three level outcome ('loyal', 'sometimes loyal', 'multiple practice user'); (2) that the assumption of IIA holds, and that the data could be modelled using an MNL; (3) missing observations could have been excluded, or substituted using extreme values; and (4) that only data from individuals with two or more GP visits in the previous 12 months should be included in the analysis of the patient-GP loyalty relationships. The results for these analyses are presented in Appendix 5.

Reshaping the dependent variable in the MNP to three levels resulted in four key changes: *Smoking* status was no longer statistically significant; chronic health issues significantly influenced being 'loyal' (as well as 'multiple practice use'); both choice variables (*GPChoice* and *GPGroup*) were significant; and *BBilling* was no longer significant. Comparing the results from the three level MNP to those from the bivariate probit, it can be observed that *Smoking* is not significant in either specification, and that the remaining socio-demographic variables of influence were the same in the two specifications. However, they differ with respect to the influence of the attitude variables; namely price sensitivity (*BBilling*) is no longer significant in the three level MNP. The absence of evidence for price sensitivity in the three level

MNP is of concern given previous research which suggests individuals do consider some indicator of price (either in terms of the fee charged, or out-of-pocket costs) when choosing between GPs.[94,172,177,184,196] This provides further support for the use of the bivariate probit specification over the MNP, even if the latter is described using a three level outcome rather than four.

The multinomial outcome could also have been modelled as an MNL, which would potentially have been more stable. Indeed, running the multinomial analysis as an MNL produced qualitatively the same results as the MNP, without violating IIA (see Table A 6, Appendix 5). However, the choice of the MNP for the initial estimations allows for a more direct comparison with the analysis estimates produced by the bivariate probit.

The impact of missing observations was tested by recoding individuals with missing data initially classified to the Unknown category in three ways: first to the lowest category for each variable (labelled the 'Lower Bound' analysis); second to the highest category for each variable ('Upper Bound'); and, finally excluding them altogether ('Complete Case'). The resulting probability estimates were robust to the method used to classify missing observations; the largest differences produced were at the third decimal place for the Complete Case analysis (Table A 7).

The final sensitivity analysis investigated how the modelled relationship might alter if the sample was limited to those individuals who visited a GP practice two or more times within the 12 months prior to the CHERE Survey. Restricted to the bivariate probit specification, the results were largely consistent with those of the main analysis (see Table 13) in that *Age, Area, Education* and *BBilling* all influence the patient-GP loyalty relationship. However, *Gender* and *GPChoice* were no longer significant, while *Smoking* (status unknown) and *GPGroup* became significant. Arguably, the main change is that *GPChoice* is no longer significant since *GPGroup* was close to being significant in the main analysis (in the analysis for those with at least one visit, p=0.059

for *GPGroup*). It is possible that for some individuals these two variables are capturing the same effect of being able to have a choice of GP when they go to a practice.

3.4 Discussion

Previous analyses of patient-GP loyalty relationships have typically focused on socio-demographic determinants, based on individuals' characteristics or what they have experienced in past visits. In this chapter, these analyses were extended to explore how the importance individuals place on certain aspects of the experience of care influenced the patient-GP loyalty relationship. The influence of the experience of care was assessed using individuals' attitudes to the meta-health effects of convenience and choice.

The impact of GP practice location, appointment time, practitioner choice and type explored in this chapter have been previously included as attributes in a number of DCE studies investigating the factors associated with GP choice.[94-96,177-182,184,185,187,190-193] These factors have typically been found to be positive determinants of GP choice; individuals are more likely to choose GP visits or GP practices that are close by[181,191,192], have flexible opening hours[96,181,191], and where they can choose their practitioner.[94-96] The research in this chapter confirms these findings with respect to the influence of being able to choose a practitioner on the choice of patient-GP loyalty relationship, but does not find an influence for convenience factors. It does so using attitudes (ratings of importance) of these meta-health effects rather than preferences.

A number of these DCE studies explicitly address the link between meta-health effects and patient-GP loyalty, notably through trade-offs between convenience (in terms of the appointment time or opening hours) or choice (of practitioner), and some element of continuity in the patient-GP relationship.[184,188-190] These analyses suggest some interesting trade-offs between meta-health effects and patient-GP loyalty. In a survey of the Swedish general population, Hjelmgren and Anell (2007)[190] find that individuals are more likely to choose models of primary care in which they have a choice of

provider. However, the influence of establishing continuous relationships through registration was observed only for individuals for whom the nearest hospital was between 10 and 30 kms away (the furthest distance possible).[190] It is possible that in this instance, individuals interpreted registration as being a curb on their ability to choose a practitioner, rather than as an indicator for continuity of care. In contrast, Hole (2008)[188] and Cheraghi-Sohi et al. (2008)[184] find that when choosing between different GP consultation types, individuals in the UK are willing to pay more in terms of being able to see a practitioner who knows them well (reflecting continuity) than they are to secure an appointment at the time of their choosing (reflecting convenience). The results from Zickafoose et al. (2015)[189] suggest that individuals place a higher value on continuity, in terms of always seeing the same physician for the care of their children, than other aspects of primary care services (such as opening hours, visit type, and the availability of phone consultations) with the exception of being always able to access same day visits for medical care (rather than prevention).

The evidence from the DCEs reviewed is that individuals place more emphasis on aspects of being loyal to a GP, compared with the meta-health effects of convenience or choice. However, these analyses investigated these factors as covariates in choosing between GP appointments or types of practices, rather than investigating one as a possible influencer of the other. Moreover, they rely on stated preferences only, without including information on past behaviour.

While it would be possible to construct a DCE to investigate the nature of patient-GP loyalty relationships, the existence of the CHERE Survey provided an existing data source with which to examine those relationships. From that survey, individual-reported data on past experiences, a measure of revealed 'preference', are used in this chapter to measure patient-GP loyalty relationships. These data are combined with stated attitude ratings data on the importance of meta-health effects in driving GP choice. In so doing, the analyses in this chapter not only address the role of meta-

health effects in affecting the patient GP-loyalty relationship, but also whether those effects can be captured by attitudes.

Two different estimation methods have been used to explore the patient-GP loyalty relationship: a bivariate probit and an MNP. While the qualitative results from these specifications are largely consistent, the bivariate analysis appears to reflect the underlying data better, and be more amenable to estimation and interpretation. This suggests that a bivariate probit estimation can represent the patient-GP loyalty relationship better as measured by the cross-sectional sample used in this chapter. In terms of choice behaviour, the relationship in the bivariate probit implies that individuals first choose whether or not to be loyal to a GP, and having made that choice, choose whether or not to use multiple practitioners.

The main results from the bivariate probit indicate that being older and retired, and living outside a major city, are associated with a higher probability of having a usual GP practice, or being loyal. Those who are younger, or price sensitive, have a lower probability of being loyal and a higher probability of multiple practice use. By and large, these socio-demographic influences on the patient-GP loyalty relationship outweigh those of the meta-health effects. Only one meta-health effect was significant in affecting the patient GP-loyalty relationship; respondents who felt that having a choice of GP was important had a higher probability of being loyal (approximately 65%). This suggests that such individuals might be willing to be restricted to using a single GP practice provided that they had the freedom to see the GP of their choice within such a practice. Finding that choice rather than convenience was important in this context is consistent with what has been observed using stated preference studies[184,188,189] and qualitative research.[176]

3.4.1 Limitations

The data from the CHERE Survey used for the dependent variable in this analysis were derived from two questions that asked respondents about their prior decisions about

having a usual GP practice (remaining loyal) or using multiple practices. However, individuals will also make loyalty decisions at the GP level. Within the CHERE Survey, the level of concordance between individuals who stated they had a usual practice and a usual GP was 85%. The survey did not provide further information on whether individuals also saw other practitioners. Use of the question on the usual GP, rather than the usual practice, as the dependent variable would have limited the analysis of the patient-GP loyalty relationship to a purely binary one and not permitted the identification of those who are by and large "loyal" but exercise the option available within the Australian system of visiting multiple practices and practitioners.

In addition, information was not available on other GP practices from which respondents might have selected when making their decisions about whether to use multiple practices. The availability of more information about the practices chosen by individuals, and others from which they could have chosen, would allow the inclusion of practice specific variables into such modelling, potentially enhancing its relevance. Moreover, while the respondent sample was restricted to those with at least one visit, the main conclusions from the bivariate probit were supported in a sensitivity analysis restricted to those with at least two prior GP visits (see Appendix 4). Some differences in the influence of the attitude variables measuring the meta-health effects among those respondents suggest that following studies might consider how such attitudes change with increasing GP use.

The attitude variables included in the analyses were selected on the basis of those which corresponded most closely to the reasons reported by most respondents for using more than one GP practice in the 12 months prior to the survey. It is possible that the reasons included do not capture all the factors that are important to individuals when making a choice about ongoing GP relationships.

The usefulness of the data from the CHERE Survey to investigate the question of patient-GP loyalty is limited by its cross-sectional nature. However, the combination

of information available from the survey with respect to individual socio-demographic characteristics, GP utilisation behaviours and attitudes makes it ideal to investigate the question at hand. While there were some differences between the individuals included in the CHERE Survey and the Australian population; largely in terms of age (older in the CHERE Survey), and income (higher), the latter was not significant in this analysis. Age was found to be strongly associated with the patient-GP loyalty relationship, but the difference in the distribution of age between the CHERE Survey and the Australian population does not impact on the external generalisability of these results since the influence of between age group differences would still be anticipated to hold.

3.4.2 Conclusion

The results from the analyses presented in this chapter suggest that the patient GP-loyalty relationship as informed by data from a survey of Australian individuals can be modelled using a bivariate probit relationship. Age, area of residence, education, gender, being retired and the presence of chronic health issues all influence how individuals choose to relate to GP practices. This is consistent with what has been previously identified in the literature for Australia.[170,201] Where individuals are price sensitive, motivated by the ability to access bulk-billing, multiple practice use is more likely to ensue. This finding is of particular relevance in a policy environment where the potential introduction of a co-payment for GP visits would reduce access to bulk-billing.[218] The meta-health effect of choice was also a significant determinant of the patient-GP loyalty relationship but had less influence than the key socio-demographic factors. Convenience, as expressed by GP practice location and the availability of appointment hours, was not a significant influencer of the patient-GP loyalty relationship.

The results in this chapter add to the literature on patient-GP loyalty relationships by exploring the potential influence of meta-health effects, and by doing so through the use of expressed attitudes towards those effects. The findings indicate that individuals who place importance on the ability to choose their practitioner, or on having multiple

practitioners available within a practice, have a higher probability of forming links with a usual practice. Exploring further interactions in the link between attitudes, socio-demographic characteristics and patient-GP loyalty would help to inform potential changes in how GPs are funded – such as expanded patient enrolment or capitation arrangements. The results described in this chapter indicate that for some individuals, the accompanying loss of choice associated with capitation arrangements say would adversely influence the nature of subsequent patient-GP loyalty relationships for some individuals.

4 Stated Preference Methods for Exploring the Value of Meta-Health Effects: DCEs

Chapter Summary

Recent decades have seen an increase in the use of DCEs for the assessment of patient and community preferences associated with health care interventions. This chapter provides an overview of the steps involved in the design and analysis of DCEs. It outlines the process of survey design, including the choice of attributes, experimental design, and the formation of choice tasks for DCE studies. How data from choice tasks are analysed to quantify respondents' preferences is subsequently described. This focuses on the theoretical underpinnings of analysing choice data and the analysis methods that have been applied in this thesis.

4.1 Introduction

The use of DCEs to assess preferences associated with health care interventions has increased over the last 15 years.[44,45] In recent years, this has seen the release of guidelines focused on the design, analysis and reporting of such experiments in a health care setting.[47,219,220] DCEs describe a good or service by its component attributes, such as the mode and frequency of treatment administration. Those attributes are combined into alternative combinations, or profiles. Individuals completing DCE surveys make repeated comparisons of alternative profiles, stating their preferred profile each time. The analysis of the choices individuals make reveals how they trade-off between the attributes that make up the profiles, and provides an estimate of the value of those attributes.[47-50]

There is an intuitive appeal to the use of DCEs to measure the value associated with meta-health effects insofar as they allow the outcomes of interest to be described explicitly as attributes within the choice profiles within a DCE study. Thus, within a DCE health and meta-health effects can be included as separate attributes within the same valuation task to assess the relative importance of health compared with meta-health effects.[9] Examples of the use of DCEs to assess preferences for meta-health

effects include preferences for primary care services[4,94-96,179,181,184,185,187-193], assessment of health promotion[10], the need for blood transfusions[145], continuity of intrapartum care[221], the acceptability of the use of growth hormone[222], modes of administration for RA[90,223], and follow-up monitoring programmes for women with breast cancer.[224,225]

This chapter provides an overview of the practical aspects of constructing a DCE, such as DCE design, the nature of the choice problem and the options considered by respondents. The modelling approach to the analysis of choice data is then presented. Specific details of the application of these methods to the DCEs to investigate preferences for meta-health effects in the context of decisions for the treatment of RA (RA Therapy study) and the ongoing management of breast cancer risk (Mastectomy study) are presented in Chapter 5 and Chapter 6 respectively.

4.2 Constructing DCEs

There are three key steps when constructing DCEs: choosing the attributes and levels for inclusion; designing the experiment; specifying the choice task and the choice question. These are described in the following sections.

4.2.1 Attribute choice

The choice of attributes and levels for a DCE is a critical step as it defines that for which individuals will express a preference. The choice of attributes can be informed by existing literature or by qualitative research with relevant individuals to uncover the drivers of decisions in the context of interest.[47,48,50] The number of attributes (k) should be sufficient to reflect what is relevant to that choice situation, but not so large as to overly burden respondents.[xix47] Similarly, the number of levels (l) per attribute and their range should be broad enough to encompass situations faced in a real-world situation and to induce respondents to trade between profiles.[47,48,50] The selection of the

[xix] Profiles that are overly complex due to the number of attributes run the risk of non-response, or respondents engage in simplifying heuristics e.g. they choose on the basis of one or two attributes only, or always choose one profile regardless of the attributes on offer.[47]

attributes and their corresponding levels for the RA Therapy and Mastectomy studies is detailed in Chapters 5 and 6 respectively.

4.2.2 Experimental design

Once the attributes and their levels have been specified, the next step is to determine how they will be combined in profiles for inclusion in the DCE survey. This relies on the use of an experimental design to generate the choice profiles that form the basis of the choices respondents make.

The total possible number of profiles that can be formed by the combination of attributes and levels for a given design is known as the full factorial.[226] Full factorial designs allow for all the main effects, those due to the attributes alone, and interactions, the combined effects of attributes, to be estimated.[226] For a design of k attributes each of l attributes, the full factorial is produced by the product of each level (l) raised to the power of the k attributes with that level.[47,50] For example, the full factorial for a design with three two level attributes and one four level attribute is $3^2x4^1 = 36$ profiles. To produce a DCE survey, those profiles must be combined into choices sets. The required number of choice sets increases with the size of the full factorial design depending on the number of alternatives to be offered.[47,50] For example, having two alternatives in a choice set with 36 combinations in a full factorial would result in 630 possible choice sets.

While full factorial designs have the potential to provide a great deal of information, it is standard to reduce this to statistically meaningful but more manageable designs called fractional factorial designs.[226] For any given DCE, the smallest possible design is determined by the number of parameters to be estimated in the analysis of the choice model; there must be the same number or more choice sets than parameters (coefficients and standard deviations) to be estimated.[50]

There are essentially two approaches by which the combination of attributes and levels from the full factorial design can be reduced for specifying the design for use in a DCE. The first is the use of existing catalogues of orthogonal main effects plans (OMEPs). These are pre-designed mathematical combinations of possible attributes and levels from which an appropriate design can be selected for the specific discrete choice question.[47,226] Once a given OMEP has been chosen, this can then be used to produce the required number of choice sets by randomly pairing profiles, shifting attribute levels within the design to create a new set of profiles, or selecting alternate profiles from multiple OMEPs.[47,226]

However, the use of catalogue derived OMEPs can lack flexibility; there might not be existing OMEPs to cover the combination of attributes and levels for a specific choice task.[47] Greater flexibility is offered through the use of computer based design software. These software generate designs by continued iterations from a pre-specified OMEP, such as those in SAS, or Street and Burgess designs, or by testing alternative attribute and level combinations to describe pre-specified utility functions, such as in the case of Ngene.[47]

The goal for computer aided experimental designs is to identify designs that are efficient; those which minimise the variance associated with the choice estimates, best characterised by the D-error.[xx226-228] Huber and Zwarina (1996)[228] note that efficient designs have a minimal D-error and possess the following four properties: level balance - all levels of each attribute appear the same number of times; orthogonality – the levels of each attribute are independent[xxi]; minimal overlap – the probability that an attribute level appears more than once in a choice set is as low as possible; and, utility balance – alternatives within a choice set appear equally attractive to respondents, one does not dominate the other.[226-228]

[xx] D-error is estimated by the determinant of the covariance matrix on the choice coefficients, raised to the power $1/k$, where k is the number of attributes.[227]

[xxi] In the context of choice experiments, the property of orthogonality relates to independence of the choice estimates for each attribute.[226]

For the research in this thesis Ngene has been used to generate the experimental designs on the basis of minimising the error variance associated with the ratio of two coefficients; described as WTP-efficiency.[47,229] The process for generating and assessing the efficiency of the resulting designs within Ngene is provided in appendices supporting those studies (see Appendix 7 and Appendix 12).

4.2.3 Choice profiles

The choice tasks used in both the RA Therapy and Mastectomy studies follow a common approach for DCEs; in each choice set, respondents are asked to state which of the profiles on offer they prefer. An alternative approach is to ask respondents to state which of the profiles on offer they consider to be the best and which the worst; called best-worst scaling.[230] This form of preference elicitation can be applied to compare two or more goods or services described as complete profiles of attributes (so called 'case 3' profiles). Proponents of best-worst scaling techniques suggest that the preference information derived using standard DCE approaches or case 3 best-worst approaches can also be extracted using single profile tasks.[33,231] In such tasks, individuals are presented with single profiles describing a good or service and asked to choose the best and worst attribute. Where the attributes in those profiles appear without levels, it denotes a 'case 1' experiment, whereas in 'case 2' experiments the levels on the attributes are included.[232]

The potential appeal of single profile methods (case 1 or case 2) is that they might be easier for respondents to complete because they consider the attributes only and do not make comparisons between profiles.[230] However, Yoo and Doiron (2013)[233] find that case 2 and case 3 valuations do not produce the same values for a given attribute of interest. For example, Yoo and Doiron (2013)[233] find that non-salary based outcomes were more important in choice decision in the case 2 (trade between attributes) than in case 3 (choice between profiles) situation. They postulate that this might be due to

differences in how much attention respondents pay to attributes, reference points, or to the very nature of the question in the two cases.[233]

There are potentially lessons from this example in relation to the form of choice scenario to be used to explore the value of meta-health effects using DCEs. The value of meta-health effects might be expected to differ depending on whether individuals are evaluating them as part of an overall profile (or good) that forms part of a trade (case 3), or as an individual attribute whose importance they are asked to assess relative to other attributes (case 2). It is likely that meta-health effects would be given more importance in a case 2 experiment (a trade between attributes) than a case 3 experiment (a trade between product profiles in which the meta-health effect is just one attribute).

Moreover, it is unclear whether the valuations derived from case 2 experiments elicit preferences in a manner that would be consistent with informing economic evaluations for decision-making. In their assessment of conjoint analysis, Louviere et al. (2010)[114] note that unless a ranking is made of a complete profile it is not consistent with collecting data that fulfil the requirements for choice under neoclassical utility theory:

> "*they collect data in ways that cannot be analysed to be consistent with neoclassical economic theory because ratings and attribute importance measures do not readily translate into choice or matching..on which utility theory is based (p61).*"[114]

Thus, given the potential for case 2 methods to over-emphasise the value of meta-health effects, and that there is some question as to whether they capture value in a manner that is consistent with its use to inform economic evaluations for decision-making, complete profile comparisons are the appropriate method to use in the context of the research undertaken for this thesis.

4.2.3.1 Labelled or unlabelled designs

The other aspect to consider in designing the DCE survey is whether the alternatives on offer in each choice set will appear labelled or unlabelled.[48,50,234] Labelled designs assign a specific name to the alternatives within choice sets, and those names might be anticipated to reflect some aspect of respondents' preferences beyond what is captured in the attributes themselves. In the context of health care, these might be brand names of medications, or the names of conditions e.g. cancer or heart disease. Unlabelled designs use generic terms to differentiate between alternatives e.g. option 1 or 2, and are not anticipated to influence respondents' preferences.[48,50,234] The use of labelled and unlabelled designs in the RA Therapy and Mastectomy studies is discussed in Chapters 5 and 6 respectively.

4.2.4 The choice question

In any comparison of two or more profiles, the respondent must be asked a question about which profile they prefer or would choose. This raises the possibility that the choice question in a DCE could be re-framed to capture more explicitly the meta-health effect under consideration. Marti et al. (2012)[154] suggest that the preference elicitation question in best-worst tasks can be restated to reflect any concept in which the analyst is interested, provided they can be located on a similar 'scale' (e.g. from bad to good). This suggests that rephrased best-worst DCE tasks could be used to estimate preferences by creating a scale specific to a meta-health such as reassurance, hope or trust (e.g. *Which of these options makes you feel most reassured?*). However, such a question might unduly restrict what is being maximised by the individual in making their choice. For example, including a question about which profile leads to the greatest reassurance assumes that the individual is concerned with reassurance, rather than allowing the choice of profiles to reveal what is being maximised. Moreover, it has the potential to exacerbate the impact of 'focusing' in affecting the values derived, where focusing refers to drawing the respondents' attention to the concept of interest, thereby inflating its importance relative to other sources of value. To avoid the potential for focusing, and of inappropriately identifying what is being maximised by

respondents, the choice questions in the DCEs in both Chapter 5 and 6 are: "*Which one of these options would you choose?*"

4.2.4.1 The use of forced choice

Choice tasks within DCEs can be presented as a forced choice, or can include an opt-out that allows the respondent to choose the status quo, or neither option, over the alternatives on offer. While having an opt-out might be relevant for predicting choices in a market, one disadvantage is that given an opt-out option some individuals might always choose to 'opt-out' or not trade. Such responses would provide either limited or no information in the context of this research on how individuals value meta-health effects relative to health effects. Thus forced choices, in which respondents have to choose between the options presented in each scenario, have been used for the DCEs in the RA Therapy and Mastectomy studies in order to capture individuals' preferences between attributes for health and meta-health effects.

4.2.5 DCE sample size

One benefit of the repeated measures nature of the data within a DCE is that fewer participants are required to provide sufficient variation to examine the relationships of interest within the data. Guidance from the literature is that as few as 20 respondents per choice set are required to provide estimates for simple relationships between the choice task and the attributes included.[48] However, investigating more complex relationships involving covariates, interactions between attributes and/or covariates, or comparing across survey versions, requires an increase in sample size to allow for sufficient variability to estimate the relationships within the data with precision (minimising the size of the confidence intervals around the coefficient estimates, and underlying measurement error due to how respondents understand the choice task and its complexity).[47]

Results of simulations using data from a number of DCE experiments suggest that the benefits to improving precision begin to diminish at sample sizes above 150 respondents, and become negligible above 300 respondents.[47] That is, the benefits of increasing the sample size to the precision with which relationships between the data are estimated begin to diminish after 150 respondents, and stabilise after 300.[47] Both the RA Therapy and Mastectomy studies utilise three versions of their respective designs to test for the presence of framing effects. Accordingly, both studies have a target sample size of 150 respondents per version, making 450 in total for each study.

4.3 DCE Model Specifications

The underlying premise of the stated preference experiments used in this research is that faced with a choice between two or more health profiles or scenarios, individuals will choose the alternative from which they expect to derive the greatest utility. This is under pinned by Lancastrian choice theory which states that the utility of a good or service can be decomposed into the utility of its underlying attributes.[46] The combination of attributes which is expected to deliver the highest utility is preferred by the individual.[46] In making their choices between alternatives, the resulting preferences reflect systematic or explainable preferences for the attribute, and an element of random error which can not be explained.[48] This is expressed as a utility function of the type described by McFadden's random utility theory:[48,235,236]

$$U_{ijt} = V_{ijt} + \varepsilon_{ijt} \quad \text{i=1,..I individuals, j =1..J alternatives, t = 1,..T choices} \quad (10)$$

where U is the utility associated with the choice of product, V is the vector of observable variables that explain that choice, and ε the vector of unobservable factors, or random error term.[48,235] These apply to the ith individual, over J alternatives, for the tth decision out of T choices.

The simplest model with which to estimate this form of choice model is the conditional logit regression, a special case of the multinomial logistic regression, in which choice is

between two options.[235,236] Application of the conditional logit regression assumes that the errors are independently and identically distributed (iid), and that tastes are homogenous across individuals (leading to a restriction that there is an independence of irrelevant alternatives across the data[235,236]), expressed by the following:

$$U_{ijt} = \beta x_{ijt} + \varepsilon_{ijt} \qquad (11)$$

where x are the attributes and covariates expected to influence choice and β the vector of homogenous attribute coefficients: β can vary between attributes, but for a given attribute is assumed to be the same across individuals. This specification does not include a constant term, which is appropriate for situations in which the experiment uses a forced choice design. In the absence of an alternative specific constant in a labeled experiment, interpretation of a constant term would be difficult. This is discussed further in Chapters 5 and 6 in the context of the research specific DCE designs.

As summarised by Keane (2014)[236], for each choice option, *j*, the probability of choice has a closed form solution of the standard form:

$$P(j|X_{it}) = \exp(\beta x_{ijt}) / \sum_{j=1}^{2} \exp(\beta x_{ijt}) \qquad (12)$$

4.3.1 Preference heterogeneity

While the analysis of choice data using conditional logit regression can incorporate adjustment for repeated observations over the same individual (*i*), it imposes the assumption of homogeneity of preferences across β. However, homogeneity in preferences between individuals is unlikely to hold, so it is appropriate to use more complex methods of analysis that take account of between individual differences in the impact of each attribute on choice.

4.3.1.1 Latent class analysis

One method of exploring preference heterogeneity is to assume that individuals can be grouped based on the similarity of their behaviour with respect to choice. Latent class analysis proceeds on this premise, that choice can be investigated by taking into account the likelihood that individuals can be allocated into segments or groups based on the similarity of their choice behaviour. The probability that an individual will choose a particular option is therefore conditional on their membership to a group and can be summarised by modifying expression (12)[237]:

$$P(j/g) = \exp(\beta_g x_{ijt}) / \sum_{j=1}^{2} \exp(\beta_g x_{ijt}) \qquad (13)$$

This shows that the probability of an individual choosing option *j* is conditional on their membership of group *g*, of which there are 1 to *G* groups to be determined by the model performance criteria, and the attribute coefficients vary between groups.[237] Estimating the probability of choice is essentially via the conditional logit regression, allowing for differences in the choice coefficients between groups.[236-238] Latent class analyses can predict, for each individual in the analysis, the probability that they belong to each of the *G* classes. Latent class analysis is applied to the DCE data from the Mastectomy study in Chapter 6.

4.3.1.2 Mixed logit regression

The mixed logit regression relaxes the assumption of preference homogeneity, and explicitly introduces an element for individual specific variability to account for preference formation. This can be expressed by modifying equation (11) to decompose the attribute coefficients into an overall 'mean value', delimited by β for each attribute and the individual specific variation (or variance) from that mean, η. The errors in the resulting equation, (14), are still assumed to be iid.[236]

$$U_{ijt} = (\beta + \eta_i) x_{ijt} + \varepsilon_{ijt} \qquad (14)$$

Mixed logit regressions are applied in the analysis of the DCE data from both the RA Therapy study (Chapter 5) and the Mastectomy study (Chapter 6).

4.3.1.3 Generalised multinomial logit

While the mixed logit model allows for individual heterogeneity in the analysis of choice, it does not account for potential differences between individuals in the unexplained variability in how differently they consider attributes. That is, such models do not correct for the presence of 'scale' factors in the error variance.[48,236,239] Here, scale refers to unobserved shifts (up or down) in the error variances around the coefficients that potentially confound observed choices; individuals may be more or less variable in the choices they make, but the extent of that variability is not observed. Such variability might arise due to differences between samples answering a DCE, differences in the description and level of attributes within a DCE, or both.[240,241] The potential for scale to be influenced by differences in attribute levels and descriptions implies that the detection of scale effects between samples might indicate framing effects.

Fiebig et al. (2010)[239] have expanded on the mixed logit model of individual choice to take account of scale, introducing the generalized multinomial logit (GMNL).[xxii] Equation (14) can be re-expressed to show the presence of individual specific scale, σ_i:

$$U_{ijt} = (\beta + \eta_i)x_{ijt} + \varepsilon_{ijt}/\sigma_i \qquad (15)$$

Ordinarily, this scale factor σ_i cannot be observed, so it is normalised to one to allow identification of the mixed logit model.[236] Within the GMNL, Fiebig et al. (2010)[239] directly account for the influence of scale on the estimation of the individual coefficients and their error variances, allowing for alternative modelling structures

[xxii] Hess and Rose (2012) caution that the GMNL and similar models which attempt to separately measure scale fail to do so because they conflate the scale associated with the choice coefficients and the error terms.[242]

depending on how scale, σ_i, relates to the coefficient estimates. This is expressed by an additional scaling parameter, γ (typically bounded by 0 and 1), the value of which determines how the variance scaling factor σ_i relates to the coefficient estimates.[239] Rearranging (15) to remove scale from the errors, the manners in which scale can relate to the coefficient estimates are captured by: [236,239]

$$U_{ijt} = [\sigma_i\beta + \gamma\eta_i + (1-\gamma)\sigma_i\eta_i]x_{ijt} + \varepsilon_{ijt} \quad (16)$$

Where γ=1, it can be seen that (16) becomes:

$$U_{ijt} = [\sigma_i\beta + \eta_i]x_{ijt} + \varepsilon_{ijt} \quad (17)$$

This is the GMNL-I in which only the mean attribute values are scaled (scale $\sigma_i\beta$) while the individual variations (η_i) maintain a constant (unscaled) variance.[236,239] Where γ=0, it can be seen that (16) becomes:

$$U_{ijt} = \sigma_i(\beta + \eta_i)x_{ijt} + \varepsilon_{ijt} \quad (18)$$

which is described as the GMNL-II, in which scaling applies to both elements of the coefficient estimate; the mean estimate β and individual variations from that mean (η_i).

Specifying GMNL-I or GMNL-II imposes different assumptions on the treatment of preference heterogeneity. Where GMNL-I is specified, it implies an independence between the individual specific scale (σ_i) and the variance of the preference heterogeneity captured in the individual coefficient variation components (η_i).[239,243] Where GMNL-II is specified it implies a proportionality between scale and preference heterogeneity.[239,243] Keane and Wasi (2013)[236], note that it is not necessary to specify gamma and that, functionally, it can be allowed to vary to reveal the underlying data preference structure.

GMNL is applied to the analysis of the DCE data from the RA Therapy study (Chapter 5).

4.3.2 Estimating mWTP

Both the RA Therapy (Chapter 5) and Mastectomy (Chapter 6) studies include cost attributes in their DCEs. Estimates of mWTP can thus be formed as the ratio of the coefficient of interest to that of the cost coefficient.[48,244]

Fiebig et al. (2010)[239] recommend that where estimation of coefficients is performed via a random parameters model (such as a mixed logit or GMNL), the calculation of the mWTP should take into account the distribution of the resulting estimates and be estimated via simulation rather than the ratio of point estimate of means. However, the process of estimating mWTP via simulation requires that distributions be specified for both parameter distributions. There are no standard rules for the choice of distribution, i.e. whether it should be the same distribution for the two parameters used in estimating mWTP (price and the relevant attribute being traded).[245,246] Moreover, the use of different distributions (e.g. triangular compared with normal) will result in different estimates of mWTP.[245,246] In addition, Daly et al. (2012)[246] note that while simulations of mWTP will typically be conducted using multiple draws, each new simulation (consisting of multiple draws) will result in different estimates. Given the difficulties in choosing an appropriate distribution, and the potential for multiple simulations to produce different results, Daly et al. (2011)[247] conclude that in the case of estimating simple ratios of coefficients produced by DCEs (as is the case in estimating mWTP), simulation is not required and indeed will not produce reliable estimates of the standard errors of the estimates. Thus, the approach in this thesis is to estimate mWTP without simulation, as the ratio of the coefficients on the attribute of interest to that of cost. Confidence intervals of the mWTP were estimated using the Delta method, which has been preferred to other methods (such as the Krinksy Robb

estimation method) when estimating ratio values (such as mWTP) using random parameter models.[248-250]

Lancsar et al. (2007)[251] suggest that one method of investigating relative attribute impact is to rank order the resulting mWTP estimates produced using the choice coefficients derived from those attributes. Such rankings are likely to be specific to each choice task under consideration since they will be influenced by the levels of the cost attribute and the cost-elasticity of the other attributes in the task. Using this method, rankings of mWTP were produced within both the RA Therapy and Mastectomy studies, with the largest absolute mWTP as the reference (set to one), and all other mWTP expressed as a ratio of that value. While the DCEs in the RA Therapy (Chapter 5) and Mastectomy (Chapter 6) studies contain different attributes, and costs with different levels, comparisons of their ranked mWTP are relevant in terms of the ordering of health or meta-health effects.

4.3.3 Capturing the effects of demographics

As well as the design attributes in a DCE, individual specific factors such as age, gender, income and education are likely to influence the choices made by individuals. Typically, the influence of demographic variables on choices is assessed by interacting the demographic variable of interest, say income, with an attribute which it is likely to affect, say costs.[56,229,235] For a given individual, there will be variability in the attribute levels they observe for cost (the design attribute) over the choice tasks they complete. Thus, including an interaction of cost and income would allow the effect of income on choice to be modelled.

This approach to the inclusion of demographic variables has not been pursued for the analyses presented in Chapters 5 and 6 for a number of reasons. The first is due to the number of possible interactions required by such an approach compared with what was included in the original experimental design. Rose and Bleimer (2006)[252] show that

there is a loss of efficiency when estimating choice models that include demographic effects that have not been captured in the underlying experimental design. Neither the experimental design for the RA Therapy study (see Appendix 7) nor that for the Mastectomy study (Appendix 12) included interactions for demographic effects. Second, both experiments have sample sizes in the order of 450 participants. These would be unlikely to be sufficient to allow the detection of interactions with multiple demographic variables, as well as the main effects being investigated in this research. Accordingly, alternative approaches have been used to explore the role of demographics in choice in both the RA Therapy and Mastectomy studies. These are described in the methods section of the respective chapters.

5 Meta-Health Effects in Treatment Choice: the Role and Value of Convenience in Therapies for Rheumatoid Arthritis

Chapter Summary

Convenience is the meta-health effect most often investigated in valuation studies. In this chapter, the role and value of convenience were investigated in the context of therapy for RA using a DCE. This included an assessment of whether preferences varied with the amount and type of information presented.

Attributes and levels for the DCE were derived from the literature: mode and frequency of delivery; efficacy; treatment discontinuation; and cost (own and government). A WTP efficient design was generated in Ngene using priors from a pilot study. The DCE was administered to a community based on-line panel, with random allocation to three survey versions that differed in the amount and type of information presented. Each respondent answered 12 choice sets (from a possible 48, 4 blocks), unlabelled, using forced choice. Results were analysed using mixed logit and GMNL. Framing effects were tested by comparing coefficient estimates and mWTP values across the survey versions.

The survey was completed by 442 individuals. Treatment mode was important, but respondents placed more emphasis on the role of efficacy and safety in making their choices. In general, choice was not influenced by government cost, but did vary by own cost. mWTP for convenience relative to health effects was influenced by whether the data were analysed as continuous or categorical variables; the latter reducing the ranking of meta-health effects in value determination. Framing effects influenced the results; exclusion of efficacy differences increased the value of convenience relative to health effects, and providing respondents with more information increased the importance of convenience gains and government costs in choices.

The results show that meta-health effects are important in choosing RA therapies. They contribute to the existing literature by demonstrating that the values derived for those effects

are subject to the influence of framing. Presenting as much information as possible for a given attribute reduces the potential bias in the values derived for meta-health effects compared with health effects.

5.1 Introduction

The results of the systematic review of the literature in Chapter 2 show that convenience is the most commonly investigated meta-health effect, and is associated with the highest relative values in valuation studies (see Chapter 2). Higgins et al. (2014)[93] investigated the importance of convenience as a meta-health effect, noting the potential for multiple ways for convenience to influence the QoL values reported in studies, differences in the manner in which those values are obtained, and a lack of consistency in reporting how convenience effects are analysed.

The research presented in this chapter investigates the importance of convenience as a meta-health effect relative to other sources of value, within the context of a chronic health condition. The condition chosen is rheumatoid arthritis (RA), which is reported to occur in 2% of Australians, and is associated with ongoing health care use for its management.[215] Recently in Australia, a new oral formulation of the biological disease modifying anti-rheumatic drug (bDMARD) tofacitinib, has become available for the treatment of patients with RA.[253] This adds an oral administration route to the already available intravenous (IV, e.g. infliximab, abatacept, rituximab), and subcutaneous injection (e.g. etanercept, adalimumab, abatacept) administration modalities. The mode of treatment administration: oral, IV infusion or subcutaneous injection, is potentially a source of (in)convenience for patients and is amenable to testing as a meta-health effect. Similarly, the frequency of administration of these therapies can also impact on their acceptability to patients. Administration frequency for these treatments for RA ranges from twice daily (for the oral therapy) to six monthly (rituximab as an IV therapy).[253]

Thus both mode and frequency of administration can be investigated in terms of the impact of convenience on patient well-being. The primary goal in this research is to assess the value of these factors relative to other factors (the health effects of safety and efficacy) associated with treatment for patients with RA. One approach that is being used increasingly to assess the value of such factors in health care is stated preference discrete choice experiments (DCEs).[45,48] Both the systematic review reported in Chapter 2 and the review by Higgins et al. (2014)[93] show that DCEs are the predominant form of measurement technique used to value QoL effects associated with convenience, and meta-health effects more generally.[90,93,101,112,115-120,130,131,134]

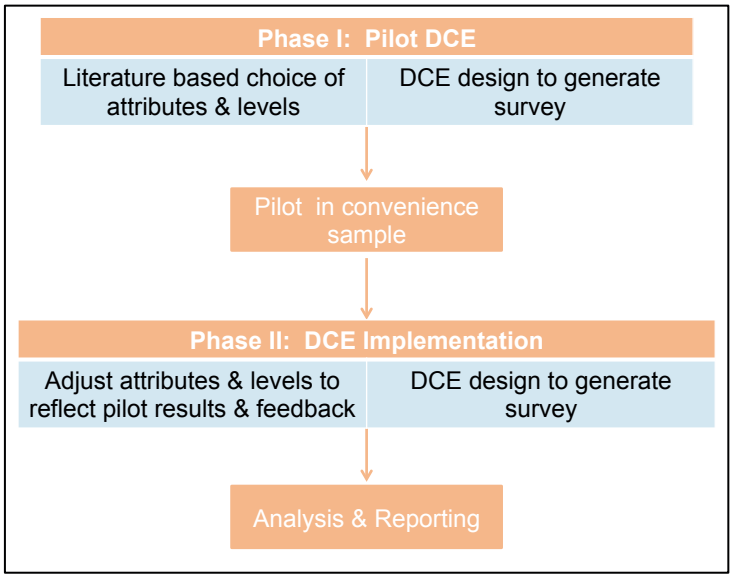
As described in Chapter 4, within a DCE the interventions of interest (e.g. different types of treatment for RA) are described by their characteristics (called attributes) of interest and importance, such as the mode and frequency of administration, or efficacy. Surveys are then constructed in which respondents are asked to make choices between repeated comparisons of alternative combinations of these attributes. This allows the value respondents place on those attributes, such as mode of administration or efficacy, to be determined based on the stated choices they make.[45,47,251] The analysis of those repeated choices allows values to be derived for the included attributes. These values can be used later in designing new treatments, or as inputs to decisions on how health care resources might be allocated.[48,130]

The use of DCEs to understand treatment choices for RA has been investigated previously for countries other than Australia.[90,223,254,255] Similarly, DCEs have been well represented in studies investigating the value of convenience within other health care settings.[101,112,115-120,130,131,134,256] The research presented in this chapter differs from and extends previous applications of the method to treatment choices in RA in three ways. First, the research sets out to value convenience explicitly relative to other factors affecting treatment choice, such as health effects. Second, the research includes both the cost to the individual and the cost to government in the DCE attributes. This addresses one of the aims of the research; whether the role of convenience, a meta-

health effect, is influenced by the cost to government. This allows assessment of whether the marginal willingness to pay (mWTP) for convenience effects differs depending on who is paying for the treatment under consideration. Finally, a key subtheme of this research is to test whether the manner in which the question is asked, or framed, impacts on the values individuals place on the meta-health effects and health effects.

The research proceeds in two phases. In phase I, a pilot DCE questionnaire was developed and tested among a convenience sample. The results from the pilot, including qualitative feedback on the survey and subsequent seminar presentations of the results, were utilised to refine the DCE survey. In phase II, the refined DCE survey was implemented among a sample of the general Australian community drawn from an online panel. The outline for the research is depicted in Figure 17.

Figure 17: Overview of RA Therapy study



Abbreviations: DCE, discrete choice experiment.

The results of the repeated choices by individuals were analysed taking account of individual heterogeneity in responses, in terms of both the analysis of choice itself through the use of mixed logit regression and GMNL modelling[56,236,239,257], and by exploring the potential for demographic factors to influence the probability of

individuals choosing a given choice scenario. The value of the determinants of choice was expressed as the mWTP associated with the trade-offs between cost and other attributes, such as mode of treatment administration, derived from the analysis of choices. All analyses investigate the influence of framing on the values derived for health effects and meta-health effects.

5.2 Methods

This section outlines: the development of the DCE survey; the underlying modelling approach; data collection; and data analysis. While the exploration of framing was a subtheme, it is described first as it provides definitions that appear in subsequent sections.

5.2.1 Development of the DCE

5.2.1.1 Exploring framing

As described in Chapter 1, framing effects reflect a combination of how individuals conceptualise a problem (such as a choice), how that problem is presented to them and the information they are provided.[83,84] The influence of framing effects in the choice of RA therapy was investigated in this research through the use of three framing sets (versions) of the survey that differed in the amount and type of information presented to respondents:

1. 'Base': the base design in which there was a common approach across attributes in terms of the amount of information presented in describing the attributes and their levels.
2. 'Attribute': the amount of information provided describing the mode of treatment administration was increased relative to all other attributes to give more detail about what is required to administer treatment.
3. 'No-Efficacy Difference': scenarios were constructed such that they did not differ in terms of efficacy; there was complete overlap in the *Efficacy*

attribute (the levels were the same between treatment alternatives within a choice set) and described as having the same effect between treatments.

The choice decisions resulting from these framing sets were compared in an approach similar to that utilised by Kragt et al. (2012)[258] and Rolfe et al (2002)[259]; by comparing the coefficient estimates, overall model performance and resulting mWTP values from differently framed versions of the same DCE. In addition, the use of the GMNL model to investigate choices in this study allows a comparison of scale effects (the extent to which there are unobserved shifts in error variances across individuals) in the differently framed versions; the presence of framing effects might thus result in differences in the presence of scale across the versions.

The choice of frames to test in this study is justified on two grounds. First, as revealed in Chapter 2, drug manufacturers are seeking higher prices on the basis of gains in convenience that arise from advancements in the mode and frequency of administration. Thus it is pertinent to examine the influence on the resulting values for convenience of how such advancements are described. This is tested via the 'Attribute' frame. Second, there have been studies investigating the value of convenience that focus on that value without varying the potential efficacy effects.[63,64,113] The 'No-Efficacy Difference' frame is designed to test how such an assumption might influence the values, such as mWTP, derived for convenience relative to health effects.

It was hypothesised that providing more information about each treatment modality in the 'Attribute' frame would result in increased differences between the modalities (e.g. tablet versus injection) via an increased mWTP for convenience, but not necessarily in the overall value for a given modality. Within the 'No-Efficacy Difference' frame, it was hypothesised that not including efficacy differences in the DCE would result in higher values for convenience relative to the health effects. Both of these hypotheses were examined in the analysis of the survey data.

5.2.1.2 Attributes and health state description

5.2.1.2.1 Phase I – the pilot study

The attributes included in the pilot DCE were informed by those included in the DCEs for bDMARDs conducted by Augustovski et al. (2013) and Poulos et al. (2014).[90,223] They focused on attributes addressing the mode and frequency of treatment administration, treatment efficacy, adverse effects (AEs), and out-of-pocket (OOP) costs. The attributes and levels are listed in Table 19, with the levels listed as applying to an alternative frame being those that apply to the 'Attribute' frame (for Mode) and the 'No-Efficacy Difference' frame (for *Efficacy*) only.[xxiii]

Table 19: Attributes and levels – pilot study

Domain	Attribute Description	Level Description	Alternative Frame
Mode	The treatment is: (3)	A 15 minute intravenous infusion. An injection. A tablet.	An intravenous infusion using a needle under your skin and infusion line. Your doctor gives you your treatment. This takes around 45 minutes each time, including 15 minutes for the treatment. An injection under the skin on your thigh. Sometimes you need help to do this so you go to see your doctor. A tablet. Sometimes the arthritis in your hands makes it difficult to handle the tablets.
Frequency	You take the treatment: (3)	Twice a day (*tablets only*) Weekly Fortnightly Monthly (*not tablets*)	n.a.
Efficacy	The chance that the treatment will control your condition, allowing you to do your usual activities is: (4)	Twenty in one hundred Forty in one hundred Sixty in one hundred Seventy in one hundred	The same for the two treatments
Safety	The chance that you need to stop taking your medicine because you experience side effects is: (3)	One in one hundred Five in one hundred Ten in one hundred	n.a.

[xxiii] Respondents allocated to the 'Attribute' frame saw the longer description for mode and not the shorter descriptions which applied only to the 'Base' and 'No-Efficacy Difference' frames.

Domain	Attribute Description	Level Description	Alternative Frame
Cost	This treatment costs $2,000 per month. You pay: (4)	Nothing, the government pays all the costs. $1,000 and the government pays $1,000. $1,500 and the government pays $500. All of the costs.	n.a.

Note: Numbers in parentheses indicate the number of levels associated with each attribute.
Abbreviations: n.a., not applicable.

The wording of the *Efficacy* attribute was developed to reflect the change associated with achieving a treatment response according to the American College of Rheumatology (ACR) criteria for a 50% improvement in RA symptoms (ACR50). A review of bMDARD trials for ACR50 response rates (monotherapy and combination regimens) reports values within this range, and that they are essentially the same across the bDMARDs.[260,261]

The impact of AEs on treatment discontinuation has been used to capture the effects of tolerability. This is because the bDMARDs are not considered to differ in the extent of their tolerability, and there are many possible AEs associated with the various bDMARDs, including: serious lung infections, serious reactions at the site of the needle which is used to administer treatment, and some cancers such as lymphoma and melanoma.[262] Choosing which specific AEs to include as attributes seemed arbitrary and not germane to the question at hand (focusing on understanding how value is determined for convenience).

Levels for the cost of treatment are based on the PBS dispensed price for the bDMARDs for RA. A one month supply of etanercept or abatacept had a dispensed price for maximum quantity of $1,774 as at October 2014. This provided the upper bound of the cost attribute.

Specifying how costs were to be shared between individuals and the government is novel in the context of DCEs in this area. The levels ranged between the situation under which all costs were born by government (emulating patients who access subsidised medicines via the PBS safety-net), to the individual being responsible for all drug costs (as would occur if a drug was to be purchased as a non-PBS subsidised private script). OOP costs were used as the payment vehicle in this case, because this research is interested in the factors influencing individual decision-making. Moreover, attempting to express the government costs associated with the treatment alternatives as changes in personal income tax contributions required to fund drug subsidies would have resulted in low values that would be unlikely to influence individual choices.

The levels associated with the mode and frequency of administration of the treatments themselves are self-explanatory, and are designed to cover the spectrum of possible treatments available to patients with RA.[260-262] For example, abatacept is administered by IV infusion once a month, while tofacitinib is administered orally twice daily.[253] Both of these treatments would be captured by the mode and frequency attributes described. Potential differences in the mode of administration are highlighted in the level descriptions included in the 'Attribute' frame (see Table 19). The longer attribute descriptions for mode of delivery, along with the introductory health vignette (see Text Box 1) were developed by reviewing patient experiences available from RA websites (http://www.nras.org.uk/stories and http://empowered.org.au/rheumatoid-arthritis/), along with treatment descriptions in the literature[253,260,261] and the product information for a number of DMARDs and bDMARDs used in the treatment of patients with RA.[263-267]

All respondents saw the same introductory health vignette presented in Text Box 1. This was designed to provide respondents with a minimum level of information regarding RA, its prognosis if untreated, and an outline of the treatment choices. As much as possible, the language used in the vignette was designed to be neutral in describing the condition and its treatment.

Text Box 1: Rheumatoid arthritis treatment description – pilot study

Health State Scenario
Your doctor has told you that you have rheumatoid arthritis.

This is a condition that affects the joints of your hands and your knees. Your doctor explains that without treatment you will experience increasing pain in your knees and hands. The joints in your hands will swell, distorting how your hands look and how well you can use them to grip or hold objects. Your knees will become stiff and sore, making it difficult to use stairs and to get up after you've been sitting. This makes it difficult for you to do your usual daily activities.

The doctor asks you to decide between two possible treatments for your condition. The treatments differ in terms of how you take them, how often you take them, the effect they have on your condition, whether you experience side effects that mean you stop taking them, and how much they cost you per month. Regardless of the treatment you choose you will need to have blood tests every three months to check your condition, and whether or not your treatment is being effective.

You will now see 12 questions describing two different treatments. Each time, please indicate which treatment option you would prefer.

5.2.1.2.2 Phase II – the full study

Following the pilot survey and subsequent feedback during a CHERE seminar, modifications were made to the way in which costs and AEs were included in the DCE. It was noted that because the cost attribute presented total cost as the complement of individual and government cost, it would not allow the individuals' own mWTP to be disentangled from what they thought the government should pay. Thus further consideration was given to the construction of this attribute in terms of how to present realistic levels with respect to individuals' OOP costs, and elicit the impact on individuals' choices associated with the impact of government subsidies of treatment costs.

To allow for the explicit valuation of the effects of own and government costs on choice, the cost attribute was restated as *Own OOP* only, with the addition of a separate government cost attribute. Both attributes noted that costs were paid on a monthly basis. The levels for *Own OOP* were 0, $40, $250 and $500; while for *Govt. Costs* they were 0, $500, $1,500 and $3,000. The levels for *Own OOP* were based on the PBS patient charges current at the time of creating the design (2015) of $36.90 for

general patients (corresponding to the $40 level of *Own OOP*), with the higher levels designed to test mWTP limits. The cost levels for *Govt. Costs* were based on the PBS prices of the bDMARDs, etanercept and abatacept as described previously. The revised cost attributes and levels are presented in Table 20

Table 20: Attributes and levels - final survey

Domain	Attribute Description	Level Description	Alternative Frame
Mode	The treatment is: (3)	A 15 minute intravenous infusion. An injection. A tablet.	An intravenous infusion using a needle under your skin and infusion line. Your doctor gives you your treatment. This takes around 45 minutes each time, including 15 minutes for the treatment. An injection under the skin on your thigh. Sometimes you need help to do this so you go to see your doctor. A tablet. Sometimes the arthritis in your hands makes it difficult to handle the tablets.
Frequency	You take the treatment: (3)	Twice a day (*tablets only*) Weekly Fortnightly Monthly (*not tablets*)	n.a.
Efficacy	The chance that the treatment will control your condition, allowing you to do your usual activities is: (4)	Twenty in one hundred Forty in one hundred Sixty in one hundred Seventy in one hundred	The same for the two treatments
Safety	The chance that you need to stop taking your medicine because you experience side effects is: (3)	One in one hundred Five in one hundred Ten in one hundred	n.a.
Cost	The cost to you per month of treatment is: (4)	Nothing $40 per month $250 per month $500 per month	n.a.
	The cost to the government per person, per month of treatment is: (4)	Nothing $500 per month $1,500 per month $3,000 per month	n.a.

Note: Numbers in parentheses indicate the number of levels associated with each attribute.
Abbreviations: n.a., not applicable.

The inclusion of *Own OOP* and *Govt. Costs* allows for the estimation of mWTP from two perspectives; the consumer and the government. However, for any given cost attribute, the resulting mWTP values derived will be influenced by the range of the levels included within the choice experiment.[268] Thus, the comparability of the mWTP arising from the *Own OOP* and *Govt. Costs* will be limited by the difference in the ranges included for the levels of the two attributes.

To address suggestions that more information was required for respondents on the AEs associated with treatment, the AEs experienced most commonly among patients being treated for RA (serious lung infections, serious reactions at the site of the needle which is used to administer treatment and some cancers such as lymphoma and melanoma)[262] were included in the introductory health scenario (see Text Box 2). This information was common across all three frames.

Text Box 2: Rheumatoid arthritis treatment description – final study

Treating Rheumatoid Arthritis
Your doctor has told you that you have rheumatoid arthritis.

This is a condition that affects the joints of your hands and your knees. Your doctor explains that without treatment you will experience increasing pain in your knees and hands. The joints in your hands will swell, distorting how your hands look and how well you can use them to grip or hold objects. Your knees will become stiff and sore, making it difficult to use stairs and to get up after you've been sitting. This makes it difficult for you to do your usual daily activities.

The doctor asks you to decide between two possible treatments for your condition. The treatments might differ in terms of how you take them, how often you take them, the effect they have on your condition, how much they cost per month and the chance you will experience side effects that mean you have to stop taking them. These side effects might include serious lung infections, serious reactions at the site of the needle which is used to administer treatment and some cancers (like lymphoma and melanoma). Regardless of the treatment you choose you will need to have blood tests every three months to check your condition, and whether or not the treatment is being effective.

You will now see 12 questions describing two different treatments. Each time, please indicate which treatment option you would prefer.

5.2.1.3 DCE experimental design and survey development

Full details of the development of the experimental design for the pilot and final study are provided in Appendix 7. In summary, both used a WTP efficient design constructed in Ngene.[229] This necessitates the use of coefficient values, or priors, in order to generate the design. Those for the pilot study were obtained from the literature[90,223], while those for the final study were obtained from the pilot study.

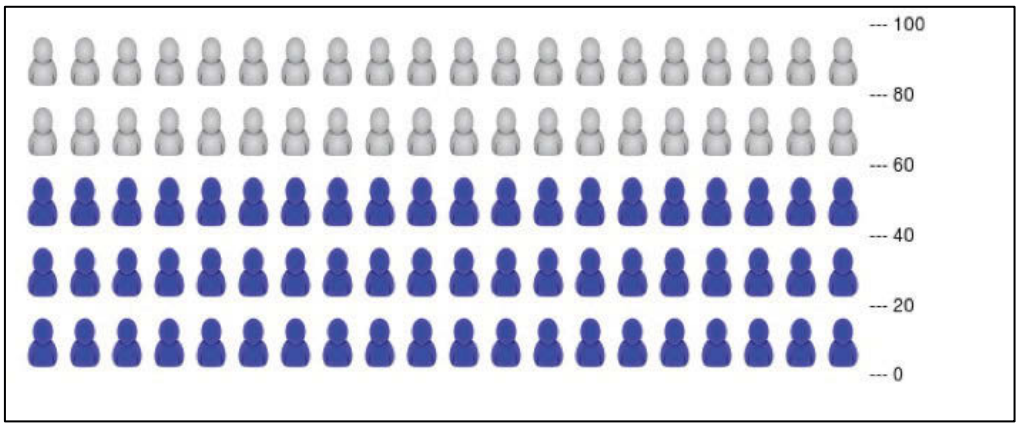
The design included two restrictions on the levels of mode and frequency: where mode was a 15 minute IV infusion, frequency was restricted to exclude the possibility of including twice a day administration; and where mode was tablet, frequency was restricted to exclude the possibility of allowing monthly administration. The inclusion of restrictions results in the loss of level balance within the design with respect to the attributes of mode and frequency.

For the pilot study, the combination of attributes and levels produced a $4^2 \times 3^3$ design[xxiv], with 432 possible combinations. The inclusion of an additional attribute for government cost altered the experimental design to a $4^3 \times 3^3$ with 1,728 possible combinations. The minimum possible number of profiles for estimation of a level balanced design (as it is divisible by 4 and 3 without remainder) is 12.[269] For both the pilot and the final studies, a design using a total of 48 rows was used in Ngene, in four blocks (incorporated into the design) of 12 choice sets, with an option pre-specified in Ngene to minimise the chance of dominant alternatives across choice sets. This design was replicated for each of the three frames ('Base', 'Attribute' and 'No-Efficacy Difference'), giving a total of 12 possible blocks. Respondents to the survey were randomised to one of each of the 12 blocks. The order of presentation of choice sets within blocks was also randomised. All choice sets were forced choices; and an unlabelled design was used.

[xxiv] The inclusion of restrictions between mode and frequency collapses the apparent four levels in frequency to only three.

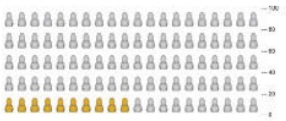
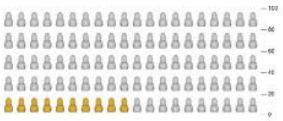
Both the pilot and final surveys were activated online using the Qualtrics platform.[199] To enhance the communication of the *Efficacy* and *Safety* attributes, presented as proportions, icon-arrays were used. This is based on research that has found that the presentation of risk or proportion evidence as numbers can be misinterpreted by some respondents, and that the inclusion of icon-arrays, graphical depictions of numerical events, enhances comprehension.[270] All icon-arrays were generated online and downloaded from the online generator, www.iconarray.com. Individual icon-arrays were generated for each proportion to be used within the survey. An example of an icon-array is presented in Figure 18. Thus, to facilitate the presentation of the choice scenarios in the format desired, with the inclusion of icon-arrays, and importantly, preserving the experimental design, each scenario, for each version was loaded into Qualtrics as a separate picture. This is because Qualtrics does not allow an existing DCE design or icon-arrays to be uploaded automatically. The process used to convert the Ngene designs into the required picture formats for loading into Qualtrics is contained in Appendix 6. An example of a formatted choice set is presented at Figure 19.

Figure 18: Example icon-array, rheumatoid arthritis study



Note: Icon-array depicting a 60% chance (60 out of 100) that the treatment will control your condition, allowing you to do your usual activities.

Figure 19: Example choice set (No-efficacy difference frame), rheumatoid arthritis study

	Option A	Option B
The treatment is:	An injection	An injection
You take the treatment:	Fortnightly	Weekly
The chance the treatment will control your condition, allowing you to do your usual activities is:	The same for the two treatments: 40 in 100 	
The chance you stop taking your treatment because you experience side effects is:	10 in 100 	10 in 100 
The cost to you per month of treatment is:	$250 per month	$250 per month
The cost to the government per person, per month of treatment is:	$3,000 per month	$1,500 per month

Which one of these options would you choose?

☐ Option A

☐ Option B

5.2.2 Modelling approach

The approach to the analysis of the choice data was provided in Chapter 4. Main effects only were estimated, based on the following: the meta-health effects for mode (*IV:Tablet*, *Injection:Tablet)*, and frequency of delivery (*BD:Monthly, Weekly:Monthly, Fortnightly:Monthly);* the health effects (*Efficacy* and *Safety*); and costs of treatment *; OwnOOP; and Govt.Cost.* The estimation of main effects assumes separability of the influence of each of the attributes. It is possible that the effects from the attributes for mode and frequency of delivery, and costs of treatment (own and government costs) respectively would overlap. This will be examined in the first instance by comparing

the overall coefficient values and significance, with the potential for further testing using two-way interactions.[xxv]

As previously described the mode and frequency variables are both treated as categorical. Therefore, *IV: Tablet* and *Injection: Tablet* represent effects coded variables for the effects on choice of preferred medication of the mode of administration of IV or injection relative to tablet. Similarly, *BD:Monthly*, *Weekly: Monthly* and *Fortnightly: Monthly* represent the effects coded variables for the impact on choice of the preferred medication of the frequency of administering treatment twice a day (BD), weekly or fortnightly compared with once a month. All other variables are treated as continuous in the base analysis. They represent the likelihood that treatment will resolve condition symptoms (*Efficacy*), lead to AEs that require treatment cessation (*Safety*), own OOP costs per individual per month of treatment (*OwnOOP*), or the cost to government per individual per month of treatment (*Govt.Cost*).

The analysis begins with a conditional logit regression, then explores preference heterogeneity using a mixed logit regression (equation 15), and GMNL (equation 16). The GMNL is initially implemented without specifying a value for γ. However, given the number of observations available for each of the survey frames (approximately n=150), rather than estimate an unrestricted model, a restriction of independence between scale and preference heterogeneity (see the discussion in Chapter 4) will be imposed ($\gamma = 1; \text{GMNL} - \text{I}$) as has been applied elsewhere.[243,271]

5.2.2.1 Capturing the effects of demographics

Within this chapter, the analyses focus on estimates of the main effects determining choice. The influence on choice of individual specific factors such as age, gender,

[xxv] The use of two way interactions for mode and frequency would require the addition of 10 possible interactions between the levels of mode and frequency, allowing for the two restrictions in the design. Testing during the design of the pilot survey indicated that such an interaction would greatly increase the sample size required for estimation without loss of efficiency.

income and education is explored in terms of their ability to explain variations in the predicted probability of each choice arising from the GMNL estimations. This is shown in expression (19), where $\widehat{U}_{ijt}$ is the GMNL predicted probability of making a choice for each individual over each option *j* at each choice time *t*, adding in the socio-demographic variables (*Age*, *Income*, *Education*, *Gender*, *Arthritis*, *Ongoing Medication*), while also accounting for the underlying attributes.

$$\widehat{U}_{ijt} = \alpha + \beta_1 Age_i + \beta_2 Income_i + \beta_3 Education_i + \beta_4 Gender_i + \beta_5 Arthritis_i + \beta_6 Medication_i + \beta_7 IV{:}\,Tablet_{jt} + \beta_8 Injection{:}\,Tablet_{jt} + \beta_9 BD{:}\,Monthly_{jt} + \beta_{10} Weekly{:}\,Monthly_{jt} + \beta_{11} Fortnightly{:}\,Monthly_{jt} + \beta_{12} Efficacy_{jt} + \beta_{13} Safety_{jt} + \beta_{14} OwnOOP_{jt} + \beta_{15} Govt.Cost_{jt} + \varepsilon_{ijt} \quad (19)$$

Two specifications were estimated, one in which α is permitted to vary, thereby capturing the average predicted probability, and the other in which α =0, in effect providing an additional estimate of the contribution of each attribute, as well as the demographic variables, to the predicated probability of each choice. Each demographic variable enters the model specification as a categorical variable, using effects coding. The relevant category structure is described in the results section of this chapter.

The underlying choice models in this chapter utilise logistic regression on the basis that at each choice moment, an individual can make a decision to choose an option or not; depicted as a binary outcome (0 or 1). However, the predicted probabilities produced by the GMNL are the probability on the continuum between 0 and 1 that an individual will choose a given option (Option A or Option B) at each choice task. As such, the predicted probability values have been modelled using a linear function over a continuous scale, rather than as a logistic function over a binary or discrete scale.

5.2.3 The survey data and analysis

5.2.3.1 Data collection

Respondents for the pilot survey were recruited by direct email invitation and by loading the survey invitation link onto the investigator's FaceBook page. This meant that the respondents were more highly educated, had a higher household income and a higher proportion of females compared with the Australian population. Nonetheless, the results were primarily intended to refine the prior parameter values used in the DCE design (see Appendix 7), and to test the language and face validity of the questionnaire. A minimum of 30 respondents was targeted for the pilot study.

Respondents for the main survey were recruited using the Pureprofile online panel.[198] This panel has over 60,000 active members Australia-wide monthly. Availability of the survey was posted to the portals of all Pureprofile members over the age of 16 years, without further restriction. The survey was live during April 2015. Survey responses were monitored to achieve a balanced distribution of respondents in terms of age and gender across each of the survey blocks, with a target of 450 completed surveys (see Chapter 4).

5.2.3.2 Analysis plan

The pilot data were analysed using a standard conditional logit regression, with the principal analysis focusing on the results from the combined or 'pooled' data across all three frames. Tests were conducted for differences between the coefficients produced by the three frames (see the section below on poolability). Exploratory analyses were conducted of the mWTP to assess whether the experimental design was behaving in the expected manner, as well as exploring heterogeneity in choice behaviour using latent class modelling.

Initial analyses of the main survey were of the pooled data using mixed logit regression. Differences between frames were tested to identify whether subsequent

analyses should be conducted using pooled or unpooled data. Subsequent analyses utilised GMNL to take account for scale variance. Results from the GMNL were used to estimate mWTP. The resulting mWTP were ranked in terms of their size relative to the largest value in each framing set. The ability to test models in various demographic sub-groups, such as those with or without arthritis, was informed by examining design coverage by key demographics. Finally, the relationship between the predicted probabilities and demographics was tested.

5.2.3.2.1 Maximisation procedure

All analyses were conducted using STATA version 12, and copies of the STATA code are available upon request. Specification of the conditional logit was according to the available STATA command *clogit*.[272] Use of the mixed logit was by invoking the STATA user written command *mixlogit*.[257] Use of the GMNL was by invoking the STATA user written command *gmnl*.[243] The relationship between the predicted choice probabilities and demographic variables depicted in expression 7 was tested using OLS with the STATA command *regress.*

The online survey company Pureprofile utilises a general sampling technique where all panel members are invited to participate and essentially self-select onto the survey. Thus, it is not possible to adjust the model estimation procedure for the participation rate in the survey since the denominator, those who actually view the survey invitation notice, is not known. Therefore, all regression analyses within STATA have been conducted using robust standard errors, or their equivalent (by invoking the cluster option) to account for the survey nature of the data.[213]

Initial runs of the regression analyses for the mixed logit regression encountered multiple iterations that were not concave and did not achieve convergence, potentially due to the sample size. Accordingly, two strategies were implemented to achieve better convergence:

1. An alternative convergence algorithm within STATA was used by invoking the

"difficult" option (available with all maximum likelihood estimation commands like the mixed logit and GMNL commands).[213]

2. The starting values for higher replications (500 and 1000) of each mixed logit regression and GMNL specification were based on the resulting matrix of values for the preceding iteration of that model with fewer replications (the STATA default is 50 replications).[213]

Maximisation was according to the Newton Raphson algorithm, the default within STATA for these maximization procedures.[213] It is typically standard to report model fit using the pseudo-R^2. However, as noted by Greene (2008), while it is possible to estimate such a measure using the log-likelihood, it requires that the model contain a constant term (for comparison with the log-likelihood from a model of the dependent variable on the constant only).[56] Since the models used in this chapter do not contain a constant term, the estimation of a pseudo-R^2 for the mixed logit regression and GMNL estimation procedures is not appropriate (a pseudo-R^2 is produced automatically for the conditional logit estimation procedure within STATA).

Use of other standard model fit criteria such as the overall log-likelihood, model significance, and information criteria (AIC/BIC) is reported for each estimation procedure. Note that Gu et al. (2013)[243] caution that the use of the AIC/BIC in conjunction with the GMNL procedure in STATA may be misleading due to the number of observations stated in that procedure (GMNL), and used in the information criteria estimates, the former also taking into account the total number of choices per alternative, rather than just the total number of choices as would typically be the case when estimating those criteria (the direction of the bias is unclear). However, as the comparisons are only being made between the various frames all using the same GMNL estimation procedure, it is likely that any bias will operate in the same direction so as to allow the standard interpretation of AIC and BIC to apply (the lower the AIC/BIC value the better the model fit).

5.2.3.3 Test of poolability

The impact of framing effects on the coefficient estimates can be tested globally by assessing whether the data from the three versions of the survey (Base, No-Efficacy Difference and Attribute) can be pooled.[84,258,259] This has been tested in two ways: by comparing the overall data fit using a likelihood ratio (LR) test; and a Chow-test to examine the influence of belonging to a framing set on resulting attribute coefficients. Both these tests use estimates from the mixed logit regression, and therefore have taken into account underlying preference heterogeneity. Neither test adds an additional adjustment for scale, as is required under a Swait and Louviere test applied to multinomial logistic regression estimates that do not account for heterogeneity.[273] The LR test compares a sum of log-likelihood values from each framing set with a pooled version log-likelihood to examine the benefits of estimating the data combined across, or separately for, each framing set.[259,271] Pairwise LR tests, comparing pairs of framing sets, were also conducted to examine whether the observed differences are due to one particular framing set. In the Chow test, sample specific variables (effects coded) were introduced into the pooled regression, both as main effects and interacted with each of the attributes and levels. In order for the mixed logit regression to be estimated in STATA, it cannot contain more than 20 random covariates. For this reason, each of the sample specific variables and interactions were included as fixed rather than random effects, hence there is no estimate of the heterogeneity associated with the coefficient estimates for those interacted attributes (there are no reported measures of the standard deviation for those estimates).

5.2.3.4 Estimating mWTP

For all attributes, estimates of the mWTP were formed using *Own OOP* as the numeraire (a mWTP was not estimated for *Govt.Costs* since the interpretation of the ratio of *Govt.Costs* to *Own OOP* is not clear). This utilised the estimates produced by the GMNL specifications. mWTP were tabulated for those coefficients that were statistically significant, but all mWTP results are presented graphically. Estimates of

mWTP were formed by invoking the willingness to pay user written command within STATA, *wtp*.[272]

5.2.3.5 The influence of data coding on estimated value

One of the potential sources of influence on the size of the coefficient estimates and mWTP for the meta-health effects relative to the health effects is the way they have been coded for analysis. As is fairly standard to the analysis of such preference data in this area[92,159,223], it has been assumed that the meta-health effects enter the analyses as categorical variables, while the health effects and costs are treated as continuous variables.

However, this difference in coding may not take into account how much information each attribute conveys about the overall product when being assessed by respondents. For example, the coefficient on mode reflects preferences over an entire aspect of a product e.g. IV compared with tablet. When comparing scenarios that differ based on mode (e.g. IV versus tablet) respondents see and evaluate the entirety of those attributes. In contrast, the coefficient estimate for *Efficacy* is for a one-point change in the chance of improvement, and has been derived assuming that those data are continuous (or linear). The resulting choice estimates and mWTP are interpretable as absolutes for a one unit change (in this case 1% point). In reality, survey respondents consider differences between categories of *Efficacy* outcomes e.g. 20% compared with 60%, and not one-point changes. Assuming that health effects can be modelled as continuous implies that respondents attend to, or consider, all the levels within those attributes – but this might not be the case.[274] In addition, real-life product choices might involve larger differences in efficacy such that the relative value of health effects to meta-health effects changes beyond that evaluated when health effects are modelled as continuous variables. Comparing the contribution to value for meta-health effects with that of health effects is therefore confounded by the differences in the units of measure for those attributes introduced by the differences in coding.

To test the effect of relaxing the assumption regarding the linearity of health effects on the choices individuals make, and importantly on the relative weighting placed on meta-health effects relative to health effects, the GMNL specifications were re-estimated using categorical coding (effects coded) for *Efficacy* and *Safety*. In each case the base attribute value was the most desirable level, such that deviations from that level would be seen as undesirable. *Own OOP* and *Govt.Cost* were retained as continuous variables to allow mWTP to be estimated readily.[xxvi] Arguably, the impact on the mWTP of differences in the units in which attributes are measured could be addressed by multiplying the mWTP for a one-point change by the relevant difference in that measure between two relevant product profiles. This suggests that to estimate the mWTP to move from an efficacy of 20% to 70%, the mWTP associated with a one-point change should be multiplied by 50 (relying on the linearity across all costs). This method is tested, both by examining the linearity of the health effect coefficients when coded as categorical variables, but also by comparing the mWTP estimated from the categorically coded health effects with those based on multiplications of the mWTP for a one-point change

5.3 Results

5.3.1 The pilot

The pilot survey was completed in October 2014. Results are provided in Appendix 8. A total of 33 respondents completed the survey, providing responses that could be analysed to investigate choices. Feedback on the survey language and design indicated that it was well understood by respondents; the language and the tasks were clear (Table A 13). Specific comments on attribute language and levels were considered for possible amendments to the main survey.

[xxvi] Rather than as a function of multiple possible cost functions, dependent on the level of the cost attribute applied.

The significant factors influencing choice were the mode of treatment administration (shifting from *Tablet* to *IV*), *Efficacy* and *OOP* (see Table A 18). This was reflected in the resulting mWTP values, with the highest value produced for the convenience attribute of moving from *Tablet* to *IV* (see Table A 22). However, the ranking of those values was highly dependent on coefficient significance, and whether or not *Efficacy* and *Safety* were treated as continuous or categorical data (see Table A 24). The latter resulted in *Efficacy* being ranked as the most important attribute, which is consistent with the feedback on attribute importance provided by respondents (Table A 15).

5.3.2 The final survey

5.3.2.1 An overview of the findings

Analysis of the DCE results shows that convenience, reflected in the mode of administration, was important in affecting choice. However, health effects in terms of efficacy and safety were more important in determining preferences. While preferences were influenced by own costs, they generally did not vary with government cost except when more information was provided. The mWTP when assessed using own costs shows higher values for convenience gains, such as a move from IV administration to tablet, than for one percentage point gains in efficacy. This was influenced by whether the treatment response and safety data were analysed as being continuous or categorical. Incorporating efficacy and safety as categorical effects resulted in those effects having higher mWTP relative to convenience.

Analysis across the three framing versions shows that the type and manner in which information is included influences the role of meta-health effects (convenience) and the values elicited. Excluding efficacy differences increased the importance of convenience relative to health effects, both in terms of the preference orderings and the subsequent mWTP. Similarly, providing respondents with more information about the way in which treatment was administered increased the importance of convenience gains and government cost in forming choices.

5.3.2.2 Demographics

The final survey was completed by 422 individuals. Only one respondent, in the 'Attribute' frame did not complete the questions on demographics. Information about health status, including the presence of chronic health issues, was available for all 422 respondents as this was completed prior to the general demographics questions. Selected demographic characteristics for respondents are provided in Table 21. In general, compared with the Australian population, fewer survey respondents were in excellent health, more had chronic health issues, and the sample was skewed towards the middle age groups.[214] The incidence of arthritis among respondents was higher than that reported for the Australian adult population.[215][xxvii]

Table 21: Final survey - demographics

	Pooled n = 421	**Base** n=141	**Attribute** n=139	**No-Efficacy Difference** n=141	**Australian Population**
Health Status, %					
Excellent	8.29	7.09	8.57	9.22	20.2
Very good	40.05	40.43	39.29	40.43	35.4
Good	33.89	36.17	32.86	32.62	30.0
Fair	13.98	14.18	14.29	13.48	10.4
Poor	3.79	2.13	5	4.26	4.0
Chronic disease, %	59.00	60.28	59.29	57.45	45.00
Arthritis, %	20.38	21.99	21.43	17.73	14.29
Ongoing Medications, %	43.13	44.68	47.86	36.88	n.a.
Female, %	49.17	46.10	51.80	50.35	50.60
Age, %					
16-24	2.38	1.42	3.60	2.13	16.31
25-44	38.48	36.88	38.13	40.43	34.73
45-64	39.67	36.88	41.01	41.13	31.42
65-74	15.68	20.57	12.23	14.18	9.37
75 or more	3.8	4.42	5.04	2.13	7.98
Median household income, $	1,150-1,529	1,150-1,529	1,150-1,529	1,530-1,919	1,234
Education, %					
School only	28.50	30.50	25.18	29.79	n.a.
University	33.73	36.88	33.09	31.21	n.a.
Vocational & Other	37.77	32.62	41.73	39.01	n.a.

Abbreviation: n.a., not applicable.

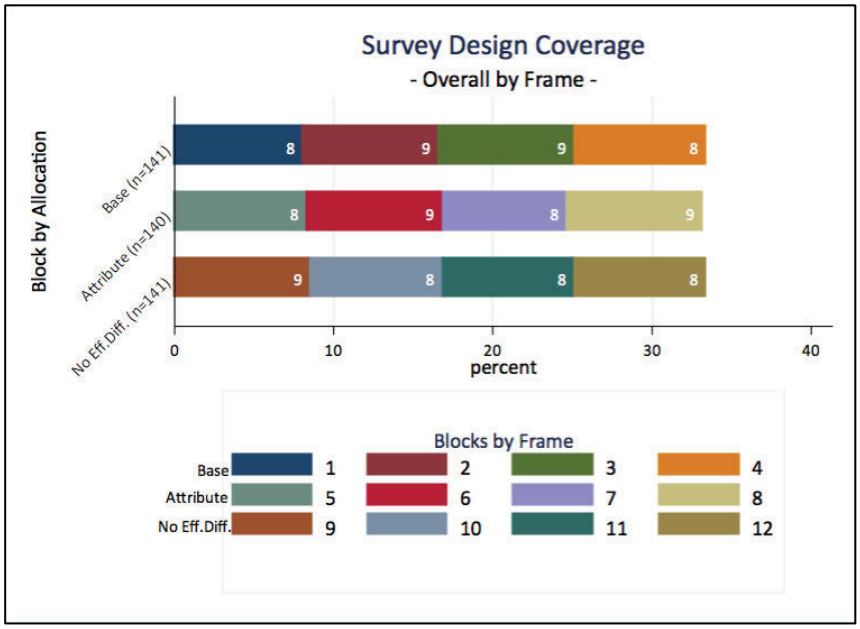
[xxvii] Recruitment to the survey was not monitored for RA status. A higher incidence of RA among survey participants might reflect interest in the survey among Pureprofile members with RA.

While the survey sample was monitored for balance in terms of the allocation of respondents to the respective blocks and for overall gender balance, stratification of respondents by gender or other demographic factors was not part of the randomisation process. Potential differences between the frames in terms of the demographic composition of respondents was tested using one-way tests of analysis of variance (Bartlett's test of equivalence of variance), with the demographic factor as the response variable and the frame as the factor or explanatory variable. The results of these tests showed that there were statistically significant differences across the frames with respect to age $\chi^2(2)= 45.21$ ($p<0.001$); arthritis status $\chi^2(2)= 26.12$ ($p<0.001$); gender $\chi^2(2)= 2.7e+0.4$ ($p<0.001$); and health status $\chi^2(2)= 45.43$ ($p<0.001$). There were no statistically significant differences for income $\chi^2(2)= 0.87$ ($p=0.647$); education $\chi^2(2)= 5.15$ ($p=0.076$); or chronic health status $\chi^2(2)= 0.37$ ($p=0.830$).

5.3.2.3 Survey coverage

Recall that the survey design consisted of 48 choice tasks, allocated into four blocks of 12 choice sets, and replicated over the three different frames ('Base', 'Attribute' and 'No-Efficacy Difference'). Blocks were numbered 1-12, with blocks 1-4 being allocated to the 'Base' frame, 5-8 to the 'Attribute' frame and 9-12 the 'No-Efficacy Difference' frame. Within any frame, there was complete design coverage; all 48 choice sets were presented and the coverage of respondents across the blocks was balanced (see Figure 20). That is, the proportion of total respondents allocated to each block of questions was essentially the same both within and across the three frames.

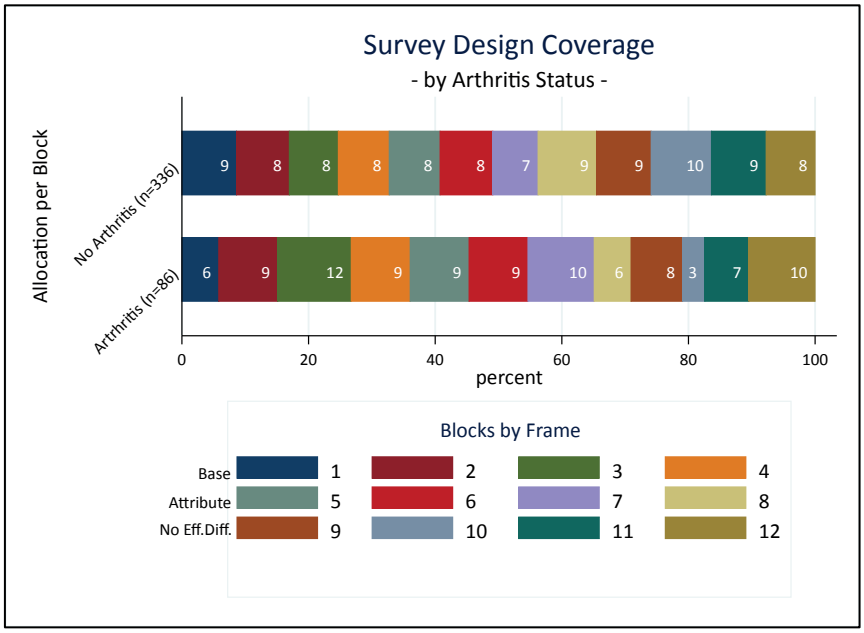
Figure 20: Design coverage by frame



Note: Numbers shown in bars are percentages of the total survey sample.

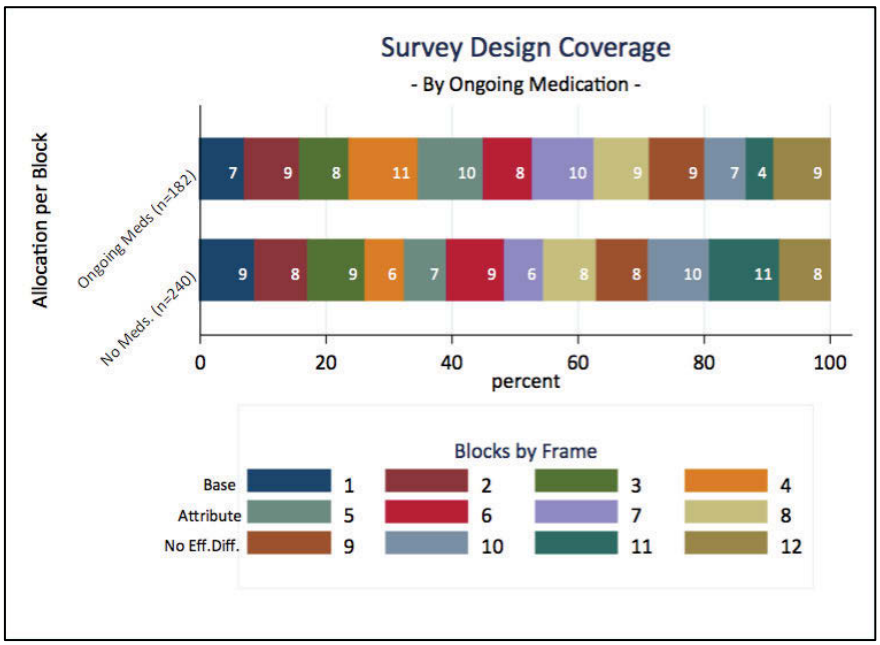
Design coverage was also examined for two potential stratification factors that might influence choice; RA status and the use of ongoing medications for a chronic health condition. From the data in Figure 21 and Figure 22 it can be observed that there were imbalances in design for both these factors. For example, 10% of those without RA were allocated to block 10 ('No-Efficacy Difference'), yet only 3% of those with RA were allocated to this block. Similarly, 11% of those on ongoing medications were in block 4 ('Base'), but only 6% of those who were not on ongoing medications were allocated to this same block. Tests of association between RA status and block allocation, and medication status and block allocation were statistically significant; Bartletts's $\chi^2(11)= 302.91$ ($p<0.001$) for RA status and $\chi^2(11)= 39.65$ ($p<0.001$) for medication status. Given the imbalance in coverage for both of these factors, neither is used to stratify the analysis of choice. However, they are used to describe underlying rankings of attribute importance.

Figure 21: Design coverage by arthritis status



Note: Numbers shown in bars are percentages of the respective sub-samples.

Figure 22: Design coverage by ongoing medication use



Note: Numbers shown in bars are percentages of the respective sub-samples.

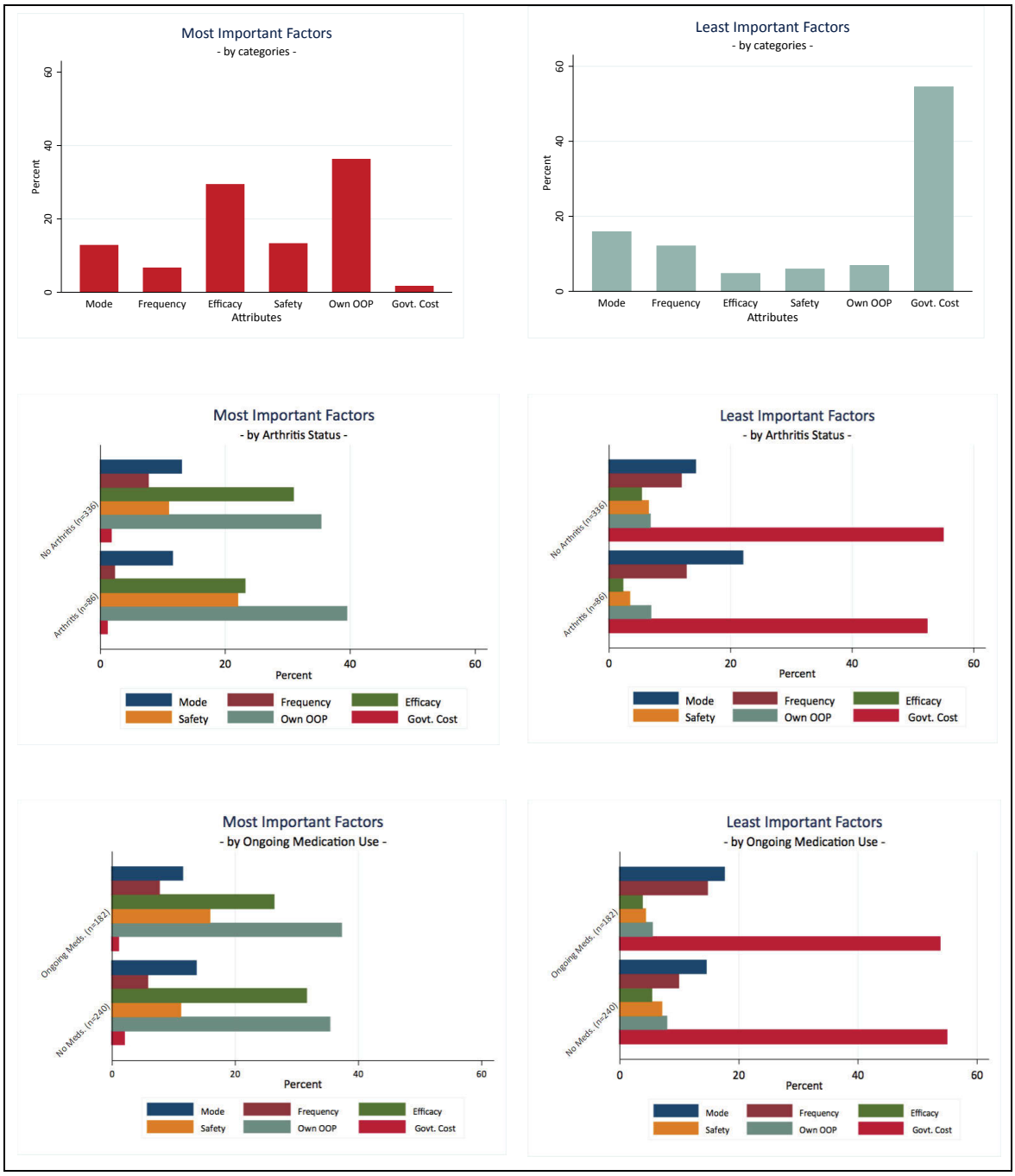
5.3.2.4 Attribute rankings in decisions

After completing the 12 choice sets, respondents were asked to report which of the six attributes they considered to be the most important when making their choices, and

which the least important. As can be observed from the results in Figure 23, the overwhelming majority of respondents (over 50%) considered *Govt.Cost* to be the least important factor in their decision-making, followed by the convenience factors. Conversely, *Own OOP* was most commonly considered the most important factor in decision-making by respondents (35%), followed by *Efficacy*.

Attribute rankings were also investigated based on RA status and use of ongoing medications. Taking into account ongoing medication use, the distribution of attribute rankings differed somewhat from those of the overall population; respondents not taking ongoing medications were more evenly distributed in terms of considering *Efficacy* and *Own OOP* as the most important factor. Similar differences were noted based on whether or not respondents reported having RA. That is, while respondents with arthritis continued to rank *Own OOP* as the most important factor, they ranked *Efficacy* and *Safety* almost equal as the next most important factor. Respondents without RA ranked *Efficacy* as the clear second most important factor. With the exception of *Govt. Cost*, the clearly least important attribute, the convenience attributes of mode and frequency of administration were ranked as being of less importance than the health effects and *Own OOP*.

Figure 23: Attribute rankings



Abbreviation: OOP, out-of-pocket.

The apparent differences in the ordering of factor importance based on RA status and ongoing medication use suggests that these factors might be used to form subgroups for analysis. However as noted previously, the design coverage was not balanced across these factors so they have not been used to stratify the analysis of choice. Both factors are used as covariates to explore the influence of demographic factors on the predicted probabilities of choice.

5.3.2.5 Determinants of choice: mixed logit regression

The results of the mixed logit regression analysis are presented in Table 22. Five specifications are presented: two for the pooled analyses across the three framing sets (with and without a constant term); and three separate specifications for each of the sample frames – 'No-Efficacy Difference', 'Base' and 'Attribute'. Across all five specifications it can be observed that *Efficacy* and *Safety* (health effects), and *Own OOP* are always significant and influence choice in the direction expected. However, the influences of the meta-health effect convenience (as shown by mode and frequency) and *Govt. Cost* vary depending on the framing set. For example, across all three framing sets, respondents exhibited a preference for the tablet over the IV infusion (evident by the negative and significant coefficient on *IV:Tablet*), although preferences for the tablet over the injection are significant only in the 'No-Efficacy Difference' frame. Similarly, respondents are influenced only by the extreme frequency of twice-daily relative to monthly administration of treatment (*BD: Monthly,* the negative sign favouring monthly over twice-daily treatment); all other levels of the frequency attribute were not statistically significant. The results for the standard deviations indicate significant heterogeneity across individuals in the influence of all attributes (although not all levels for the frequency attribute in all frames) in treatment choice.

Table 22: Mixed logit regression

	Pooled Constant	Pooled	Base	Attribute	No-Efficacy Difference
Coefficient Means					
Constant	-0.200				
	(0.059)**				
IV: *Tablet*	-0.796	-0.756	-0.750	-1.016	-0.647
	(0.103)**	(0.103)**	(0.185)**	(0.173)**	(0.190)**
Injection: *Tablet*	-0.038	-0.067	0.051	0.144	-0.324
	(0.060)	(0.06)	(0.123)	(0.105)	(0.101)**
BD: *Monthly*	-0.463	-0.453	-0.255	-0.564	-0.500
	(0.092)**	(0.094)**	(0.152)	(0.169)**	(0.160)**
Wk: *Month*	-0.014	-0.046	0.074	-0.098	-0.150
	(0.050)	(0.049)	(0.083)	(0.092)	(0.088)
Ft: *Month*	0.038	0.052	-0.090	0.114	0.161
	(0.051)	(0.051)	(0.101)	(0.082)	(0.091)
Efficacy	0.082	0.081	0.090	0.067	n.a.
	(0.008)**	(0.009)**	(0.013)**	(0.011)**	n.a.
Safety	-0.125	-0.124	-0.145	-0.074	-0.176
	(0.014)**	(0.014)**	(0.023)**	(0.021)**	(0.032)**
Own OOP	-0.006	-0.006	-0.006	-0.005	-0.007
	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Govt. Cost	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)**	(0.000)**	(0.000)	(0.000)*	(0.000)
Standard Deviations					
IV: *Tablet*	0.840	-0.788	1.013	-0.652	0.892
	(0.113)**	(0.106)**	(0.202)**	(0.234)**	(0.239)**
Injection: *Tablet*	0.472	-0.486	0.586	0.506	0.532
	(0.084)**	(0.081)**	(0.174)**	(0.131)**	(0.149)**
BD: *Monthly*	-0.209	0.128	0.253	0.378	0.112
	(0.127)	(0.419)	(0.180)	(0.345)	(0.142)
Wk: *Month*	0.239	-0.16	0.104	-0.335	0.190
	(0.100)*	(0.628)	(0.111)	(0.217)	(0.185)
Ft: *Month*	-0.240	-0.168	-0.228	0.039	0.495
	(0.105)*	(0.178)	(0.241)	(0.069)	(0.143)**
Efficacy	0.071	0.071	0.070	0.070	n.a.
	(0.007)**	(0.007)**	(0.009)**	(0.013)**	n.a.
Safety	0.145	0.141	0.089	-0.110	0.227
	(0.016)**	(0.015)**	(0.034)**	(0.027)**	(0.032)**
Own OOP	0.005	-0.005	-0.005	-0.005	-0.005
	(0.000)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Govt. Cost	-0.001	0.001	-0.001	0.001	0.001
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Observations	10,128	10,128	3,384	3,360	3,384
N	422	422	141	140	141
Wald Chi	190.15	187.89	83.45	83.19	56.23
d.f.	10	9	9	9	8
p-value	<0.001	0.00	0.00	0.00	0.00
Log-likelihood	-2,520.06	-2,528.49	-781.75	-853.24	-857.28
AIC	5078.12	5092.99	1599.49	1742.48	1746.56
BIC	5215.36	5223.01	1709.78	1852.63	1844.59

Notes: * p<0.05, ** p<0.01, *** p<0.001
All categorical data are effects coded to minimise the potential for contamination of the base effect in the grand mean.
Base categories are shown in italics.

All analyses were conducted in STATA 12, with robust standard errors (in parentheses) to account for the survey nature of the data.
Analyses conducted using 1,000 replications.
Abbreviations: AIC, Akaike Information Criterion; BD, twice daily; BIC, Bayesian Information Criterion; d.f., degrees of freedom; Ft, Fortnight; IV, Intravenous; n.a., not applicable; Wk, Week.

Including a Constant Term

One of the specifications tested here was the inclusion of a constant term in the analysis of the pooled data. Within that analysis, the constant term was significant and negative. However, caution is required in the interpretation of this coefficient due to the forced choice nature of the experimental design and the manner in which the constant has been included. That is, for the purposes of analysing stated preference data, the effect of a given variable on choices made by an individual is considered only where there is some variation of that variable in all or some subset of choices in the analysis set for each individual.[56] The *mixlogit* command in STATA does not have an option to include a constant term. Thus one was generated by creating a sequence of 1s corresponding to when Option A appeared in the data and 0s when Option B appeared in the data. This in essence creates an 'alternative specific constant' tied to Option A, but that in this instance carries no alternative specific information other than the fact that Option A always appeared on the left of the screen when respondents were completing the choice tasks. The negative and statistically significant coefficient on this constant therefore suggests some right-sided bias on the part of respondents in their preferred choice sets.

Overall, the mean coefficients and s.d. on the attributes are largely unaffected by the inclusion of the constant term in that specification; the estimates produced by the pooled regression with and without the constant are consistent with one another. Thus, given the difficulty of interpreting the result on the constant term, its apparent lack of influence on the attribute estimates, and that a constant was excluded from the initial design, a constant is excluded from the remaining analyses of the choice data.

Logarithmic Costs

Within this analysis, *Own OOP* and *Govt. Cost* have both entered the analysis as continuous variables, expressed as dollars. The effect of coding these two cost attributes in logarithmic form was tested in a sensitivity analysis (Table A 25). These results show that both the cost variables remain significant, as do the coefficients for the main attributes of significance; *IV: Tablet*, *Efficacy* and *Safety*. There are minor changes to two of the frequency attributes across all three framing sets (separately); *Fortnight: Monthly* is now significant for the 'Base' frame, and *Daily: Monthly* is no longer significant for the 'Attribute' frame. Similarly, within the 'Attribute' frame, *Injection: Tablet* becomes significant with the inclusion of logged costs. However, these changes are not consistent across the frames. Moreover, the cost attributes are acting as proxies for possible cost levels (consistent with expressing them in dollars) rather than an actual distribution of existing market prices (consistent with expressing them as logarithmic). Since the estimates on the main attributes of significance remain unchanged, cost can be assumed to be linear without unduly influencing the resulting mWTP.[275] Accordingly, the remaining analyses are based on the use of cost specified in dollars, not logarithms.

5.3.2.5.1 Tests of pooling

The results of the LR tests for whether that data can be pooled across the three framing sets ('Base', 'No-Efficacy Difference' and 'Attribute') are presented in Table 23. Based on the results from the two-way LR tests the estimates from the three framing sets are different; the LR test of poolability is rejected. In addition, the LR test of the comparison of all three samples with those from the pooled regression is statistically significant, indicating that the data from the three framing sets should be analysed separately.[259,271]

Table 23: LR tests of ability to pool samples

	LLH from Mixed logit	**d.f.**	**LR Test X^2 Statistic (d.f.)**		
			No-Efficacy Difference	*Attribute*	*Pooled*
Base	-781.75	9	151.06 *** (1)	142.98*** (1)	
No-Efficacy Difference	-857.28	8		8.08**(1)	
Attribute	-853.24	9			
Summed	-2492.27	26			72.45*** (17)
Pooled	-2,528.50	9			

Notes: * p<0.05, ** p<0.01, *** p<0.001
LR Test for the comparison of the framing sets was constructed as 2*(LLHFrame1-LLHFrame2) with d.f. equal to the difference of the d.f. of the respective framing sets. This means that for the Base and Attribute framing sets which each have 9 d.f. a X^2 test statistic could not be estimated for d.f.=0, so is approximated with d.f.=1.

Abbreviation: d.f., degrees of freedom; LLH, log-likelihood; LR, likelihood ratio.

The results of the Chow test support the findings of the LR tests that the data from the three framing sets are best considered separately. The overall test on whether all the sample specific variables (interacted with the attributes) are jointly zero was rejected with $X^2(19)= 39.40$; p=0.004. The overall frame specific variables (*Attribute: Base* and *Efficacy: Base*) were not statistically significant. However, a number of the frame specific interactions were significant; there is a statistically significant difference between the 'Attribute' and 'Base' frames in the effect of mode (*IV: Tablet* and *Injection: Tablet*) and *Safety* on choice, and for the No-Efficacy Difference and 'Base' frames in the effect of mode (*Injection: Tablet*) on choice.

Table 24: Pooled regression with sample specific interactions (Chow test)

Variables	Main Effects		Framing Interactions: Compared with Base	
	Coefficients	**s.d**	**No-Efficacy Difference**	**Attribute**
IV: *Tablet*	-0.809	0.876	0.188	-0.268
	(0.100)**	(0.110)**	(0.128)	(0.130)*
Injection: *Tablet*	-0.044	-0.465	-0.268	0.180
	(0.062)	(0.071)**	(0.080)**	(0.087)*
BD: *Monthly*	-0.456	-0.206	-0.027	-0.124
	(0.092)**	(0.162)	(0.120)	(0.133)
Weekly: *Monthly*	-0.049	0.280	-0.085	-0.053
	(0.049)	(0.095)**	(0.069)	(0.072)
Fortnightly: *Mont*	0.061	0.217	0.094	0.060
	(0.053)	(0.118)	(0.072)	(0.073)
Efficacy	0.082	0.073	n.a.	-0.011
	(0.008)**	(0.008)**	n.a.	(0.006)
Safety	-0.127	0.149	-0.025	0.053
	(0.013)**	(0.015)**	(0.018)	(0.016)**
Own OOP	-0.006	0.005	-0.001	0.001
	(0.000)**	(0.000)**	(0.000)	(0.000)
Govt. Cost	-0.000	0.001	0.000	-0.000
	(0.000)**	(0.000)**	(0.000)	(0.000)
Attribute: *Base*	-0.015			
	(0.082)			
Efficacy: *Base*	0.038			
	(0.084)			
Observations	10,128			
N	422			
Wald Chi	241.09			
d.f.	28			
p-value	0.00			
Log-likelihood	-2,501.77			
AIC	5077.54			
BIC	5344.80			

Notes: * p<0.05, ** p<0.01, *** p<0.001
All categorical data are effects coded to minimise the potential for contamination of the base effect in the grand mean.
Base categories are shown in italics.
All analyses were conducted in STATA 12, with robust standard errors (in parentheses) to account for the survey nature of the data.
Analyses conducted using 1,000 replications.

Abbreviations: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; d.f., degrees of freedom; n.a., not applicable; s.d., standard deviation.

5.3.2.6 Understanding choices: GMNL

5.3.2.6.1 Determinants of choice

On the basis of the significant heterogeneity observed in the mixed logit regression, the choice data were tested using the modelling specification in expression (16), without specifying a value for γ (the scaling parameter applied in the GMNL model that specifies the degree of independence between the coefficient and error variances), and by invoking the *gmnl* command within STATA. However, such a specification is data intensive, and as the analysis was run by framing set ('Base', 'Attribute' and 'No-Efficacy Difference'), it failed to converge with this specification. The alternative, in which γ =1, specifying the GMNL-I model, was therefore executed and produced results for all three framing sets (with 1,000 replications). The coefficient estimates are provided numerically in Table 25 and graphically in Figure 24.

Across all three framing sets, respondents have a preference for the simplest mode of administration over the more complicated, illustrated by the negative and significant coefficient on the *IV: Tablet* variable. However, differences in preferences for the injection mode of administration arise only for the 'No-Efficacy Difference' frame; being also in favour of the tablet modality – respondents in this framing set were less likely to choose an option for which the mode of delivery was an injection compared with a tablet. A similar pattern can be observed for the frequency attributes; across all three frames respondents are less likely to choose an option administered on a twice-daily basis compared with a monthly basis (*BD: Monthly)*. This preference for less frequent administration extended to weekly compared with monthly treatments (*Weekly: Monthly*) in the 'No-Efficacy Difference' frame.

For both frames in which efficacy differences appeared, respondents were more likely to choose those options associated with a higher likelihood of an improvement in their condition; reflected in the positive and significant coefficient on *Efficacy*. Respondents were less likely to choose options associated with a higher chance of their experiencing side effects that might necessitate cessation of treatment. In the 'Attribute' frame,

Safety does not achieve significance at the 5% level, but this might be a sample size issue since it is significant at the 10% level (p=0.059). Respondents were also less likely to choose options that resulted in higher monthly OOP costs to them. This was statistically significant across all three frames. However, the effect of *Govt. Costs* was only statistically significant for the 'Attribute' frame.

Table 25: GMNL estimation – testing framing effects

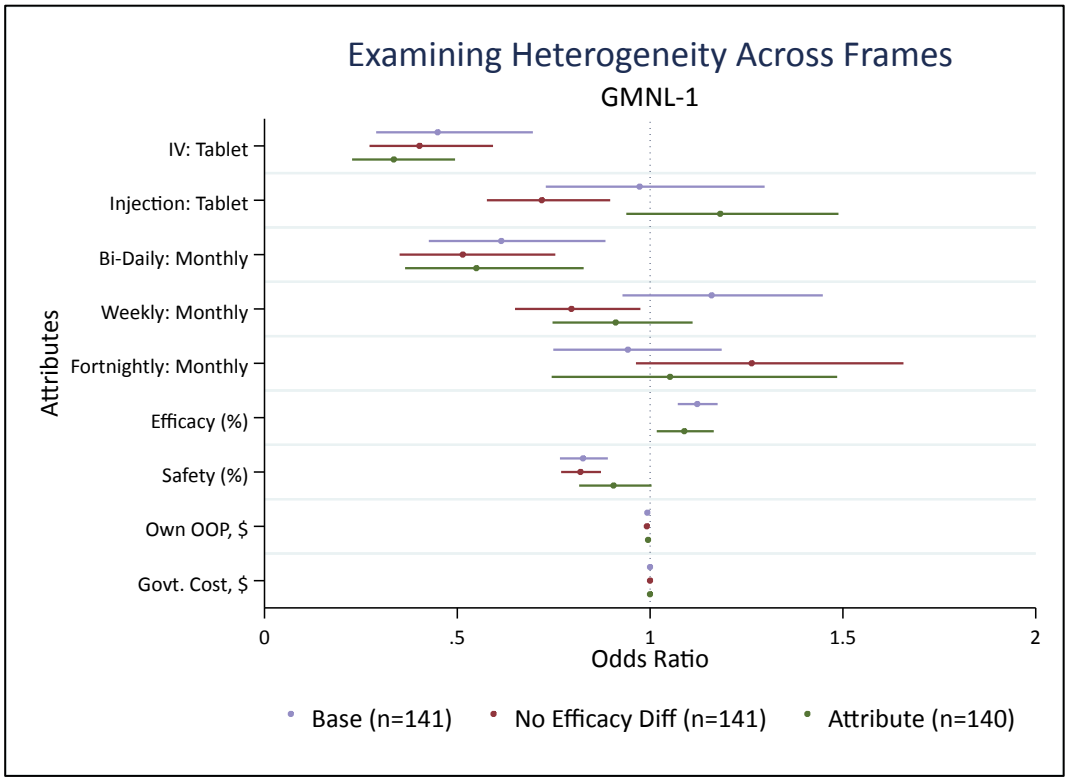
	Base	Attribute	No-Efficacy Difference
Coefficient Means			
IV: *Tablet*	-0.801	-1.093	-0.912
	(0.224)**	(0.198)**	(0.198)**
Injection: *Tablet*	-0.028	0.167	-0.330
	(0.147)	(0.118)	(0.113)**
BD: *Monthly*	-0.488	-0.599	-0.666
	(0.186)**	(0.209)**	(0.196)**
Weekly: *Monthly*	0.148	-0.094	-0.228
	(0.113)	(0.101)	(0.104)*
Fortnightly: *Monthly*	-0.059	0.050	0.234
	(0.117)	(0.176)	(0.138)
Efficacy	0.115	0.085	n.a.
	(0.023)**	(0.035)*	n.a.
Safety	-0.191	-0.100	-0.199
	(0.038)**	(0.053)	(0.032)**
Own OOP	-0.007	-0.005	-0.009
	(0.001)**	(0.001)**	(0.001)**
Govt. Cost	-0.0002	-0.0002	-0.0002
	(0.000)	(0.000)*	(0.000)
Standard Deviations			
IV: *Tablet*	-1.009	0.662	0.887
	(0.183)**	(0.217)**	(0.232)**
Injection: *Tablet*	0.502	-0.547	0.533
	(0.169)**	(0.134)**	(0.145)**
BD: *Monthly*	0.137	0.588	-0.126
	(0.143)	(0.190)**	(0.150)
Weekly: *Monthly*	-0.013	-0.394	-0.211
	(0.101)	(0.148)**	(0.139)
Fortnightly: *Monthly*	-0.144	0.007	0.459
	(0.237)	(0.083)	(0.162)**
Efficacy	-0.053	0.059	n.a.
	(0.011)**	(0.017)**	n.a.
Safety	0.066	0.117	0.191
	(0.031)*	(0.030)**	(0.029)**
Own OOP	0.003	0.004	-0.004
	(0.001)**	(0.002)*	(0.001)**
Govt. Cost	0.001	0.001	-0.001
	(0.000)**	(0.000)**	(0.000)**
Tau (constant)	0.951	0.755	0.782
	(0.197)**	(0.635)	(0.158)**
Observations	3,384	3,360	3,384
N	141	140	141

	Base	Attribute	No-Efficacy Difference
Wald Chi	57.21	66.76	70.18
d.f.	9	9	8
p-value	<0.0001	<0.0001	<0.0001
Log-likelihood	-774.61	-848.23	-851.35
AIC	1587.21	1734.46	1736.70
BIC	1703.62	1850.73	1840.85

Notes: * p<0.05, ** p<0.01, *** p<0.0001
All categorical data are effects coded to minimise the potential for contamination of the base effect in the grand mean.
Base categories are shown in italics.
All analyses were conducted in STATA 12, with robust standard errors (in parentheses) to account for the survey nature of the data.
Analyses conducted using 1,000 replications.

Abbreviations: d.f., degrees of freedom; s.d. standard deviation; AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; IV, intravenous; n.a., not applicable; OOP, out-of-pocket.

Figure 24: GMNL coefficients



Notes: The omitted level for each independent variable is shown in italics alongside the category name

Abbreviations: GMNL-I, generalised multninomial logit-I; IV, intravenous; OOP out-of-pocket; Wk, weekly.

Heterogeneity in Preferences

In addition to considering the respective mean coefficient estimates, it is also worth considering the extent of variation across the three framing sets in terms of the heterogeneity exhibited by respondents in how they were influenced by the attributes. There was variability in response to all attributes insofar as in all framing sets, the standard deviation associated with choice (the measure of heterogeneity) was significant for at least one level of each attribute.[xxviii] However, differences occurred across the framing sets in terms of the heterogeneity exhibited with respect to the frequency of administration attributes. There was no heterogeneity observed among respondents in the 'Base' frame for this variable, yet those in the 'Attribute' frame exhibited significant heterogeneity across two levels (*Bi-Daily: Monthly* and *Weekly: Monthly*) and those in the 'No-Efficacy Difference' frame across the *Fortnightly: Monthly* level.

Despite the individual attribute differences in heterogeneity across the framing sets, overall the mean estimate of τ, the indicator of the presence of scale effects, is not different across the three framing sets, and is only significant in the 'Base' and 'No-Efficacy Difference' framing sets (Table 25).

Evidence of Framing

Comparison of the choice coefficients across the three framing sets suggests that the type and manner in which information is presented influences choices. This is most apparent for the convenience related attributes which are 'closer' to the base level: *Injection: Tablet*, *Weekly: Monthly*, and *Fortnightly: Monthly*. For all three of these coefficients there is some hint of preference reversal between the framing sets. Those in the 'Attribute' frame appear to prefer injections to tablets (a positive coefficient), as compared with the other framing sets; respondents in the 'Base' frame preferred a weekly administration compared with monthly (where we would expect the opposite);

[xxviii] The negative signs on these standard deviations should be ignored; they are all interpretable as being positive variations.[243]

and those in the 'Attribute' and 'No-Efficacy Difference' framing sets preferred a fortnightly administration over monthly (which again, goes against expectations).

While the confidence intervals on these estimates overlap, this may be an artefact of sample size. Indeed, results of pairwise-LR tests of the three sets of estimates suggest that they cannot be considered to be the same: 'Base' vs 'Attribute' $\chi^2(1)=$ 147.25,$p<0.001$ (estimated based on d.f.=1, rather than d.f.=0 for which a test statistic is not available); 'Base' vs 'No-Efficacy Difference' $\chi^2(2)=$ 153.48,$p<0.001$, and 'Attribute' vs 'No-Efficacy Difference' $\chi^2(2)=$ 6.24,$p=0.04$. However, the presence of framing effects is not supported by the scale coefficients which did not differ across the three versions.

5.3.2.6.2 Comparing GMNL with Mixed Logit

Overall, the use of the GMNL-I to characterise the choice of RA treatments performs well compared with the mixed logit regression in that it not only captures individual heterogeneity, but it directly addresses scale, and does so more efficiently. This is evidenced by the improved (smaller) LLH, AIC and BIC statistics for the GMNL-I compared with the mixed logit regression for all three framing sets (see Table 26).

Table 26: Comparison of GMNL-I and mixed logit model performance

	Base		**Attribute**		**No-Efficacy Difference**	
	Mixed logit	GMNL-I	Mixed logit	GMNL-I	Mixed logit	GMNL-I
LLH	-781.75	-774.61	-853.24	-848.23	-857.28	-851.35
AIC	1599.49	1587.21	1742.48	1734.46	1746.56	1736.70
BIC	1709.78	1703.62	1852.63	1850.73	1844.59	1840.85

Abbreviations: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; GMNL-I, generalised multinomial modelling; LLH, log likelihood.

5.3.2.6.3 Valuing determinants of choice

The mWTP has been estimated using the coefficient estimates from the GMNL-I estimation procedures for all framing sets. Point estimates of the mWTP and 95% CI are presented in Table 27 for those variables with a statistically significant choice

coefficient in at least one framing set. Estimates for the mWTP for all variables, regardless of their statistical significance, are presented graphically in Figure 25. In terms of the statistically significant estimates, the highest absolute value was for *IV: Tablet* in the 'Attribute' frame at $203.64. Within that frame, the mWTP for the two health effects was substantially lower at 7.8% and 9.1% of the relative value of *IV: Tablet* for *Efficacy* and *Safety* respectively. In contrast, while the mWTP for *IV: Tablet* was also the largest value for the other two frames, the relative size of the health effects was not as low at 14.4% for *Efficacy* in 'Base', and 23.9% and 21.9% for *Safety* in 'Base' and 'No-Efficacy Difference' respectively. Thus, there appears to be a difference across framing sets in the point estimates of the absolute mWTP and in their interpretation for the importance placed on the relative product attributes e.g. mode of administration compared with health effects.

Table 27: Estimated mWTP

	Base		Attribute		No-Efficacy Difference	
	mWTP $ (95% CI)	**Wgt (rnk)**	**mWTP** $ (95% CI)	**Wgt (rnk)**	**mWTP** $ (95% CI)	**Wgt (rnk)**
Mode						
IV: *Tablet*	-110.72 (-176.17 to -45.28)	1 (1)	-203.64 (-295.2 to -112.09)	1 (1)	-106.87 (-146.81 to -66.93)	1 (1)
Injection: *Tablet*	-3.83 (-43.43 to 35.76)	0.03 (6)	31.11 (-12.74 to 74.96)	0.15 (3)	-38.66 (-65.24 to -12.08)	0.36 (3)
Frequency						
BD: *Monthly*	-67.47 (-112.57 to -22.37)	0.61 (2)	-111.58 (-189.65 to -33.51)	0.55 (2)	-78 (-122.13 to -33.86)	0.73 (2)
Weekly: *Monthly*	20.44 (-9.4 to 50.29)	0.18 (4)	-17.44 (-55.98 to 21.09)	0.09 (5)	-26.77 (-49.22 to -4.33)	0.25 (4)
Health Effects						
Efficacy (per 1%)	15.94 (10.5 to 21.38)	0.14 (5)	15.82 (3.71 to 27.92)	0.08 (6)	0 (0-0)	
Safety (per 1%)	-26.43 (-35.71 to -17.15)	0.24 (3)	-18.58 (-37.11 to -0.05)	0.09 (4)	-23.36 (-31.39 to -15.33)	0.22 (5)

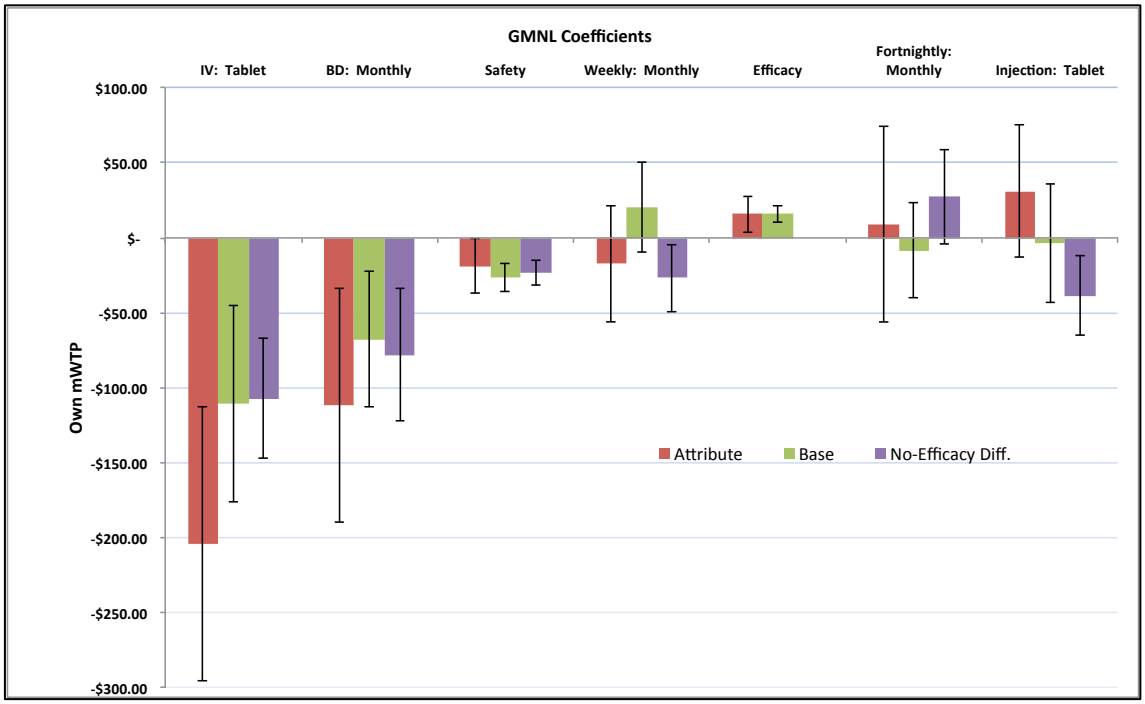
Note: mWTP is not shown for the *Fortnightly:Monthly* level of the frequency attribute as this was not significant in any of the frames.
All coefficient ranks are based on the absolute values of the mWTP compared with that of the largest mWTP for each frame (that associated with the move from Tablet to IV Infusion), and excludes the influence of fortnightly frequency in attribute ranking.

Abbreviations: CI, confidence interval; IV, intravenous; mWTP, marginal willingness to pay; rnk, rank; Wgt, weight.

Ranking Based on mWTP

The ranking of the attributes in terms of their relative value, as well as differences in those rankings across the framing sets, can be observed from the graphical representation of the mWTP (Figure 25). For example it is clear that *IV: Tablet* was associated with the highest mWTP across all frames, and while the confidence intervals of those estimates overlapped, the point estimates were nearly two fold higher for the 'Attribute' frame compared with the other two framing sets. In contrast, the value placed on the health effects was relatively consistent across all three framing sets, although the higher value for *Safety* relative to *Efficacy* is somewhat surprising given that respondents nominated *Efficacy* as the most important attribute. The same attribute ranking also results from the use of *Govt. Cost* to estimate the mWTP (Figure 26). However, the notable difference in this instance is that there is generally much greater consistency across the three frames in the point estimates of the mWTP values obtained for any given attribute.

Figure 25: mWTP – assessing convenience



Note: mWTP values are sorted from largest to smallest on the basis of the absolute values for the 'Base' frame.
Estimates formed by dividing the coefficient for each attribute by the coefficient on *Own OOP,* invoking the STATA *wtp* command. All confidence intervals produced using the Delta method.

Abbreviations: BD, twice daily; GMNL, generalised multinomial modelling; IV, intravenous; mWTP, marginal willingness to pay.

Figure 26: mWTP – based on government costs

Note: mWTP values are sorted from largest to smallest on the basis of the absolute values for the 'Base' frame.
Estimates formed by dividing the coefficient for each attribute by the coefficient on *Own OOP*, invoking the STATA *wtp* command. All confidence produced using the Delta method.

Abbreviations: BD, twice daily; GMNL, generalised multinomial modelling; IV, intravenous; mWTP, marginal willingness to pay.

Categorical Coding of Health Effects

The results incorporating *Efficacy* and *Safety* coded as categorical variables are provided in Table 28 and graphically in Figure 27. Of note, the overall regression for the 'Attribute' frame is no longer statistically significant (p=0.31), mirrored in that many of the attributes that were significant previously are no longer significant. This would appear to be due to the effect of the difference in frame, coupled with the small sample size and the number of replications (1,000); estimating the same specification with fewer replications (500) produced a statistically significant model (p<0.001; see Table A 26 in Appendix 9). Overall, convenience effects remained significant in the 'Base' and 'No-Efficacy Difference' frames, but not the 'Attribute' frame, as compared with the specifications in which *Efficacy* and *Safety* were included as continuous. In addition, variables that would have been expected to be significant (*Efficacy*) are not

statistically significant in the 'Attribute' frame, but remain so in the other two frames. However, some of the results for *Efficacy* and *Safety* are not as might be expected. For example, *Safety 5%:1%* has a positive (but not significant) coefficient for the 'Base' and 'Attribute' frames. Similarly, *Efficacy 60%:70%* is positive and significant in the 'Base' frame. In both cases, negative signs would have been expected. The high degree of variability across the frames induced through the use of categorical coding for *Efficacy* and *Safety* is evident in the graphical representation of the data (Figure 27).

Table 28: Categorical coding of health effects – GMNL-I

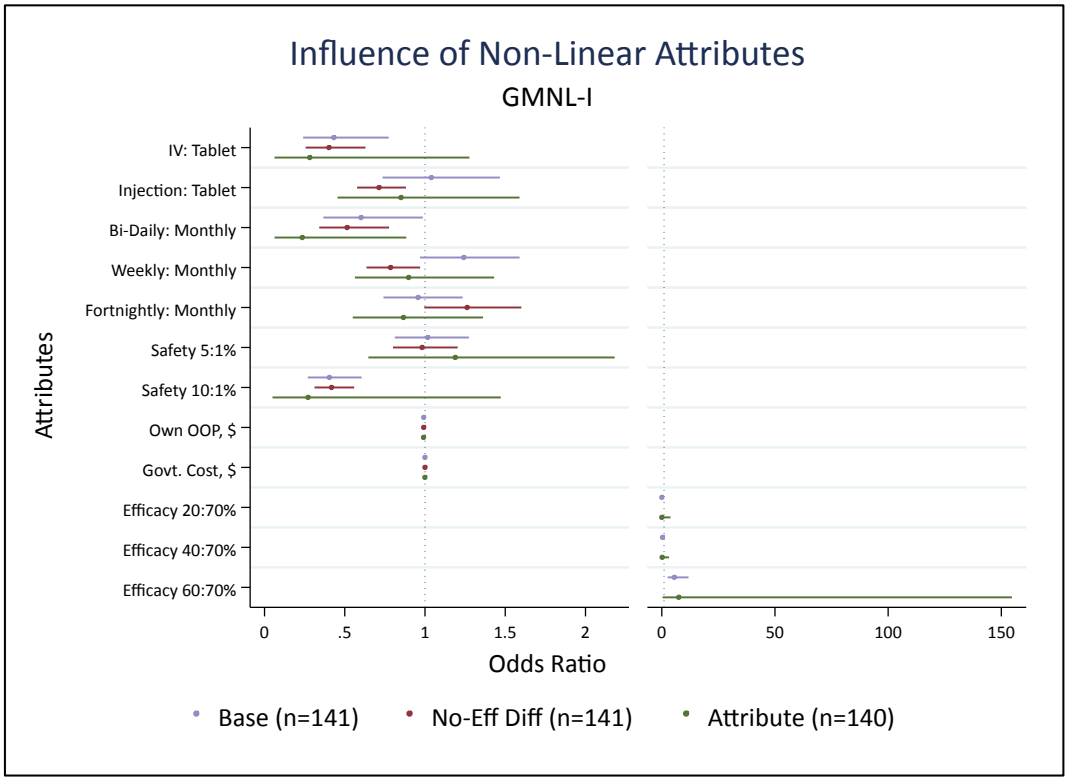
	Base	Attribute	No-Efficacy Difference
Coefficient Means			
IV: *Tablet*	-0.840	-1.268	-0.913
	(0.298)**	(0.772)	(0.229)**
Injection: *Tablet*	0.038	-0.162	-0.338
	(0.176)	(0.319)	(0.108)**
BD: *Monthly*	-0.508	-1.449	-0.665
	(0.252)*	(0.676)*	(0.210)**
Weekly: *Monthly*	0.216	-0.108	-0.242
	(0.126)	(0.238)	(0.108)*
Fortnightly: *Monthly*	-0.044	-0.145	0.233
	(0.130)	(0.231)	(0.121)
Efficacy 20%: *70%*	-3.522	-5.453	n.a.
	(0.749)**	(3.468)	n.a.
Efficacy 40%: *70%*	-1.073	-1.773	n.a.
	(0.310)**	(1.502)	n.a.
Efficacy 60%: *70%*	1.710	2.013	n.a.
	(0.387)**	(1.545)	n.a.
Safety 5%: *1%*	0.017	0.172	-0.018
	(0.115)	(0.310)	(0.104)
Safety 10%: *1%*	-0.907	-1.306	-0.874
	(0.207)**	(0.864)	(0.149)**
Own OOP	-0.008	-0.009	-0.008
	(0.002)**	(0.003)**	(0.001)**
Govt. Cost	-0.0002	-0.001	-0.0002
	(0.0001)	(0.001)	(0.0001)
Standard Deviations			
IV: *Tablet*	-1.163	0.902	0.847
	(0.228)**	(0.208)**	(0.211)**
Injection: *Tablet*	-0.589	0.523	0.474
	(0.169)**	(0.128)**	(0.118)**
BD: *Monthly*	-0.083	0.499	-0.012
	(0.101)	(0.261)	(0.221)
Weekly: *Monthly*	-0.100	-0.376	-0.224
	(0.141)	(0.140)**	(0.177)
Fortnightly: *Monthly*	-0.354	0.083	-0.497
	(0.162)*	(0.092)	(0.169)**

	Base	Attribute	No-Efficacy Difference
Efficacy 20%: *70%*	-0.916	0.680	n.a.
	(0.223)**	(0.264)**	n.a.
Efficacy 40%: *70%*	0.604	0.535	n.a.
	(0.251)*	(0.189)**	n.a.
Efficacy 60%: *70%*	0.139	0.248	n.a.
	(0.105)	(0.281)	n.a.
Safety 5%: *1%*	0.143	0.292	-0.013
	(0.195)	(0.141)*	(0.230)
Safety 10%: *1%*	-0.277	-0.483	-0.895
	(0.217)	(0.192)*	(0.148)**
Own OOP	-0.003	0.004	0.004
	(0.001)**	(0.001)**	(0.001)**
Govt. Cost	0.001	0.001	0.001
	(0.000)**	(0.000)**	(0.000)**
Tau (constant)	1.043	-1.657	-0.677
	(0.216)**	(0.388)**	(0.141)**
Observations	3,384	3,360	3,384
N	141	140	141
Wald Chi	51.32	13.78	67.60
d.f.	12	12	9
p-value	0.00	0.31	<0.0001
Log-likelihood	-773.72	-848.71	-849.12
AIC	1597.431	1747.419	1736.247
BIC	1750.601	1900.412	1852.656

Notes: * p<0.05, ** p<0.01, *** p<0.0001
All categorical data are effects coded to minimise the potential for contamination of the base effect in the grand mean.
Base categories are shown in italics.
All analyses were conducted in STATA 12, with robust standard errors (in parentheses) to account for the survey nature of the data.
Analyses conducted using 1,000 replications.

Abbreviations: BD, twice daily; d.f., degrees of freedom; AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; IV, intravenous; n.a., not applicable; OOP, out-of-pocket.

Figure 27: GMNL-I coefficients with categorical health effects



Abbreviations: GMNL, generalised multinomial logit; IV, intravenous; OOP out-of-pocket.

Linearity of the *Efficacy* and *Safety* attributes was tested within each framing set using Wald tests[56] in two ways: an unscaled Wald test of whether the coefficients for each of the levels within the respective attributes (*Efficacy* and *Safety*) were the same within a given attribute; and a scaled Wald test to account for the fact that each level represented a larger increment from the base level, making it possible that the coefficient values should be scaled in order to test for equivalence. To illustrate the scaled Wald test, for the *Efficacy* variable the increment between the 20% level and 70% is 50 percentage points, while for the 40% level it is 30%. To account for this difference in the increment considered by respondents, the scaled coefficient test adjusts the coefficient on the 40% level by a factor of 50/30.

The results of the Wald tests of linearity (Table 29) indicate that for the 'Base' and 'No-Efficacy Difference' frames it would be reasonable to treat *Efficacy* and *Safety* as

categorical variables; that is the coefficients for the individual attribute levels are significantly different from one another. In contrast, the Wald tests for the 'Attribute' frame cannot be rejected, indicating that the attribute coefficients are not statistically significantly different and can therefore be treated as continuous. The difference in the outcomes of linearity tests reinforces the importance of acknowledging the differences that can be introduced into analyses of choice by the manner in which information is presented.

Table 29: Tests of coefficient linearity

	Base	Attribute	No-Efficacy Difference
Efficacy Attribute			
Wald	20.19 (p<0.001)	2.58 (p=0.108)	n.a.
Scaled Wald	20.36 (p<0.001)	2.07 (p=0.151)	n.a.
Safety Attribute			
Wald	13.28 (p<0.001)	1.77 (p=0.184)	16.40 (p<0.001)
Scaled Wald	6.95 (p<0.01)	1.35 (p=0.245)	6.73 (p<0.01)

Notes: The Wald test was estimated using the following linear combinations:
1. For the efficacy attribute: Efficacy 20%:70% - Efficacy 40%:70% - Efficacy 60%:70% = 0
2. For the safety attribute: Safety 5%:1% - Safety 10%:1% = 0
The Scaled Wald test was estimated using the following linear combinations:
1. For the efficacy attribute: Efficacy 20%:70% - (Efficacy 40%:70%)*(5/3) – (Efficacy 60%:70%)*5 = 0
2. For the safety attribute: (Safety 5%:1%)*(9/4) - Safety 10%:1% = 0

Abbreviation: n.a., not applicable.

Impact of Categorically Coding Health Effects on mWTP

The mWTP for those attributes for which there was a significant coefficient are presented in Table 30. The mWTP results for all attributes, regardless of significance, are presented graphically in Figure 28 and Figure 29. Although *Govt. Cost* was not significant in any of the GMNL results using categorically coded health effects, mWTP estimates using *Govt. Cost* as the numeraire are provided for completeness.

Using the results from the specifications in which *Efficacy* and *Safety* are coded as categorical variables, the estimates of mWTP for the meta-health effects remain largely unchanged from those in Table 27. This is expected since the underlying coefficients

for those variables are relatively unchanged, suggesting no interaction between them and the *Efficacy* and *Safety* attributes. However, the mWTP for *Efficacy* and *Safety* change substantially due to the shift from an estimate for the overall attribute (when coded as being continuous) to one for comparisons between levels of the respective attributes. This not only changes the magnitude of the resulting mWTP estimates, but also the resulting rank order of those attributes in terms of their contribution to overall value. Where *Efficacy* was considered by respondents, it was ranked highest in terms of the resulting mWTP, relegating the meta-health effects to the lower orders.

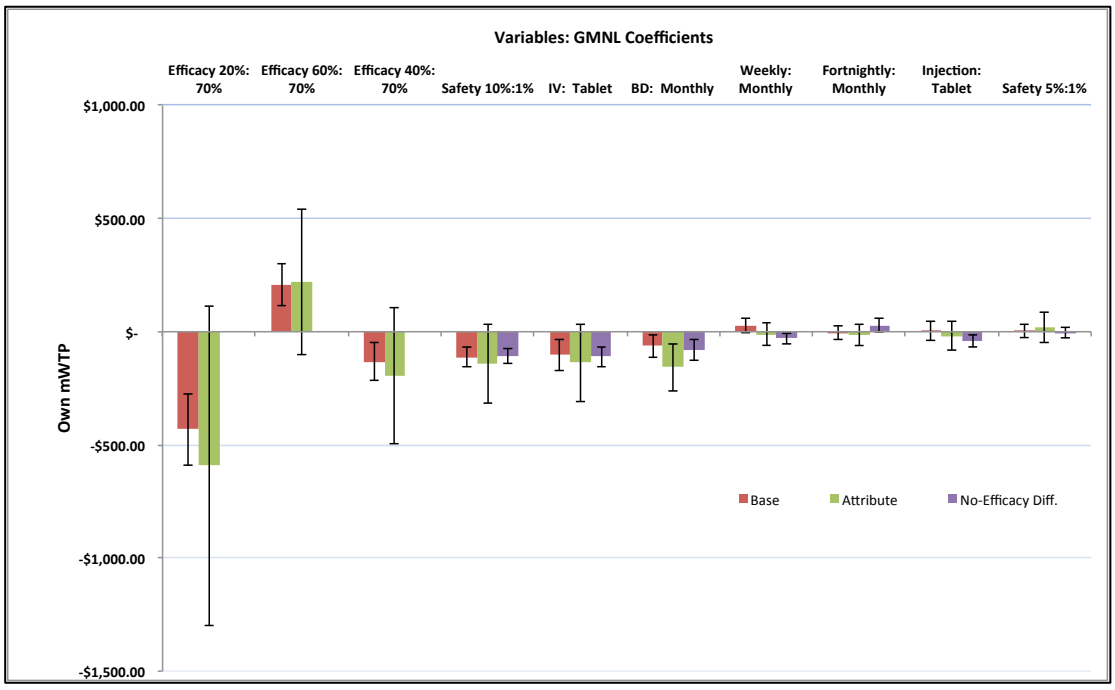
Coding the health effects *Efficacy* and *Safety* as categorical therefore results in a lower relative contribution of meta-health effects to value, or conversely, the use of continuous coding for health effects results in an overestimation of the contribution of meta-health effects to value. That is, with the exception of the 'Attribute' frame, categorical coding of health effects appears to better capture the difference between overall products as reviewed by respondents in DCEs, rather than small increments at the margin. Moreover, it does not rely on the implicit assumption made when health effects are coded as continuous that respondents interpret differences between the levels of health effects, like *Efficacy*, as being cardinal when answering choice questions.

Table 30: Impact of categorically coding health effects on mWTP

	Base		Attribute		No-Efficacy Difference	
	mWTP	**Wgt (rnk)**	**mWTP**	**Wgt (rnk)**	**mWTP**	**Wgt (rnk)**
Mode						
IV: *Tablet*	-102.96 (-171.22 to -34.7)	0.24 (5)	-137.67 (-309.04 to 33.71)	0.23 (6)	-110.7 (-155.07 to -66.34)	1 (1)
Injection: *Tablet*	4.66 (-38.07 to 47.39)	0.01 (8)	-17.62 (-81.62 to 46.37)	0.03 (7)	-40.98 (-67.23 to -14.74)	0.37 (4)
Frequency						
BD: *Monthly*	-62.31 (-112.91 to -11.71)	0.14 (6)	-157.35 (-263.92 to -50.79)	0.27 (4)	-80.58 (-126.05 to -35.1)	0.73 (3)
Weekly: *Monthly*	26.48 (-4.00 to 56.96)	0.06 (7)	-11.73 (-60.17 to 36.71)	0.02 (8)	-29.32 (-53.70 to -4.94)	0.26 (5)
Efficacy						
20%: *70%*	-431.66 (-589.68 to -273.64)	1 (1)	-592.22 (-1297.47 to 113.02)	1 (1)	n.a.	n.a.
40%: *70%*	-131.47 (-215.57 to -47.38)	0.3 (3)	-192.55 (-494.45 to 109.36)	0.33 (3)	n.a.	n.a.
60%: *70%*	209.55 (115.52 to 303.58)	0.49 (2)	218.57 (-103.54 to 540.69)	0.37 (2)		
Safety						
10%: *1%*	-111.16 (-157.25 to -65.06)	0.26 (4)	-141.8 (-318.15 to 34.55)	0.24 (5)	-105.95 (-139.82 to -72.09)	0.96 (2)

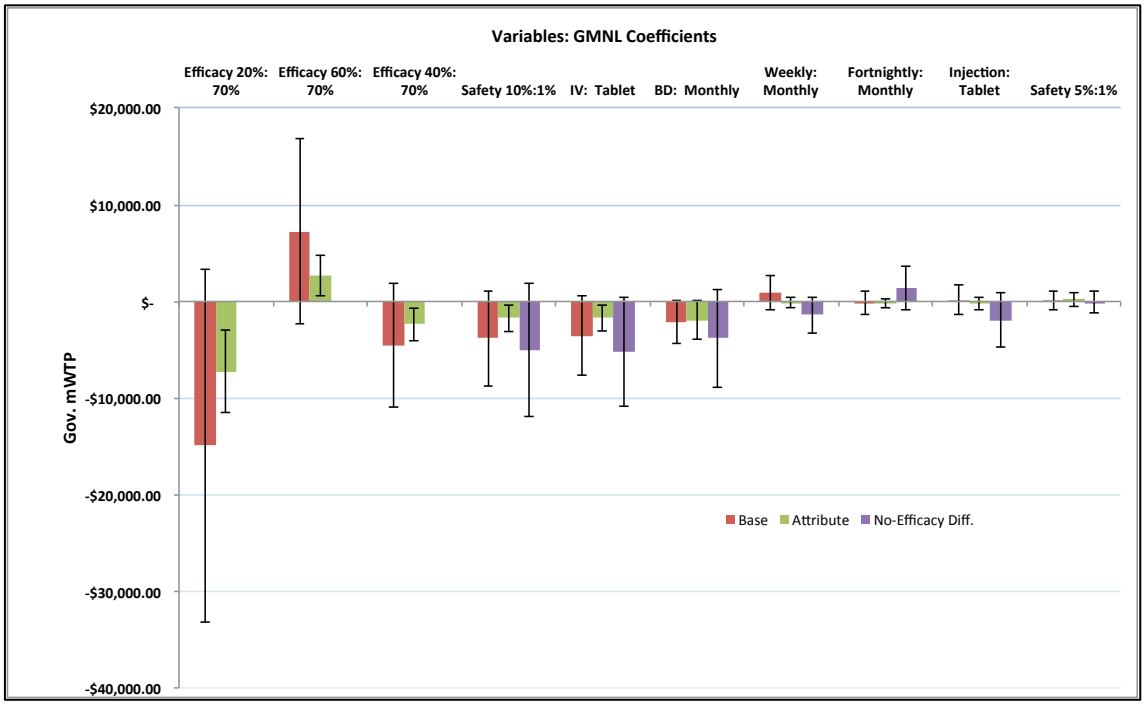
Abbreviation: BD, twice daily; CI, confidence interval; IV, intravenous; rnk, rank; Wgt, weight; WTP, willingness to pay.

Figure 28: mWTP – categorically coded efficacy & safety – own OOP



Abbreviation: BD, twice daily; GMNL, generalised multinomial logit; IV, intravenous; mWTP, marginal willingness to pay; OOP, out-of-pocket.

Figure 29: mWTP – categorically coded efficacy & safety – government costs



Abbreviation: BD, twice daily; GMNL, generalised multinomial logit; IV, intravenous; mWTP, marginal willingness to pay; OOP, out-of-pocket.

As previously noted, the impact on the mWTP of respondents comparing scenarios based on the complete profile of differences in health effects (for example *Efficacy* 40% compared with 70%) as derived in the analysis using categorical coding of those effects could also be estimated by multiplying the mWTP for a one-point change in the health effects by the respective difference between health effect categories. mWTP values estimated using this method are presented in Table 31, alongside those produced using the estimates based on the categorically coded health effects (from Table 30).

Table 31: Comparison of mWTP results for health effects

	Base, $		Attribute, $		No-Efficacy Difference, $	
	Categorical Coding	Product Method	Categorical Coding	Product Method	Categorical Coding	Product Method
Efficacy						
20%: *70%*	-431.66		-592.22			
	(-589.68 to -273.64)	-796.92	(-1297.47 to 113.02)	-790.86	n.a.	n.a.
40%: *70%*	-131.47		-192.55			
	(-215.57 to -47.38)	-478.15	(-494.45 to 109.36)	-474.51	n.a.	n.a.
60%: *70%*	209.55		218.57			
	(115.52 to 303.58)	-159.38	(-103.54 to 540.69)	-158.17	n.a.	n.a.
Safety						
5%: *1%*	2.03		18.72		-2.23	
	(-25.75 to 29.81)	-105.72	(-50.29 to 87.72)	-74.33	(-26.93 to 22.46)	-93.44
10%: *1%*	-111.16		-141.8		-105.95	
	(-157.25 to -65.06)	-237.88	(-318.15 to 34.55)	-167.24	(-139.82 to -72.09)	-210.23

Note: Estimates for the product method are produced by multiplying the difference between the relevant health profiles of interest, with the mWTP of the one-point change in that health effect as reported in Table 27.
Differences in health profiles were signed according to their relevant decrement (e.g. moving from Efficacy 70% to 20% was assigned a value of -50 points) or increment (e.g. moving from Safety 1% to 5% an increment of +4 points).

Abbreviation: mWTP, marginal willingness to pay; n.a., not applicable.

From these estimates it can be observed that assuming linearity of effects, and using the mWTP for a one-point change to estimate the mWTP associated with the difference between the health effects for different product profiles, overstates the mWTP relative to what is produced under the assumption of non-linear effects (categorically coded data). That is, where the health effects are not linear (the Base and 'No-Efficacy Difference' frames), using the mWTP for a one-point change to estimate the mWTP associated with product profile change produces an estimated mWTP that lies outside of the bounds of that produced by assuming non-linear health effects. This was not the case for the 'Attribute' frame where linearity of the health effects was preserved (and can therefore be assumed to be modelled as continuous variables).

5.3.3 Understanding the drivers of choice: predicted probabilities

As outlined in the methods section, the predicted probabilities produced by the GMNL-I estimations have been explored in two ways to better understand the factors that influence choice in the context of treatments for RA. The first is a descriptive

analysis in which the predicted probabilities from the GMNL-I estimations are summarised across the choice sets for each of the respective attribute levels (and thus only reflect the mean). These are not simulations of the predicted probabilities produced by varying the dataset to introduce new attribute levels, but rather take the mean of the predicted probabilities by restricting the choice sets to those for when each respective attribute level appears. The difference between that mean, and the overall mean predicted probability per frame, is interpreted as being the influence on the choice probability of each attribute level, given the occurrence of all the other attribute levels in the design. The second way in which the predicated probabilities have been explored is via a regression analysis to consider the role of respondents' demographic characteristics in predicting the probability of choice.

All predicted probabilities were generated using the coefficient estimates produced using the main GMNL-I results as reported in Table 25. The STATA command *gmnlpred* was invoked to produce the predicted probabilities, limiting the sample of respondents to those for the relevant survey version (e.g. 'Base'), and using 1,000 replications.

5.3.3.1 Attribute levels and predicted probabilities

The average predicated probabilities for each framing set, reported by attribute level are presented in Table 32, with the impact for each attribute on the probability of choice depicted graphically in Figure 30 to Figure 32. What can be readily observed from these data is that the largest influences on the probability of choice are *Efficacy* and *Own OOP*; both are associated with the largest increments in the attribute-specific average predicted probability compared with the overall average.

As a group, the meta-health effects have the least effect on the predicted probabilities of choice, but again we observe that the magnitude of those effects varies across the framing sets. Overall, meta-health effects are most influential in the 'No-Efficacy Difference' frame which is to be expected; in the absence of a difference in *Efficacy*,

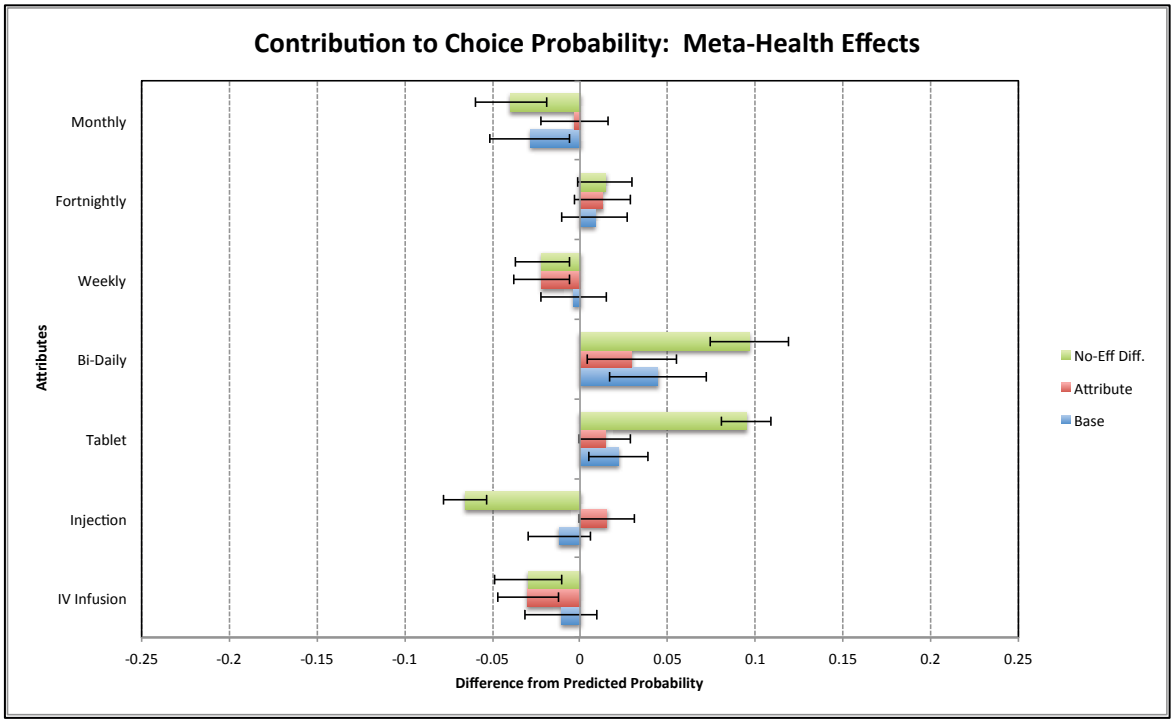
respondents shift their attention to the meta-health effects. Similarly, in the 'Attribute' frame where more information is provided about the mode of administration, we observe a larger impact for *IV* and *Injection*, compared with Base. In both cases therefore, the amount of information provided influences the responses.

Table 32: Mean predicted probabilities

	Base Mean (95% CI)	**Attribute** Mean (95% CI)	**No-Efficacy Difference** Mean (95% CI)
Overall	*0.50 (0.49 to 0.51)*	*0.51 (0.49 to 0.5)*	*0.50 (0.49 to 0.51)*
Meta to Health Effects			
IV Infusion	0.49 (0.47 to 0.51)	0.49 (0.45 to 0.47)	0.47 (0.45 to 0.49)
Injection	0.49 (0.47 to 0.50)	0.53 (0.50 to 0.52)	0.43 (0.42 to 0.44)
Tablet	0.52 (0.51 to 0.54)	0.53 (0.50 to 0.51)	0.59 (0.58 to 0.61)
BD	0.54 (0.52 to 0.57)	0.55 (0.51 to 0.53)	0.60 (0.58 to 0.62)
Weekly	0.50 (0.48 to 0.51)	0.49 (0.46 to 0.48)	0.48 (0.46 to 0.49)
Fortnightly	0.51 (0.49 to 0.52)	0.53 (0.50 to 0.51)	0.51 (0.50 to 0.53)
Monthly	0.47 (0.45 to 0.49)	0.51 (0.48 to 0.50)	0.46 (0.44 to 0.48)
Health Effects			
Efficacy 20%	0.29 (0.28 to 0.31)	0.33 (0.31 to 0.32)	0.50 (0.48 to 0.52)
Efficacy 40%	0.52 (0.50 to 0.53)	0.52 (0.49 to 0.51)	0.50 (0.48 to 0.52)
Efficacy 60%	0.51 (0.49 to 0.53)	0.52 (0.48 to 0.50)	0.50 (0.48 to 0.52)
Efficacy 70%	0.68 (0.67 to 0.70)	0.68 (0.66 to 0.67)	0.50 (0.49 to 0.51)
Safety 1%	0.58 (0.57 to 0.60)	0.56 (0.53 to 0.54)	0.62 (0.60 to 0.63)
Safety 5%	0.48 (0.46 to 0.49)	0.51 (0.48 to 0.49)	0.51 (0.49 to 0.52)
Safety 10%	0.44 (0.43 to 0.46)	0.48 (0.45 to 0.46)	0.38 (0.37 to 0.39)
Costs			
OOP $0	0.74 (0.73 to 0.75)	0.70 (0.68 to 0.69)	0.70 (0.68 to 0.71)
OOP$40	0.56 (0.54 to 0.57)	0.55 (0.53 to 0.54)	0.54 (0.52 to 0.55)
OOP $250	0.42 (0.41 to 0.44)	0.46 (0.43 to 0.45)	0.48 (0.47 to 0.49)
OOP $500	0.30 (0.29 to 0.32)	0.35 (0.32 to 0.33)	0.30 (0.29 to 0.31)
Gov.Costs $0	0.38 (0.36 to 0.40)	0.43 (0.40 to 0.42)	0.41 (0.39 to 0.42)
Gov.Costs $500	0.50 (0.48 to 0.52)	0.51 (0.47 to 0.49)	0.47 (0.45 to 0.49)
Gov.Costs $1500	0.55 (0.54 to 0.57)	0.55 (0.52 to 0.53)	0.53 (0.51 to 0.54)
Gov.Costs $3000	0.54 (0.52 to 0.55)	0.55 (0.52 to 0.54)	0.57 (0.56 to 0.58)

Abbreviation: BD, twice daily; CI, confidence interval; IV, intravenous; OOP out-of-pocket.

Figure 30: Impact on choice probability – meta-health effects



Abbreviation: IV, intravenous.

Figure 31: Impact on choice probability – health effects

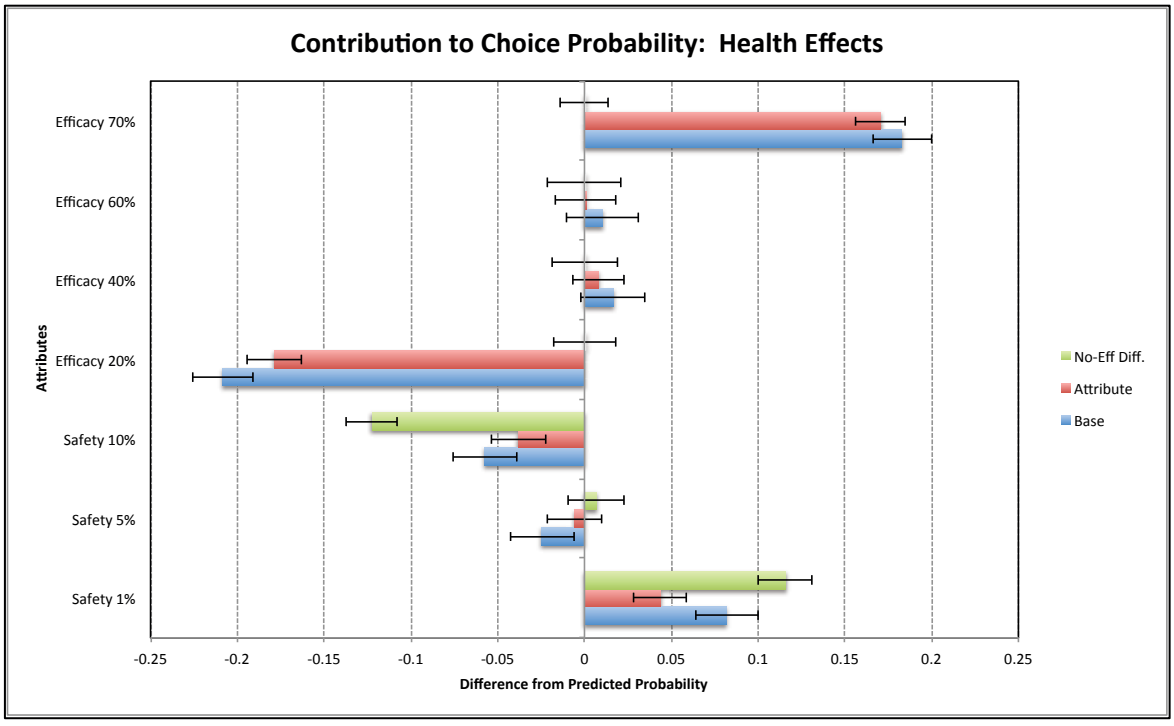
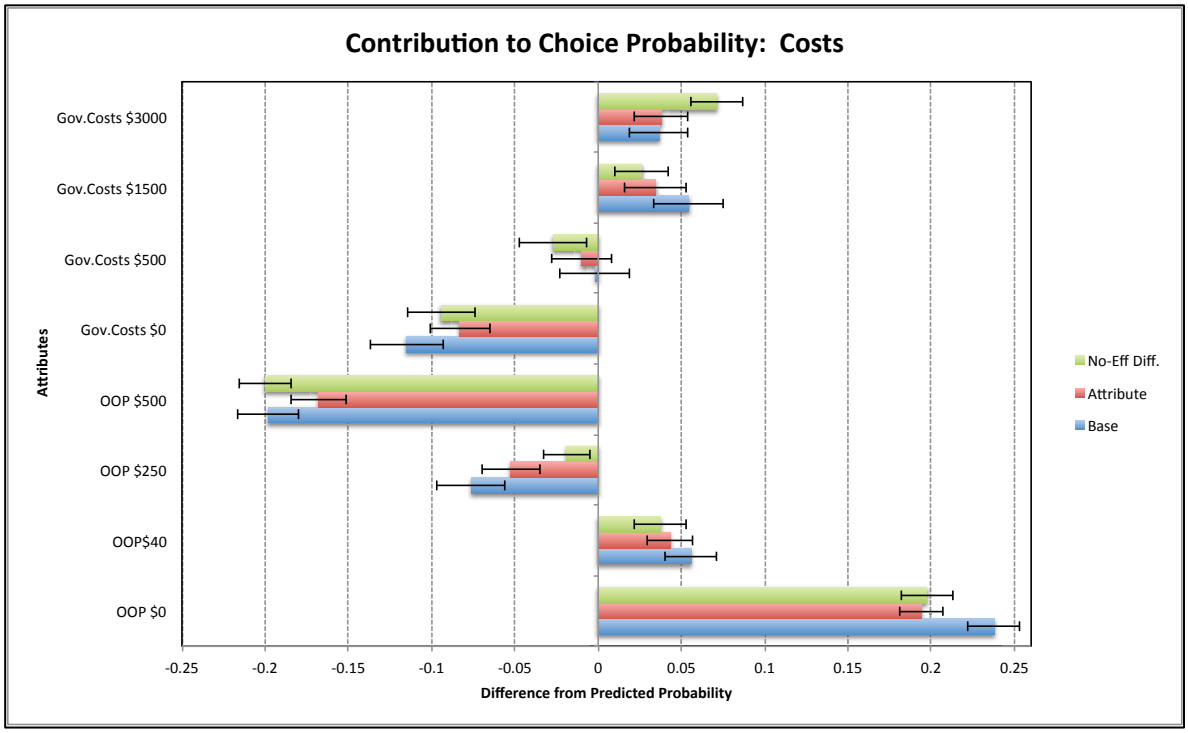


Figure 32: Impact on choice probability – costs



Abbreviation: OOP, out-of-pocket.

5.3.3.2 Demographics and predicted probabilities

The results of the regression analyses of the influence of demographic characteristics on the predicted probabilities are presented in Table 33. These are the coefficient estimates from a series of OLS regressions in which the predicted choice probabilities (for each choice task) for each individual are regressed on the corresponding demographic characteristics of interest, controlling for the experimental design attributes and the repeated nature of the choice tasks. Across all three framing sets, we observe that these demographic characteristics have no influence on the predicted choice probabilities in those specifications that included a constant (with the exception of being aged 16-24 years in the 'No-Efficacy Difference' frame, but this result is likely to be statistical chance).

If we consider that the constant in each of these specifications represents the mean likelihood of making a choice one way or the other, and that such a likelihood will be

determined by intrinsic characteristics like age, education and income, it is reasonable to re-specify these models without a constant term. Note that this improves the model fit as shown by the R^2 values. From the model results, we observe that three demographic characteristics, *Age*, *Income* and *Arthritis* status, appear to influence the underlying probability of choosing a treatment option (see Figure 33). For all three variables, the magnitude of the coefficients does not differ across the framing sets (taking into account the resulting standard errors), and the signs of the coefficients are consistent. For *Age* we observe that the choice probability falls for the youngest age group compared to the oldest age group ('Base' and 'No-Efficacy Difference' only), but increases for the other two age groups. Interestingly, across all three frames, having arthritis significantly reduced the probability of choice for any given option. *Income* effects were less consistent; having a middle income ($40,000 to $79,999 per year) compared with low income increased the probability of choosing a given option for the 'Attribute' and 'No-Efficacy Difference' frames; but being in the highest income group in the 'Base' and 'Attribute' frames reduced the probability of choosing a given option.

In the earlier comparison of the demographic data presented in Table 21 we observed that the framing sets differed in terms of the composition of respondents with respect to *Age* (Base had proportionally more respondents over 65 years) and *Arthritis* (there were fewer respondents with arthritis in the 'No-Efficacy Difference' frame), but not *Income.* The differences in the demographic spread, together with the results presented in Table 33 might suggest that the analysis of choice should have been adjusted for the demographics of age and arthritis status, beyond the heterogeneity adjustments taken into account in the GMNL-I and mixed logit regression analyses. However, the direction of effect for *Age* and *Arthritis: No Arthritis* on the probability of choice is the same across all three frames. Moreover, these show the effect on the probability of choosing any given option within a choice. This experiment was an unlabelled forced choice, thus respondents must always have made a choice, and that choice was not specific to a labelled option. Since by construction both unlabelled options occurred with equal frequency throughout the design, any differences in demographics would

not influence the number of options chosen. No further adjustments for the demographics are therefore required within the context of this analysis.

Table 33: Influence of demographics on predicted probabilities

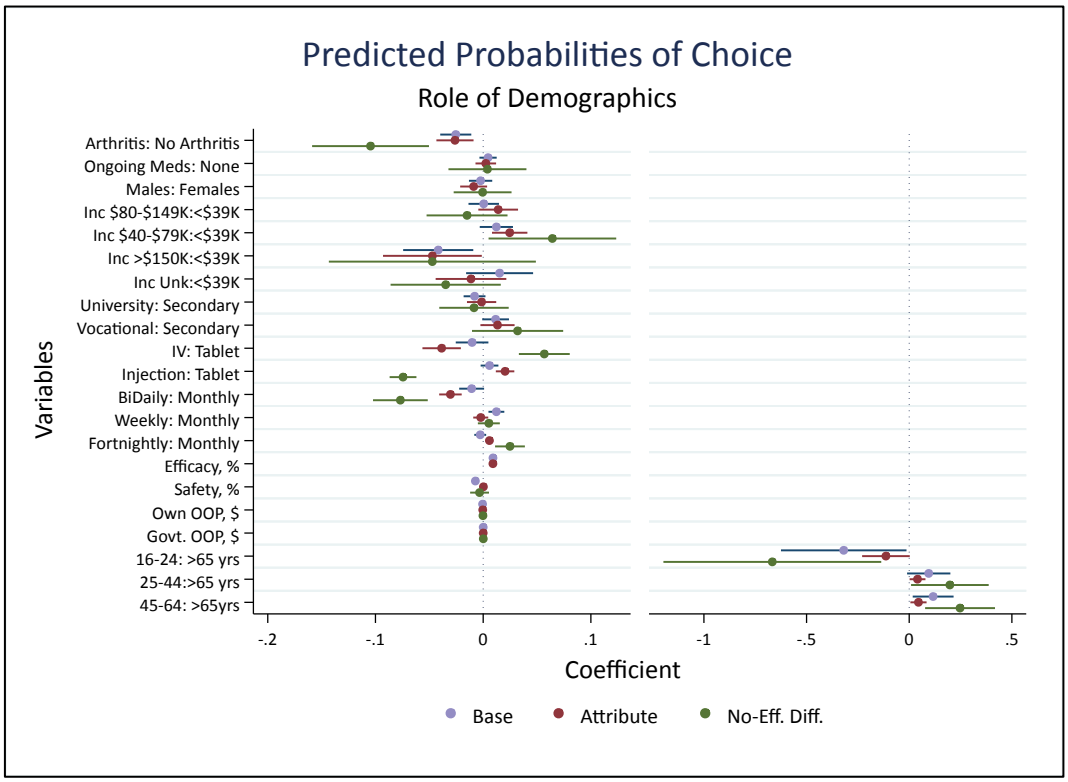
	Base		Attribute		No-Efficacy Difference	
16-24:>*65 yrs*	-0.007	-0.319	0.001	-0.113	-0.004	-0.667
	(0.005)	(0.155)*	(0.010)	(0.059)	(0.001)**	(0.268)*
25-44:>*65 yrs*	-0.001	0.095	-0.000	0.040	0.001	0.198
	(0.004)	(0.053)	(0.004)	(0.019)*	(0.001)	(0.096)*
45-64:>*65 yrs*	0.005	0.117	-0.006	0.045	0.002	0.248
	(0.004)	(0.050)*	(0.004)	(0.020)*	(0.001)	(0.086)**
Arthritis: *No Arthritis*	-0.001	-0.025	0.000	-0.026	0.000	-0.105
	(0.003)	(0.007)**	(0.003)	(0.009)**	(0.001)	(0.027)**
Ongoing Meds: *None*	-0.002	0.005	0.003	0.002	-0.000	0.004
	(0.002)	(0.004)	(0.002)	(0.005)	(0.001)	(0.018)
Males: *Females*	-0.002	-0.002	-0.002	-0.009	-0.000	-0.001
	(0.002)	(0.005)	(0.002)	(0.006)	(0.001)	(0.014)
Inc $80-$149K: <*$39K*	0.000	0.001	0.003	0.014	0.001	-0.015
	(0.005)	(0.007)	(0.004)	(0.009)	(0.001)	(0.019)
Inc $40-$79K: <*$39K*	0.000	0.012	0.003	0.025	-0.000	0.064
	(0.004)	(0.008)	(0.004)	(0.008)**	(0.001)	(0.030)*
Inc >$150K: <*$39K*	0.002	-0.042	-0.008	-0.047	-0.000	-0.047
	(0.007)	(0.016)*	(0.005)	(0.023)*	(0.001)	(0.049)
Inc Unk: <*$39K*	0.006	0.015	0.001	-0.011	0.001	-0.035
	(0.005)	(0.016)	(0.004)	(0.017)	(0.001)	(0.026)
University: *Secondary*	-0.004	-0.008	0.002	-0.001	0.001	-0.009
	(0.003)	(0.005)	(0.003)	(0.007)	(0.001)	(0.016)
Vocational: *Secondary*	0.001	0.012	-0.000	0.013	0.001	0.032
	(0.004)	(0.006)	(0.003)	(0.008)	(0.001)	(0.021)
IV: *Tablet*	-0.010	-0.010	-0.039	-0.039	-0.004	0.057
	(0.006)	(0.008)	(0.007)**	(0.009)**	(0.004)	(0.012)**
Injection: *Tablet*	0.011	0.006	0.028	0.020	-0.052	-0.074
	(0.004)**	(0.004)	(0.004)**	(0.004)**	(0.004)**	(0.006)**
BD: *Monthly*	0.013	-0.011	-0.001	-0.030	-0.011	-0.077
	(0.006)*	(0.006)	(0.008)	(0.005)**	(0.007)	(0.013)**
Weekly: *Monthly*	-0.000	0.012	-0.018	-0.002	-0.016	0.005
	(0.004)	(0.004)**	(0.004)**	(0.004)	(0.003)**	(0.005)
Fortnightly: *Monthly*	-0.012	-0.003	-0.003	0.006	-0.007	0.025
	(0.003)**	(0.003)	(0.002)	(0.002)**	(0.002)**	(0.007)**
Efficacy	0.006	0.009	0.006	0.009	n.a.	n.a.
	(0.000)**	(0.001)**	(0.000)**	(0.000)**	n.a.	n.a.
Safety	-0.015	-0.007	-0.009	0.000	-0.027	-0.003
	(0.001)**	(0.001)**	(0.001)**	(0.001)	(0.001)**	(0.004)
Own OOP	-0.001	-0.001	-0.001	-0.000	-0.001	-0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)*
Govt. Cost	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Constant	0.391		0.379		0.713	
	(0.008)**		(0.008)**		(0.006)**	
Observations	3,384		3,336		3,384	
n	141		139		141	

	Base		Attribute		No-Efficacy Difference	
F-Statistic	27,710.70	23,117.27	28,152.61	27,306.31	47,794.98	23,654.27
d.f.	21	21	21	21	20	20
p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
R-squared	0.53	0.86	0.51	0.88	0.55	0.74

Note: All coefficients are from an OLS regression.
Base categories are shown in italics.
All analyses were conducted in STATA 12, with robust standard errors (in parentheses) to account for the survey nature of the data.
Analyses conducted using 1,000 replications.

Abbreviation: d.f., degrees of freedom; BD, twice daily; CI, confidence interval; Inc, income; IV, intravenous; K, thousand; n.a., not applicable; OOP, out-of-pocket; Unk, unknown; yrs, years.

Figure 33: OLS regression output – demographics and predicted probabilities



Note: All coefficients are from an OLS regression.

Abbreviation: d.f., degrees of freedom; BD, twice daily; CI, confidence interval; Inc, income; IV, intravenous; K, thousand; OOP, out-of-pocket; Unk, unknown; yrs, years.

5.4 Discussion

Tofacitinib, the first orally administered bDMARD, was recommended for public subsidy by the PBAC in March 2015.[276] The sponsor of tofactinib lodged an application for reimbursement on a cost-minimisation basis compared with adalimumab, a subcutaneously injected treatment. The PBAC recommended listing on a cost-minimisation basis, but at a lower price for tofacitinib, relative to that of adalimumab, to remove an offset for the administration of adalimumab and to include costs expected to be incurred due to the hypercholesterolaemia associated with tofacitinib use.[276]

The results from the DCE analysis presented in this chapter suggest that patients might have been willing to pay a premium for tofacitinib over adalimumab for the convenience associated with having access to a tablet as opposed to a subcutaneous injection. However, this was demonstrated only in the framing set in which respondents were told the two treatment options did not differ in efficacy, somewhat mirroring the cost-minimisation situation under which tofacitinib (orally administered) was recommended to be non-inferior to adalimumab (subcutaneous injection). For the 'No-Efficacy Difference' frame, participants indicated their willingness to pay just under $39 to shift from a subcutaneous injection to an oral. However, using that same set of preferences, any difference in safety outcomes that might result in patients stopping treatment would reduce that WTP by approximately $23 for each one-percentage point higher risk of treatment cessation. Based on these results, a two-percentage point difference in the risk of treatment cessation, as was observed in the phase III trial comparing tofactinib with adalimumab[277], would absorb any increase in price based on the mWTP for the convenience gain.

On the basis of the preferences demonstrated in this chapter, only a shift from an IV administered therapy, such as abatacept, to a tablet such as tofacitinib, and a reduced frequency of administration, would result in a gain for which patients were prepared

to pay for the added convenience.[xxix] However, within a system such as the PBS, payments by patients are largely capped at the standard patient contributions; currently $38.30 for general patients, $6.20 for concessional patients (those on a Commonwealth health care or pension card), with reductions to $6.20 and no co-payment respectively once annual co-payment thresholds have been reached. Thus, any premiums for new drugs on the basis of a convenience gain would be paid by the Government in the form of higher prices.

This raises the question of whether public expenditure should be used to fund what are essentially gains to the individual and for which the individual has demonstrated a WTP. The research in this chapter has not addressed the question of whether or not governments should pay for gains in meta-health effects; it has asked how the amount that governments pay for treatment influences choice. The finding is that the amount that government pays, by and large, does not influence individuals' choices; individual choice is more likely to be influenced by their own OOP. While this could be interpreted to mean that individuals do not see a role for government in paying for meta-health effects or RA treatments in general, perhaps it shows that in this context individual decisions are more highly influenced by their own budgetary constraints rather than those affecting society at large. In this research, the government cost attribute was expressed as an out-of-pocket cost to government. Thus, it appealed to individuals' sense of altruism in a limited way because it did not link the costs to government back to the respondent through some payment vehicle, such as a tax contribution.[278] It is possible that linking government costs directly to individuals would have produced a different result.

In a recent review of the use of DCEs in health economics, Clark et al. (2014) note that of 179 studies reviewed, 81 (45%) investigated trade-offs between patient or consumer experience factors and health outcomes.[44] Investigating the trade-off between

[xxix] In this case the two work in opposition since the improved mode of administration – tablet – is associated with more frequent administration

convenience and health effects in this research is not in itself novel, and indeed previous stated preference research in the field of RA has included attributes that focused on convenience factors.[90,254,255] All three studies included comparable attributes as in the current analysis to examine preferences for RA treatment, although Augustovski et al. (2013) and Louder et al. (2016) focused on preferences among RA patients already taking disease modifying therapies.[90,254,255] In contrast, Harrison et al. (2015)[254] used a general community sample to produce QALY weightings for RA preferences. Augustovski et al. (2013)[90] and Harrison et al. (2015)[254] find that health effects dominate meta-health effects, but the ordering of those effects differs. In their analysis, Augustovski et al. (2013)[90] found that the order of factors influencing choice (from most to least influential) was: treatment cost, systemic AEs, frequency of administration, efficacy, mode of administration, local AEs and serious infections. Harrison et al. (2015)[254] found the order of influence to be: the extent of benefit, confidence in benefits, presence of side effects (minor and serious), mode of administration and frequency of administration.[254] In contrast, Louder et al. (2016)[255] found that meta-health effects dominate in determining choices as follows: mode and frequency of administration, side effects, costs, administration with other drugs, and efficacy.

There are some important differences between these studies. First, Harrison et al. (2015)[254] include a survival attribute, no cost attribute, and two attributes addressing efficacy (the magnitude of effect and confidence). The difference in the attribute used as the basis for subsequent estimation of marginal-rates-of-substitution complicates direct comparisons of observed trade-offs, while it is possible the two efficacy attributes might have interacted within that study to magnify the importance of health effects.[254] In addition, it is possible that individuals value meta-health effects less in the presence of a survival attribute trade-off than a cost attribute. This suggests that, for a given meta-health effect, lower values would result from a survival based utility valuation than using WTP.

Second, Louder et al. (2016)[255] present three attributes that contribute to meta-health effects (mode and frequency of administration, and whether treatment is in combination with other medications). The descriptions they include for IV and subcutaneous medications are also detailed, as in the 'Attribute' version used in this research, which suggests potentially greater consistency across respondents in what they evaluated than might have occurred with simple labels.[255] However, the levels for the efficacy attributes were relatively low and had a relatively narrow range compared with that used in this research.

Finally, Augustovski et al. (2013)[90] used an existing pool of RA patients, while the research presented in this chapter was based on a general community sample. Differences in health state valuations between patients and general community samples are well documented.[279,280] Given the differences in attribute construction and framing, it is not possible to further add to this debate by comparing the preferences observed by Augustovski et al. (2013)[90] with those reported in this chapter. However, it has been suggested that in the context of seeking to explore the value of convenience (a meta-health effect) for use in public decision-making, the public's preferences over those effects, as measured herein, are most relevant.[279] Ensuring that the public is suitably informed so as to be able to provide those preferences gives rise to the second difference between this work and that of Augustovski et al. (2013)[90]: investigating how the valuation of convenience relative to health effects varies with the information provided to individuals. This is a specific contribution to the field in understanding preferences for RA treatments, and more broadly with respect to understanding the impact of framing on preference elicitation for meta-health effects.

Framing effects have been investigated elsewhere using DCEs to consider the impact on choice of whether attributes are framed as gains or losses[258,281,282], to examine the influence of the units in which attribute levels are expressed[258], and to test the influence of including more information about a treatment[271] or attribute.[18] Like the current research, these studies also find that framing matters – individuals are influenced by

the type and manner in which information is presented. The presence of framing effects was most strongly supported by comparisons of the choice coefficients and corresponding mWTP across the survey versions. Comparisons of the scale coefficients did not support framing where health effects were coded as linear, but did support framing (with differences across the scale coefficients in the survey versions, see Table A 23) where health effects as well as meta-health effects were coded as categorical.

Within this chapter, the preferences of interest were those individuals expressed for convenience as compared with health effects. The manner in which attributes were described was varied in an attempt to influence how this was perceived by respondents and to test the influence of that manipulation on the resulting preferences for meta-health effects and health effects. This appealed to the notion that framing not only reflects traditional notions of risk (a win or a loss), but whether an individual will be positively or negatively predisposed to something of interest; such as an attribute.[89] Drawing on Levin et al.'s (1998)[89] taxonomy of framing, risk and attribute framing were tested respectively by removing the 'No-Efficacy Difference' version the likelihood of a gain or loss in efficacy associated with choosing one option over the other; and in the 'Attribute' frame by providing a more detailed account of the differences between how treatments are administered.

There are a few things to note about this method of comparing frames, and the subsequent results. The first is that in the absence of a difference in efficacy, respondents shift their focus to the meta-health effects; and in particular those related to frequency. It is within the 'No-Efficacy Difference' frame that we observe a significant effect for frequency in affecting choice and in determining value, relative to the health effects. This is what Kahneman (1984)[88] calls 'topical accounts' in which consequences of choice are dependent on the reference points within a specific context. The results observed are thus not surprising since in the absence of a difference in efficacy (no risk of a loss or gain in response), respondents focus on the other

attributes. The recommendation from these results is that incorporating information in this manner, as showing no difference, should be avoided as it biases the remaining preferences.

A similar recommendation for incorporating more rather than less information arises out of the results for the 'Attribute' frame. Here we observe higher estimated mWTP (point estimates) for the meta-health effects than in either of the other two frames, yet the values for the health effects are largely the same. In this case, providing respondents with more information about how each option is administered appears to have influenced their preparedness to trade-off between attributes. Within the 'Attribute' frame, when the levels indicated that there was more to lose associated with IV administration compared with the tablet there was a much higher mWTP; yet for the IV injection there was a positive mWTP indicating respondents would prefer the injection over the tablet. This might be because the additional explanations about treatment administration included both positive and negative information. That is, for the injection the description included the statement "*Sometimes you need help to do this so you go to see your doctor*", and the tablet included the statement "*Sometimes the arthritis in your hands makes it difficult to handle the tablets*". It is possible that respondents interpreted the statement in the tablet description as a loss relative to that in the injection description since for the injection they were receiving help (a gain), and there was no mention of specific difficulty. For the tablet, they were reminded (from the background health state) that the arthritis affects their hands, and that it makes small tasks difficult (a loss). If it is the case that respondents did interpret these levels in this manner, the relative value across all three levels is consistent with what we might expect from prospect theory.[80] That is, if we consider the tablet as the reference (or anchor) within the 'Attribute' set, the higher value for the IV reflects individuals perceiving this as a greater loss relative to that of a shift to the injection.

It is possible that the effect observed in the 'Attribute' frame is due to individuals 'filling in the blanks'; that is, in the absence of information, respondents formed a

negative view of the effects of the injection relative to the tablet. Once more information was provided the influence of that effect was moderated. Similar effects have been observed previously where describing to individuals how they were being treated produced lower values for a health state than when they were merely told they were being treated.[64] Similarly, the importance of the amount of information provided to the relative valuation of meta-health effects compared to health effects can be inferred by examining other DCEs that explore convenience with relatively simple attribute descriptions. For example, Muhlbacher and Bethge (2015)[283] find that frequency of administration of medication is the least influential attribute on choices for patients with acute coronary syndrome; yet the attribute description is short and evaluates small changes (all across the same day). Even within RA, the differences in the studies notwithstanding, Augustovski et al. (2014)[90] and Harrison et al. (2015)[254] find mode of administration to be one of the least influential factors; and both studies use a basic attribute description. Providing respondents with more information within the choice experiment itself potentially avoids differences between individuals in the reference points from which they trade-off between attributes.[xxx]

In addition to exploring the effects on value of how information is presented to respondents, this research has also investigated the impact on the value for meta-health effects of how data are coded. While previous studies have reported on the use of categorical coding for health effects,[283] the effect on the resulting values for meta-health effects of categorical coding compared with continuous coding for health effects is compared. The results show, that at the very least the assumption of linearity for health effects should be tested; where health effects are not linear – such as occurred in the 'Base' and 'No-Efficacy Difference' frames - assuming that they are linear results in an overestimate of the importance of meta-health effects at an attribute level, and a distortion of the resulting mWTP values. Johnson et al. (2011)[275] make recommendations regarding the importance of testing for the linearity of cost variables

[xxx] It is not clear that including additional information via 'hover' functions in online surveys would achieve the same ends since not all respondents would use that function.

in DCEs for the same reason; assuming linearity in effects might unduly influence the resulting mWTP values. Similarly, de Bekker Grob et al. (2013)[250] recommend that different functional forms (such as quadratic or logarithmic transformations) be applied to continuous variables to investigate potential non-linearities in utility functions in health care.

The results presented in this research suggest that the same applies with respect to analyses that include health effects and meta-health effects. Indeed, assuming linearity for health effects results in an overemphasis on the point estimates of the mWTP for meta-health effects relative to health effects, and to a distortion in the estimated mWTP for a shift between products with different health effect profiles. However, whether health effects should be linear or not appears to be influenced by the framing of the information, reinforcing the need for them to be tested within the context of each analysis.

5.4.1 Limitations

The attributes and levels used in the research reported in this chapter were developed based on similar DCEs for treatment in RA.[90,223] An alternative would have been to use qualitative research with RA patients to understand the condition from their perspective, and to reflect their views in the description of the subsequent attributes in the DCE. Indeed, Clark et al. (2014)[44] note the potential for attributes to be omitted or inappropriately specified if they are developed in the absence of qualitative research. However, they also argue that such qualitative research might not be warranted where the existing decision context is already established.[44] Arguably, this is the case in this instance insofar as treatment for RA is long established, and Augustovski et al. (2013)[90] had developed their attributes using qualitative research among RA clinicians and patients. Similarly, the significance observed in the mode of treatment in influencing choices has been attributed to convenience, but for some individuals it might also reflect needle aversion. The use of follow-up qualitative questions might be one avenue for revealing the source of such preferences in future research in this area.

A further limitation of this research is the sample size that applies within each of the framing sets. Approximately 140 respondents completed each of the three framing sets. While the analysis by Johnson et al. (2013)[47] suggest that the returns in terms of estimate precision begin to stabilise at 150, it cannot be ruled out that larger within frame sample sizes would have increased the precision of the estimates. In particular, it is possible that the differences noted between framing sets when using the categorically coded health effects suggest that larger sample sizes might have been desirable. This bears further consideration as the field of stated preference analysis continues to expand. In addition, the current analyses were of main effects only – yet it is possible that there were interactions between the levels of the mode and frequency attributes, and those of the cost attributes. While there were was an apparent dominance of mode over frequency and own OOP over government OOP, larger within frame sample sizes would permit such effects to be explored.

The question of sample size is also relevant when considering the representativeness of the survey sample; both for the general Australian population and between framing sets. It is noted that respondents to the survey tended to be in poorer health (more chronic health issues), more had RA, and were skewed toward the middle age groups compared with the Australian population. The results from the analysis of how demographic characteristics influence the probability of choice suggest that given the differences in age and arthritis status we might expect, on average, that individuals in the general community would have a higher probability of choosing a given RA treatment than observed in these data; it increases with the older age group, and is higher for those without arthritis. However, directly applying the preferences observed herein to predict increased choice probabilities is complicated by the forced-choice nature of this experiment (for the purposes of predicting uptake/demand the inclusion of an opt-out would be desirable).

5.4.2 Conclusion

The results in this chapter show that convenience factors influence treatment choice, but less so than health effects. This is consistent with what has been observed elsewhere for RA and indeed for other treatment decisions. More surprising was the finding that government costs did not influence choice; a different result might be observed if government payments were linked directly to the individual. The results also show that values elicited for meta-health effects in DCEs are subject to framing effects in the presentation of the attributes. The influence of meta-health effects (mode of administration and frequency, reflecting convenience), health effects (treatment efficacy and safety), and costs (own OOP and government costs) were all affected by the amount and type of information presented on attributes describing those effects and possible sources of difference between choice options.

The results reported in this chapter add to the literature on determining value for convenience by exploring the impact on those values of framing effects. Providing individuals with more information about treatment administration resulted in higher mWTP for convenience factors; both in absolute terms and relative to health effects. However, assuming linearity in health effects was not consistent with how respondents appeared to be making choices under conditions of more limited information. In those instances, higher values for health effects were produced by preserving the categorical nature of the health effect attributes within the analysis of the data. While it is not possible on the basis of one experiment alone to make proscriptions about how convenience should be valued, it would appear that providing as much information as practical on the attributes of interest would reduce the potential for bias. The findings for this research are important, both for the developers of similar studies and those who seek to use their outcomes, particularly as convenience is used more often as a source of differentiation when seeking funding for new health technologies.

6 Managing the Ongoing Risks of Breast Cancer Recurrence

Chapter Summary

Advancements in the understanding of breast cancer risk have resulted in an increase in the use of contralateral prophylactic mastectomy (CPM) for the management of ongoing cancer risk in women diagnosed with early stage breast cancer. The research in this chapter investigates the role of meta-health effects such as convenience, reassurance and autonomy in women's preferences for how to manage the ongoing risks of breast cancer recurrence.

A labelled DCE was used to investigate choices between CPM and routine monitoring for managing the ongoing risks of breast cancer recurrence. Based on feedback from focus group discussions with women with experience of breast cancer, the following attributes were included: mode and frequency of monitoring (convenience); risk of contralateral breast cancer; risk of ipsilateral or metastatic breast cancer; risk of pain or loss of breast sensitivity; involvement in decision-making (autonomy); and costs.

A WTP efficient design was generated in Ngene using 48 rows in 4 blocks. The DCE was administered to a community based on-line panel of women who answered 12 labelled choice sets in one of three survey versions. These differed in the amount and type of information presented. Respondents also rated the extent to which they were concerned about each attribute when completing the choice tasks. Results were analysed using conditional logit, mixed logit and latent class analysis, and reported as choice coefficients and mWTP. The value of reducing the fear of cancer recurrence was estimated by comparing the mWTP for a reduction in cancer recurrence between women who were concerned about cancer recurrence and those who were less concerned. Framing effects were tested by comparing coefficient estimates and mWTP values across the three survey versions.

There was a high degree of non-trading among the 464 women who completed the survey; over 50% of women always chose one option, typically routine monitoring. Results from both the latent class and mixed logit analyses demonstrate that women were less likely to choose an

option associated with higher risk. Meta-health effects influenced choice, but to a lesser extent than health effects. Women were more likely to choose options associated with less intrusive methods of monitoring (convenience) and where they were involved in decisions about their care (autonomy). Women who were concerned about cancer recurrence were more likely to choose CPM over routine monitoring. This resulted in a statistically significantly higher mWTP for a one point reduction in the risk of ipsilateral and metastatic cancer recurrence among those women compared with those who were less concerned about cancer recurrence. Framing effects were evident when comparing between survey versions, but were not present when assessed for the pooled dataset.

This is the first study to use a DCE to evaluate women's preferences for how to manage ongoing breast cancer risk. It demonstrates that beyond some women's strong preferences for one form of management over another, health effects influence choices to a greater extent than meta-health effects. Importantly, it was possible to estimate the value women place on achieving reductions in cancer risk and its potential reassurance.

6.1 Introduction

In Chapter 5, the role and value of convenience as a meta-health effect in the context of choosing a treatment for RA were explored. But convenience effects are not the only type of meta-health effects from which individuals will derive value in the experience of health care. For example, whether individuals have a role in deciding how and by whom they will be treated (autonomy), and the potential for care to reduce anxiety over the likelihood of future occurrences of ill health (reassurance) are also potential sources of value that might arise from the experience of health care. The objective of the research in this chapter is to explore whether these sources of value, along with convenience, are relevant to choices women make in the context of the ongoing management of breast cancer risk. The ongoing management of breast cancer risk was chosen as a topic because the complex nature of the choices women face, encompassing both health and meta-health effects, and the potentially irreversible nature of those

choices, make it particularly relevant for research investigating the importance of meta-health effects.

Recent advances in the understanding of breast cancer risk have resulted in an increase in the use of contralateral prophylactic mastectomy (CPM) for the management of ongoing cancer risk in women diagnosed with early stage breast cancer. The use of CPM is most often recommended for women who are positive for the breast cancer (BRCA) gene. However, its use in women who are not BRCA positive is increasing. This is understood to be due to the desire to reduce the risk of cancer; in part to overcome a fear of, or anxiety about, cancer recurrence.[284-289]

In a comprehensive study of the treatment decision made by 123 women who had undergone a CPM, Rosenberg et al. (2013)[285] found that factors associated with reducing the risk of disease in the other breast or achieving peace of mind were rated as extremely important in their decision-making by over 80% of women. In contrast, cosmetic effects were rated as extremely important to decision-making by only 28% of women.[285] Undergoing CPM due to the presence of the BRCA-1/2 gene or familial history of breast cancer has also been cited as an important consideration for women in reducing cancer risk and regaining control of their life.[290,291]

Factors relevant to decisions in the context of the ongoing management of breast cancer risk might also be revealed from decisions regarding primary treatment. In research with women deciding to undergo primary mastectomy, attitudes to keeping the breast, avoiding radiation (both because of its side effects and the time involved), doing everything possible ('fighting back') to beat the cancer, and minimising local recurrences, were all seen as important in forming a preference for mastectomy over lumpectomy.[292-294] Similarly, a desire to avoid ongoing surgeries associated with breast conserving treatment has been cited as a reason to choose mastectomy as the primary treatment.[295] These choices were found to be heavily influenced by the advice of the treating surgeon.[295] Post-treatment, women's attitudes to reconstructive surgery were

influenced by the extent and nature of the surgery involved, in particular whether it involved the use of prosthetic devices, and the impact on their appearance clothed.[296]

Use of CPM has been shown to be associated with increasing health care utilisation among women who have experienced breast cancer. In an analysis of 904 women with breast cancer in the US, average total health care costs post-surgery were 16.9% higher for the CPM group compared with the no-CPM group.[297] Given such increases in resource use there is a need to understand whether the outcomes produced, which include meta-health effects, justify the investment in terms of health care service use. Being able to enumerate all the effects of CPM – how it impacts on the fear of recurrence, changes in anxiety, convenience associated with different schedules for the ongoing monitoring of potential cancer recurrence – is therefore important. The enumeration of these effects in this chapter is obtained ultimately using a DCE to understand women's preferences for CPM plus routine monitoring compared with ongoing routine monitoring alone.

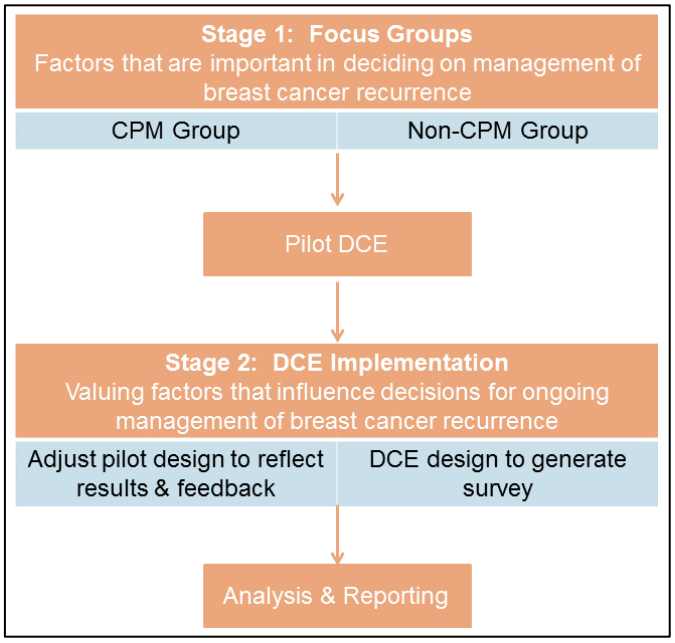
DCEs have been used previously to investigate a range of issues in the field of breast cancer. Caldon et al. (2008)[298] used a DCE to examine surgeons' preferences for mastectomy over breast conserving surgery in women with early breast cancer. Overall, mastectomy was the preferred option when trading between options that differed based on patients' age, breast cup size, tumour size, tumour location and tumour focality.[298] DCEs have also been used to investigate women's preferences for: methods of breast reconstruction[299]; participation in screening programmes[130]; types of follow-up monitoring programmes[224,225]; and, to understand the importance women with breast cancer place on different aspects of QoL.[300]

However, DCEs have not been used to investigate women's decisions regarding the management of ongoing breast cancer risk. This research set out to value explicitly how women trade-off between health effects (such as cancer risk) and the meta-health effects of convenience (such as the mode and frequency of breast cancer screening),

autonomy (involvement in decision-making) and reassurance (due to a reduction in anxiety or concern associated with cancer recurrence). The inclusion of a cost attribute allowed the mWTP for the meta-health effects and health effects to be estimated. Finally, the influence of framing effects was tested using the amount of information and health context presented to respondents within the DCE.

The research proceeded in two stages. In stage I, qualitative research was conducted among women recruited via the Breast Cancer Network of Australia (BCNA) to explore their choices regarding the decision to keep or remove the contralateral breast following a diagnosis of early stage breast cancer. The results of the qualitative research informed the development of the attributes and levels for the DCE in stage II. The DCE was implemented in two phases. In phase I, a pilot DCE was conducted among women from an online panel of members of the general Australian community. The results from that survey, together with quality assurance feedback from participants and from a subsequent seminar presentation of the results, were utilised to amend the DCE survey. In phase II, the DCE survey was implemented among women from the same online panel as the pilot survey. The outline for the research is depicted in Figure 34.

Figure 34: Approach to prophylactic mastectomy research



Abbreviations: CPM, contralateral prophylactic mastectomy; DCE, discrete choice experiment.

The results from the DCE were analysed to take account of individual heterogeneity in responses, using mixed logit regression and latent class analysis. Of particular interest in this research was the extent to which it was possible to place a value (expressed as the mWTP) on reducing the fear of recurrence in women with a diagnosis of early breast cancer. This was investigated using subgroup analyses based on women's attitudes to cancer risk to test for differences in choice behaviour and mWTP.

6.2 Methods

The approach to the research in this chapter was structured in two parts; an initial qualitative phase designed to underpin the development of the DCE, and a subsequent quantitative phase in which the DCE survey was implemented and analysed. The UTS Human Research Ethics Committee (HREC) approved both phases of this research; HREC 2014000423 for the qualitative research, and HREC 2015000161 for the

quantitative research.[xxxi] This section outlines the approaches taken in the conduct of both phases of the research.

6.2.1 Incorporating the patient perspective

There has been an increasing recognition of the importance of including consumers or patients in health services research; both to improve the translation of the research into practice, but also to enhance the relevance of the research question and approach.[301-303] With this in mind, women with experience of breast cancer, 'consumers', were an integral component of the development and conduct of this research. Consumer involvement and participation in this research was supported by the BCNA. BCNA was consulted on the overall research question and their support was sought to facilitate access to: (1) consumer representatives to participate in the development and interpretation of the research; and (2) women with experience of breast cancer to participate in the empirical research. As a result, two consumer representatives with experience of being treated for breast cancer were nominated from BCNA to assist with the development and conduct of this project. Prior to commencing the project, the consumer representatives were provided with training on the project scope and methods to be employed, and were consulted on an ongoing basis for their input on the design of the various elements of the project, as well as interpretation of the results. In addition, BCNA provided access to women in its Review & Survey Group[xxxii] to participate in the focus group research.

Discussions were initially held with BCNA regarding distribution of the DCE survey among the members of the Review & Survey Group. However, the BCNA management felt that the nature of the DCE survey – asking women to make

[xxxi] The application corresponding to the qualitative research was selected by the UTS HREC as an exemplar for this type of research in the discipline of economics and was subsequently made available on its website.

[xxxii] The Review & Survey Group comprises women with experience of breast cancer who have completed their treatment and have registered to be contacted for research regarding breast cancer.

hypothetical choices between CPM and routine monitoring – presented a risk that some women might feel undue regret or anxiety associated with the decisions they had already made. Accordingly, women in the Review & Survey Group were not invited to participate in the DCE. BCNA continued to support the research through the participation of the consumer representatives. The DCE was therefore conducted among members of the general community, for whom breast cancer is a risk and who are also potential tax-payers whose preferences regarding ongoing breast cancer management strategies are relevant in the context of public sector decision-making.

6.2.2 The qualitative research

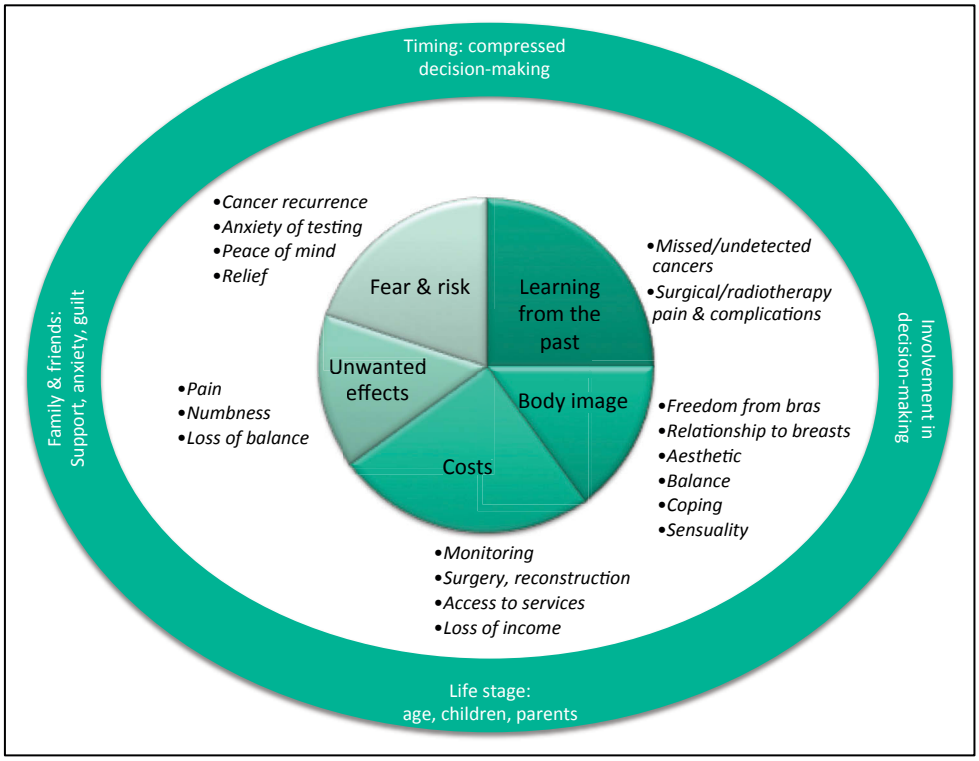
Women in the Review & Survey Group of BCNA were invited to participate in one of two focus groups depending on the choice they had made previously: women with early stage breast cancer who had decided to have a CPM, and those who were eligible but chose not to have a CPM. The methods and results of the focus group research are presented in Appendix 10. Eleven women, nine who had undergone a CPM and two who did not, participated. Analysis of the themes emerging from the group discussions revealed five key factors that were important to the decisions women made:

- their attitudes to fear of cancer recurrence, including anxiety associated with a fear of cancer recurrence and anticipatory anxiety associated with testing of the unaffected breast. Reducing anxiety or a fear of recurrence were motivators for undergoing a CPM;
- learning from the past, in terms of avoiding the pain and negative effects of surgery, motivated some women to avoid CPM. Negative associations of prior undetected or missed diagnoses of cancer that had diminished trust in breast monitoring, which for some women resulted in a desire to avoid ongoing monitoring and the intrusiveness of further testing, was associated with undergoing a CPM;

- costs, including those associated with ongoing monitoring in women once they had a diagnosis of breast cancer, surgical costs and loss of income, were noted as a factor by some women who underwent a CPM;
- unwanted effects of breast cancer treatment, including the pain associated with surgery, were associated with not having a CPM, while both groups noted the potential for numbness and loss of sensation in the breast area post mastectomy (primary or prophylactic); and
- body image (appearance) and identity, including sexuality, influenced decisions by women in the CPM group, in particular as to whether or not they underwent reconstruction or wore prostheses.

These five factors and relevant subthemes are summarised in Figure 35, including a number of minor themes focusing on communication (with doctors or family/friends), the stage of life at which women developed breast cancer and the compressed timelines for decision-making. This demonstrates the complexity of factors that women reported influenced their decisions regarding whether or not to undergo a CPM.

Figure 35: Factors influencing decisions to undergo CPM



Abbreviation: CPM, contralateral prophylactic mastectomy.

6.2.3 The quantitative research

This section outlines: the development of the DCE survey based on the results from the qualitative research; the underlying modelling approach; data collection; and data analysis. As described in Chapter 1, the investigation of framing effects was a sub-theme of this research. Nonetheless, the approach to exploring those effects is described before that of the remaining quantitative research as it provides definitions that appear in subsequent sections.

6.2.3.1 Exploring framing

The investigation of framing effects was a secondary aim of this research. The influence of framing reflects how individuals combine the effects of how they conceptualise a problem, how that problem is presented to them and the information

they are provided.[83,84] Those effects were investigated in this research through the use of three framing sets (versions) of the DCE survey:

1. 'Base': the 'base' design in which there was a common approach across attributes in terms of the amount of information presented in describing the attributes and their levels.
2. 'Goal': this utilises the same attributes and levels as the 'Base' framing set, but alters the background information provided to the respondents regarding the history and process of the breast cancer diagnosis.
3. 'No-Efficacy Difference': scenarios were restricted such that they did not differ in terms of the impact of the two options presented on the risk of breast cancer outside of the contralateral breast (the levels were the same between treatment alternatives within a choice set) and were described as having the same effect between treatments.

The justification for the 'No-Efficacy Difference' frame is the same as that applied to the RA Therapy study; there have been studies investigating the value of meta-health effects that focused on that value without varying the potential efficacy effects.[63,64,113] The 'Goal' frame is designed to test how the decisions women make might be influenced by the experience of care prior to the diagnosis of breast cancer. This is to emulate situations described during the focus groups (see Appendix 10) in which women felt their cancer was 'missed' by an earlier mammogram.

It was hypothesised that providing more information about the background of the breast cancer diagnosis in the 'Goal' framing set would alter the acceptability among some women for ongoing monitoring, resulting in a higher preference for CPM. This would be expected to reduce the acceptability of more frequent, or 'invasive' monitoring (e.g. MRI compared with ultrasound). Within the 'No-Efficacy Difference' frame, it was hypothesised that not including efficacy differences in the DCE would result in higher values for meta-health effects relative to the health effects. Testing of framing effects was investigated largely in the same manner as presented

previously[258,259]; by comparing the coefficient estimates, overall model performance and resulting mWTP values from differently framed versions of the same DCE. The difference is that the modelling approach used in the Mastectomy study (latent class analysis, see Chapter 4 and Section 6.2.6.2) did not produce an estimate for scale; thus differences in scale were not used as the basis for assessing framing effects in the research presented in this chapter.

6.2.3.2 DCE attributes and health vignettes

6.2.3.2.1 The DCE pilot

The development of the attributes for the pilot study is summarised in Table 34, and used the following process:

1. The domains for use in forming the attributes were drawn from the themes identified in the qualitative research (see Figure 35).
2. Each of the domains arising out of the qualitative research was deconstructed into those aspects that can be described as pertaining to a process aspect of care and those which reflect a response to health care (e.g. feelings of anxiety).
3. Process aspects of care were considered eligible for inclusion as attributes in the DCE given that they are potentially amenable to change via policy.
4. Domain elements that reflected a response to health care e.g. anxiety, were not considered amenable to inclusion as an attribute within the DCE since they reflect how a woman might react or feel.
5. All proposed attributes were subsequently rated as being of high, medium or low importance in terms of what is likely to most influence the choice of risk management option. The ratings were based on the feedback from the qualitative interviews.
6. Levels for the attributes were determined based on the published literature, clinical input and the results from the qualitative research.

Table 34: Mapping of qualitative domains to suggested DCE attributes

Domain from qualitative work.	Categorisation (Process or Response).	Aspect potentially captured as an Attribute.	Relevance of possible attribute to primary DCE question.
Fear & Risk			
Cancer recurrence	Process & Response	Recurrence rate associated with the proposed method of managing risk.	High. This is important as it relates to the "efficacy" and recurrence aspects being tested.
Anxiety of testing (anticipatory anxiety)	Response	Frequency of ongoing testing. Type of ongoing testing.	High. These are critical aspects of describing the ongoing monitoring.
Peace of mind	Response	Not suitable as an attribute; could be used to form underlying choice question.	
Relief	Response	Not suitable as an attribute; could be used to form underlying choice question.	
Learning from the Past			
Missed & undetected cancers	Process	False negatives/false positives of testing. Potentially as part of background scenario only.	Medium. Will it overly complicate the DCE to have recurrence and accuracy?
Surgical & radiotherapy pain/complications	Process Response	Side effects associated with TM and radiotherapy interventions. Potentially as part of background scenario only since these complications relate to primary therapy.	Low. Describe as part of the background scenario only. Response to pain etc cannot be affected by policy.
Costs			
Monitoring	Process	Costs of monitoring.	High.
Surgery/ reconstruction	Process	Costs of surgery/ reconstruction; as a separate attribute.	High.
Access to services	Process	Need to travel for care. Need to see another team for care. Time to wait for care.	Low. Raised in limited number of cases by women in the qualitative work.
Loss of income	Process	Time to return to usual duties.	Low. There is more than just the physical at play here, so it is difficult to state when women would be able to return to usual duties.
Unwanted Effects			
Pain	Process	Side effects of intervention.	Medium. Potentially incorporate with other side effects.
Numbness	Process	Side effects of intervention.	Medium. Potentially incorporate with other side effects.
Loss of balance	Process	Side effects of intervention.	Medium. Potentially incorporate with other side effects.

Domain from qualitative work.	Categorisation (Process or Response).	Aspect potentially captured as an Attribute.	Relevance of possible attribute to primary DCE question.
Body Image			
Freedom from bras	Response	Not suitable as an attribute; potentially include as an attitude question within the survey?	
Relationship to breasts	Response	Not suitable as an attribute; potentially include as an attitude question within the survey?	
Aesthetics	Process & Response	Potentially captured via prosthetics (see below).	Medium.
Breast symmetry	Process	Could be captured using an attribute for prosthetics.	Medium.
Coping	Response	Not suitable as an attribute; could form part of the choice question.	
Sensuality	Response	Not suitable as an attribute.	
Timing	Response & Process	Not suitable as an attribute.	
Life stage (age, children, family)	Response	Collect demographic information. Not suitable as an attribute.	
Medical involvement (communication)	Process	Vary degree of involvement in ongoing risk management.	Medium.
Family communication (anxiety, guilt)	Response	Not suitable for an attribute. Collect demographic information.	

Abbreviation: DCE, discrete choice experiment; TM, therapeutic mastectomy,

While timing, and its impact on decision-making, was an important factor raised by women in the focus groups, an attribute for this domain was not included due to the difficulty of capturing all the aspects of timing in a meaningful attribute. That is, there is a difference between a woman choosing to delay her decision about undergoing a CPM, and deciding to undergo the CPM but needing to delay the surgery for medical reasons (such as comorbidities and the physical and emotional enormity of the surgery[304], particularly if it includes reconstruction). Choosing to delay is not amenable to inclusion as an attribute in a DCE; an attribute such as 'You decide to delay your decision' is prescriptive and describes an action the woman would take rather than an attribute of the approach to risk management. Similarly, delaying for medical reasons relies on a medical assessment about the woman's ability to tolerate surgery and therefore is not specifically an attribute of the approach to risk management. Moreover, it is difficult to specify how long a delay might be and what

happens to women (in terms of monitoring) during such a delay. For these reasons, the impact of timing on the choice to undergo a CPM was not tested as part of this DCE.

The attributes and levels included in the pilot study are presented in Table 35. While anticipatory anxiety has been classified as a response factor, it is mediated by process factors (the type of monitoring and how often it is required). Hence, its influence on choice is included indirectly through the monitoring attribute. Similarly, how women respond to an actual cancer recurrence is a response that cannot be affected by policy, but there is the potential for more accurate methods/monitoring intervals to be introduced that influence recurrence rates (making it a process mediated attribute), or the provision of education/counselling to address potential concerns associated with fear of recurrence, and therefore it was included.

Table 35: Attributes and levels – pilot study

Domain	Attribute		Levels	Restrictions
Cancer recurrence	CBC Risk	The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:	3 in a 1,000 (0.3%) 10 in a 1,000 (1%) 20 in a 1,000 (2%) 50 in a 1,000 (5%)	Levels for CPM cannot exceed those for Routine monitoring.
	Other BC Risk	The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:	100 in a 1,000 (10%) 150 in a 1,000 (15%) 200 in a 1,000 (20%)	Appear the same in 'No-Efficacy Difference' frame (see below).
Monitoring	Monitoring	In addition to your regular self-checks, you will need to have the following tests:	Mammogram MRIs of your breast area Ultrasounds of your breast area	Mammogram does not appear in CPM
		You are scheduled to have your follow-up tests every:	Six months Year Second year	
Costs	Monitoring OOP	The cost for monitoring each year is $900, and you pay:	0, 300, 600, 900	
	Surgery OOP	The cost for surgery associated with managing your ongoing risk of cancer recurrence is $15,000, and you pay:	0, 5000, 10000, 15000	Applies to CPM only.
Unwanted effects	Pain Risk	The chance you will experience ongoing pain is:	400 in 1,000 (40%) 300 in 1,000 (30%)	Applies to CPM only

Domain	Attribute		Levels	Restrictions
			200 in 1,000 (20%) 100 in 1,000 (10%)	
	Sensitivity Risk	The chance you will experience an ongoing loss of sensitivity in your breast area is:	600 in 1,000 (60%) 500 in 1,000 (50%) 400 in 1,000 (40%) 300 in 1,000 (30%)	Applies to CPM only
Breast symmetry	Symmetry	Following your surgery:	You are able to have breast reconstruction. You are able to wear external breast prostheses. You are unable to wear prostheses or have a reconstruction.	
Involvement in decision-making (communication)	Involved	Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence.	Always Not very often	

Abbreviation: BC, breast cancer; CBC, contralateral breast cancer; CPM, contralateral prophylactic mastectomy; MRI, magnetic resonance imaging; OOP, out-of-pocket.

Health Effect Attributes

Cancer recurrence was included as the principal indicator of efficacy and risk. Breast cancer recurrence typically refers to the re-emergence of cancer in the ipsilateral (or originally affected) breast.[286] However, a woman who has had a primary breast cancer might also experience the emergence of a new primary cancer in the contralateral breast (either synchronous to the primary cancer or metachronous), or metastatic disease that emerges distant from the original primary tumour.[286,287] Accordingly, breast cancer risk was captured via two attributes; one reflecting the occurrence of contralateral breast cancer (CBC, *CBC Risk*) and the other, metastatic and ipsilateral breast cancer recurrence (*Other BC Risk*).

While it is possible that presenting separate risk attributes might complicate the choice scenario, data from systematic reviews indicate that the difference in recurrence associated with CPM is due to CBC, and not overall breast cancer recurrence.[287,288] In a cohort analysis of women with early breast cancer, the 10 year incidence of overall

breast cancer recurrence was 15.6% among those who underwent a CPM, and 18.5% for those who did not.[286] Presenting a recurrence risk for overall breast cancer is likely to show little variability.

The levels for the two attributes, *CBC Risk* and *Other BC Risk*, were drawn from the CBC and other BC incidence observed over a median of 17 year cohort of women in the US who had undergone CPM and, or therapeutic mastectomy (unilateral mastectomy; TM) alone.[286] Within that study the observed incidence of CBC among women who underwent CPM was 0.03% per year, with 1.53% experiencing a non-CBC recurrence or new primary.[286] The corresponding figures for women in the non-CPM group were 0.47% and 1.38% respectively.[286] In applying these risks as levels (see Table 35), *CBC Risk* for CPM was restricted to not exceed that of the routine monitoring only option. This restriction was suggested by BCNA and was to counter potential cognitive difficulties arising for respondents if they were presented with an option in which the risk of a contralateral BC was higher for a woman after having undergone a CPM than if not.

Breast cancer mortality depends on the treatment administered post-cancer recurrence; whether it was a non-CBC recurrence; and the baseline health of women.[305] A number of reviews have suggested that there is potentially no survival gain associated with CPM; while CPM is effective at preventing CBC, the bulk of breast cancer mortality is due to distant disease.[287,289,305] For these reasons, and given the complexity of trying to separate the potential cancer recurrence risk from the mortality risks, cancer mortality risk was excluded from the list of included attributes.

The qualitative research highlighted that the unwanted effects of pain in the breast area (*Pain Risk)* and loss of breast sensitivity (*Sensitivity Risk*) were important considerations for women in their decision-making. These attributes are somewhat more difficult to quantify for inclusion in the DCE given its focus on options for managing ongoing cancer recurrence. A number of studies have reported on the occurrence of side effects

such as pain and asymmetry in women undergoing TM for breast cancer, and were focused on women who underwent breast reconstruction.[306-308] Barton et al. (2005)[306] reported types of complications among women who did not undergo reconstruction, noting that 53% reported some form of adverse effect. Overall, pain occurred in as many as 36% of women, 17% experienced some form of infection, 17% seroma, 2% numbness in the breast, and 3% lymphoedema.[306] The incidence of infection and seroma in this case is likely to relate to the immediate post-surgical period. Osman et al. (2013)[309] showed that 30 day surgical complications occur in twice as many women undergoing CPM than TM, at 8.4% and 4.2% respectively.

However, the results from the qualitative research conducted for this study suggest that the longer-term post-surgical complications, such as pain and numbness (loss of sensitivity in the breasts) might be more influential factors in the choices women make about how best to manage ongoing cancer recurrence risk. The negative effect of CPM on breast sensitivity and sexuality has been reported as occurring in nearly 50% of women undergoing that procedure.[308,310] Effects on physical balance were also noted in the focus groups as being adverse effects of importance to some women. However, there is little in the literature regarding the occurrence of loss of balance. Moreover, the discussion during the focus groups conveyed that balance loss might depend on women's breast size, making it a difficult concept to convey in an attribute in a DCE. Accordingly, only attributes for pain and breast sensitivity were included in the DCE.

The levels used for *Pain Risk* and *Sensitivity Risk* are informed by the reported occurrence of these outcomes across a number of studies; an indicative range for pain has been chosen based on the values reported by Barton et al. (2005)[306]; and for breast sensitivity, the range includes the incidence of breast sensitivity loss as reported by Davies et al. (2015)[311] and Gahm et al. (2010)[308]. While women who undergo TM alone and routine monitoring will also be at risk of pain or loss of sensitivity in the ipsilateral breast area, this will not be associated with the choice of risk management option.

Therefore, the *Pain Risk* and *Sensitivity Risk* included in this DCE relate only to that associated with the removal of the contralateral breast.

Meta-health Effect Attributes

It is of interest to assess what impact the different levels of monitoring 'invasiveness', as a measure of convenience, has on willingness to accept monitoring. The levels used in *Monitoring* are based on the Cancer Australia Clinical Guideline for the follow-up of women completing treatment for early breast cancer.[312] These guidelines recommend at least six to 12 monthly medical examinations, with annual mammography (or ultrasound/MRI as required) for the duration of follow-up as directed by the clinician in consultation with the patient.[312] In practice, the type of monitoring women undergo (*Monitoring*: mammogram, ultrasound, MRI) for breast cancer will in part depend on the density of their breast tissue, and their underlying cancer risk. The levels for *Monitoring* do not depend on these factors in this DCE. Rather, a difference arises between the CPM and routine monitoring options in that mammography cannot be used for monitoring in women who have undergone a CPM as there is insufficient breast tissue remaining to perform a mammogram. Accordingly, mammography does not appear as a level for *Monitoring* for the CPM option.

The issue of breast symmetry arose from the qualitative research as a concept encompassing how women felt about the aesthetic and psychological impacts of having a TM alone or CPM. The inclusion of *Symmetry* as an attribute allows for the possibility that, regardless of the risk management option chosen, breast symmetry might be addressed through reconstructive surgery, the use of breast prostheses or neither. The process of breast reconstruction following TM or CPM is very complex.[287,289] The decision of whether or not to keep the unaffected breast has been linked to decisions regarding the extent and timing of breast reconstruction following mastectomy.[313-315] The wording of this attribute reflected this complexity; it was expressed as whether it was possible for women to wear prostheses or undergo reconstructive surgery.

Involvement in decision-making also arose as an important factor for some women in the qualitative research. An attribute has been included that allows for the possibility that the woman is always or rarely *Involved* in the decision-making process about her care.

Out-of-Pocket Costs

Within this DCE, separate attributes have been used for the cost components for monitoring (*Monitoring OOP*) and surgical (*Surgery OOP*) OOP costs, in a similar fashion to previous DCEs.[21,244]

In 2010, the BCNA conducted a survey of 328 women who had undergone breast reconstruction surgery (including some who did not have breast cancer), and asked for information on OOP costs associated with private hospital services.[316] A total of 271 respondents had their procedures in a private hospital, of whom 270 provided information on OOP costs. The median category for OOP costs among these women (allowing for Medicare and private health insurance rebates) was $3,000-$4,000. The majority of women (89%) paid less than $10,000 in OOP costs, with 3.3% paying more than $15,000. Based on this distribution of OOP costs it seems appropriate that the levels for *Surgery OOP* be: $0, $5,000, $10,000, and $15,000. Arguably, women choosing either option might face these surgical costs since women who undergo TM might face some costs associated with their procedure. However, since this research is focused only on choices associated with the ongoing management of breast cancer recurrence risk, the application of these costs was limited to the CPM option.

The maximum level for *Monitoring OOP* was formed by estimating the cost of a woman having two visits with a specialist in a given year, plus an MRI. The cost of an MRI has been included since this represents the highest per unit cost for any of the monitoring techniques. Two specialist visits per year are estimated to cost $171.10 (MBS Item 105,

Fee $85.55), and an MRI of both breasts is $690 (MBS 63467[xxxiii]). This produces an estimated cost of $861.10, or rounded to $900 per year in the event of an MRI. To simplify the presentation of *Monitoring OOP*, and not make it conditional on the monitoring attribute, this was permitted to vary between 0 and $900.

The Health Vignettes

Each respondent to the survey saw a health vignette prior to completing the choice tasks. The health vignettes were designed to provide respondents with a minimum level of information regarding the diagnosis of breast cancer, its treatment, and an outline of the options for ongoing management of breast cancer recurrence. The language used in the vignette was designed to be as neutral as possible in describing the condition and its treatment.

In order to investigate 'Goal' framing, two versions of the health vignette were developed; one in which more information was provided on the history of the breast cancer diagnosis and the process leading up to the decision of how to manage the ongoing breast cancer recurrence risk. Women randomised to either the 'Base' frame or the 'No-Efficacy Difference' frame saw the 'Base' frame Health Vignette (Text Box 3), while those randomised to the 'Goal' frame saw the 'Goal' frame Health Vignette (Text Box 4).

[xxxiii] MRI is not funded on the MBS for use in managing breast cancer recurrence, so this fee item is used as an indicative cost only.

Text Box 3: 'Base' frame health vignette

Managing the ongoing risks of breast cancer
Your doctor has told you that you have cancer in one of your breasts. She explains that the treatment for this type of cancer involves having surgery to remove the affected breast. After you have had surgery you might also have chemotherapy and radiotherapy.

Your doctor explains that, despite those treatments, there is still a chance the cancer might come back or that a new cancer might develop. This might happen around the area where your breast was removed, or somewhere else in your body, including your other breast.

Your doctor explains that there are some things you can do to help manage your ongoing risks of breast cancer. Two of these include having surgery to remove your currently unaffected breast plus monitoring, or keeping that breast and opting for monitoring only. She explains that these options differ in a number of ways:

- their impact on the chances of the cancer coming back;
- the types of monitoring you undergo and how often;
- the chance that you will experience pain or loss of sensation in your breast area;
- how much they cost you;
- whether you are part of the decisions about your care; and
- the cosmetic options you have available after surgery, including having reconstructive surgery or wearing artificial breasts called prostheses.

Your doctor would like you to think about and choose how you would like to manage your ongoing risks of breast cancer. Your choices are to have "Surgery to remove unaffected breast plus monitoring" or "Routine monitoring only".

You will now see 12 questions asking you to choose between these two options for managing your ongoing risks of breast cancer. Each time, please indicate which option you would prefer.

Text Box 4: 'Goal' frame health vignette

Managing the ongoing risks of breast cancer
You have noticed some changes in the look and feel of one of your breasts. Even though you recently had a mammogram that was clear you decide to see your doctor to discuss these changes. She sends you for some additional tests.

You have another mammogram and it shows a small lump in your breast. Your doctor sends you to have a biopsy where a small piece of tissue is taken from the lump in your breast using a needle. When you go back to see your doctor the following week she tells you that you have cancer in your breast.

Your doctor explains that the treatment for this type of cancer involves having surgery to remove the affected breast. After you have had surgery you might also have chemotherapy and radiotherapy. She discusses with you that despite those treatments there is still a chance the cancer might come back, or that a new cancer might develop. This might happen around the area where your breast was removed, or somewhere else in your body. This might include developing a cancer in your other breast.

Your doctor explains that there are some things you can do to help manage your ongoing risks of breast cancer. Two of these include having surgery to remove your currently unaffected breast plus monitoring, or keeping that breast and opting for monitoring only. She explains that these options differ in a number of ways:

- their impact on the chances of the cancer coming back or a new cancer developing;
- the types of monitoring you undergo and how often;
- the chance that you will experience pain or loss of sensation in your breast area;
- how much they cost you;
- whether you are part of the decisions about your care; and
- the cosmetic options you have available after surgery, including reconstructive surgery or wearing artificial breasts called prostheses.

Your doctor would like you to think about and choose how you would like to manage your ongoing risks of breast cancer. Your choices are to have "Surgery to remove unaffected breast plus monitoring" or "Routine monitoring only".

You will now see 12 questions asking you to choose between these two options for managing your ongoing risks of breast cancer. Each time, please indicate which option you would prefer.

6.2.3.2.2 The final DCE study

Analysis of the results from the pilot survey showed that there were two factors that were not significant: *Pain Risk* and *Symmetry*. However, women ranked *Pain Risk* as the fourth most important factor in their decision-making, and there were significant differences between women in the variance associated with *Pain Risk*. *Pain Risk* was therefore retained as an attribute. Feedback during a presentation of the pilot results (see Appendix 13) was that inclusion of *Symmetry* as an attribute was complicated in

that women might also have applied this attribute to the routine monitoring option, even though it was specific to CPM. Therefore, *Symmetry* was excluded as an attribute, but more information on breast reconstruction and prostheses was included as part of the background health vignettes. The resulting list of final attributes is provided in Table 36, with the health vignette for the 'Base' framing set in Text Box 5 (the same change was made to the health vignette for the 'Goal' framing set).

Table 36: Attributes and Levels – final study

Attribute		**Levels**	**Restrictions**
CBC Risk	The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:	3 in a 1,000 (0.3%) 10 in a 1,000 (1%) 20 in a 1,000 (2%) 50 in a 1,000 (5%)	Levels for CPM cannot exceed those for Routine monitoring.
Other BC Risk	The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:	100 in a 1,000 (10%) 150 in a 1,000 (15%) 200 in a 1,000 (20%)	Appear the same in No-Efficacy Difference framing set.
Monitoring	In addition to your regular self-checks, you will need to have the following tests:	Mammogram MRIs of your breast area Ultrasounds of your breast area	Mammogram does not appear in CPM
Frequency	You are scheduled to have your follow-up tests every:	Six months Year Second year	
Monitoring OOP	The cost for monitoring each year is $900, and you pay:	0, 300, 600, 900	
Surgery OOP	The cost for surgery associated with managing your ongoing risk of cancer recurrence is $15,000, and you pay:	0, 5000, 10000, 15000	Applies to CPM only.
Pain Risk	The chance you will experience ongoing pain is:	400 in 1,000 (40%) 300 in 1,000 (30%) 200 in 1,000 (20%) 100 in 1,000 (10%)	Applies to CPM only
Sensitivity Risk	The chance you will experience an ongoing loss of sensitivity in your breast area is:	600 in 1,000 (60%) 500 in 1,000 (50%) 400 in 1,000 (40%) 300 in 1,000 (30%)	Applies to CPM only
Involved	Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence.	Always Not very often	

Abbreviation: BC, breast cancer; CBC, contralateral breast cancer; CPM, contralateral prophylactic mastectomy; MRI, magnetic resonance imaging; OOP, out-of-pocket.

Text Box 5: 'Base' frame Health Vignette - Final

Managing the ongoing risks of breast cancer
Your doctor has told you that you have cancer in one of your breasts. She explains that the treatment for this type of cancer involves having surgery to remove the affected breast. After you have had surgery you might also have chemotherapy and radiotherapy. She discusses with you that there are options available to you after your treatment if you want to change how the surgery has made your body look. These options might include having reconstructive breast surgery or wearing artificial breasts called prostheses.

Your doctor explains that, despite the treatments you have, there is still a chance the cancer might come back or that a new cancer might develop. This might happen around the area where your breast was removed, or somewhere else in your body, including your other breast.

Your doctor tells you that there are some things you can do to help manage your ongoing risks of breast cancer. Two of these include having surgery to remove your currently unaffected breast plus monitoring, or keeping that breast and opting for monitoring only. She explains that these options differ in a number of ways:

- their impact on the chances of the cancer coming back;
- the types of monitoring you undergo and how often;
- the chance that you will experience pain or loss of sensation in your breast area;
- how much they cost you; and
- how often you are involved in the decisions about your care.

Your doctor would like you to think about and choose how you would like to manage your ongoing risks of breast cancer. Your choices are to have "Surgery to remove unaffected breast plus monitoring" or "Routine monitoring only".

You will now see 12 questions asking you to choose between these two options for managing your ongoing risks of breast cancer. Each time, please indicate which option you would prefer.

Women who responded to the pilot survey exhibited a high degree of "non-trading"; of the 87 women who completed the pilot survey 44 (50.6%) exhibited a strict preference for one option or the other; 34 for routine monitoring and 10 for CPM. Feedback on these results was that it was important to understand something about women's underlying attitudes to cancer risk and screening behaviours (see Appendix 13). Accordingly, questions from the Australian Health Survey[215] were included in the demographic section of the final survey to ask about women's prior cancer screening participation and smoking behaviours.

One of the key aims of this research is to understand how women value changes in the fear of recurrence, and how this might be influenced by different means of managing ongoing cancer recurrence risk. Thus, a measure of respondents' levels of cancer concern, as a proxy for fear of recurrence, was included in the main study. This was informed by the cancer concern measurement guidelines from Cancer Australia[317], focusing on measures of the level of worry about cancer recurrence, such as the Cancer Worry Scale.[xxxiv]

The following item was identified from the Cancer Australia Guidelines for modification for use in this survey: "*How much are you concerned/worried/ fearful about the cancer coming back?*" This item is drawn from work by Spencer (1999)[320] and Campbell (2000)[177]. In the former it was applied using a five point rating scale from 'not at all' to 'extremely' worried; in the latter gradations of the concern were measured over four statements that varied in the intensity with which concern impacted on daily life.

For use in this research, the question was modified such that the level of worry pertains to the attributes in the choice scenarios, rather than the respondent's own health. Following completion of the choice tasks, respondents were asked "*How much did concern or worry about each of the following factors influence the choices you made?*", with a corresponding table listing each of the attributes in the choice set (excluding breast sensitivity). Responses were rated on a five point scale from not at all to extremely. Fear of cancer recurrence was captured by a combined factor covering both *CBC Risk* and *Other BC Risk*. While concern about cancer recurrence was the primary measure of interest for this question, the other four factors were included to reduce the potential for respondents to unduly focus on cancer recurrence in their responses.

[xxxiv] Measures of fear of cancer recurrence are less relevant to this research because they are more closely associated with a prior experience of cancer, while worry might be more generalisable to an otherwise well population as reflected in the original development of the Cancer Worry Scale for use in individuals at risk of an hereditary cancer.[318,319]

6.2.4 Elements of the experiment

6.2.4.1 Choice options

Within this experiment, labelled choice options have been used: 'Surgery to remove unaffected breast plus monitoring' for CPM; and 'Routine monitoring only' for monitoring only. The word 'monitoring' appeared in both labels and the choice question, to make it clear that the CPM option did not preclude ongoing monitoring.

Women with early breast cancer might also receive endocrine therapies[xxxv] for the prevention of breast cancer recurrence or the emergence of new primary cancers.[288,311] However, the use of endocrine therapies is restricted to women with oestrogen receptor positive disease. Clinical advice was that endocrine therapies would not be considered as a substitute for CPM in women at high risk of recurrence (women who are BRCA 1/2 positive or with a family history of cancer), or women who are hormone receptor negative. As such, endocrine therapy was excluded as an alternative management option from the choices in this DCE.

6.2.4.2 DCE experimental design and survey development

The question of how to manage ongoing breast cancer recurrence has not previously been investigated using a DCE. As such, there were no coefficient values available from the literature that could be used to inform the survey design. Accordingly, for the pilot study a main effects orthogonal design was generated within Ngene without priors. For the final survey a WTP efficient design was estimated in Ngene, using the coefficient values from the analysis of the pilot survey as priors in the design (see Appendix 12).

The design included a number of restrictions:

[xxxv] Endocrine therapies include tamoxifen or the aromatase inhibitors letrozole, anastrozole and exemestane.

- *CBC Risk* for the CPM option could not exceed that of the routine monitoring only option;
- Mammogram as a level on *Monitoring* type could not appear for the CPM option;
- The attributes of *Pain Risk*, *Sensitivity Risk* and *Surgical OOP* applied only to the CPM option.

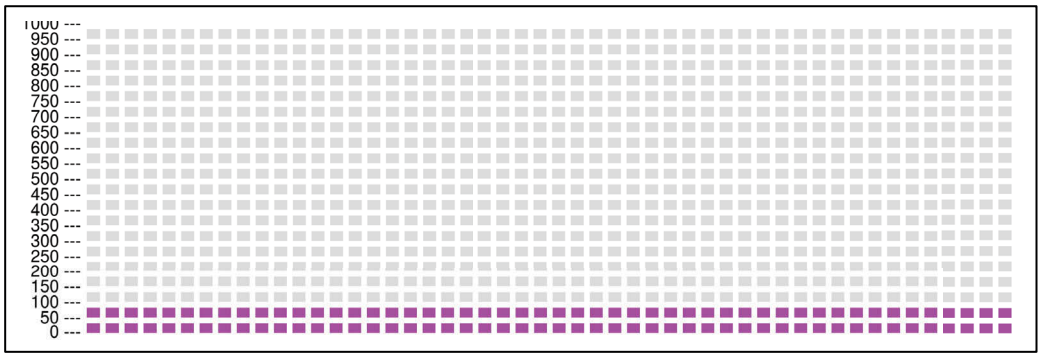
Within the pilot study, the combination of attributes and levels produced a $4^5 \times 3^4 \times 2^1$ design with 165,888 possible choice sets. The exclusion of the *Symmetry* attribute for the final survey altered the experimental design to a $4^5 \times 3^3 \times 2^1$ with 55,396 possible choice sets. The minimum possible number for estimation of an attribute balanced design (divisible by all the levels) is 12.[269] For both the pilot and the final studies, a design using a total of 48 rows was used in Ngene, in four blocks (incorporated into the design) of 12 choice sets, with an option pre-specified in Ngene to minimise the chance of dominant alternatives across choice sets. This design was replicated for each of the three frames ('Base', 'Goal' and 'No-Efficacy Difference'), giving a total of 12 possible blocks. Respondents to the survey were randomised to one of each of the 12 blocks, and to whether they saw the CPM option presented on the left or the right of the choice panel. The order of presentation of choice sets within each block was also randomised.

The DCE surveys were hosted online using the Qualtrics platform.[199] While Qualtrics has facilities for the use of conjoint and stated preference questionnaires, these are developed directly within the system and do not allow an existing survey design to be uploaded. Moreover, to enhance the communication of the four risk based attributes presented as proportions (*CBC Risk, Other BC Risk, Pain Risk, Sensitivity Risk*), icon-arrays were used as reported in Chapter 4 and Chapter 5. An example icon-array as used in this study is presented in Figure 36.

To facilitate the presentation of the choice scenarios in the format desired and preserving the experimental design, the process used for loading of the survey into

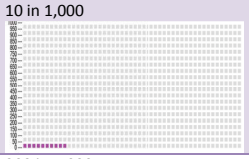
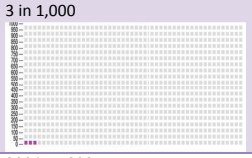
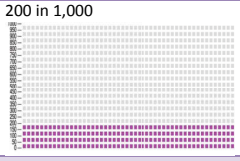
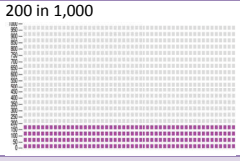
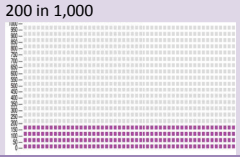
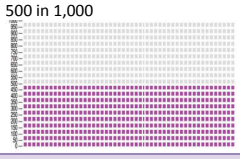
Qualtrics was to load each scenario, for each version as a separate picture. Full details of the development of the DCE experimental design and the process used to convert the Ngene designs into the required format are contained in Appendix 7. An example choice set is provided at Figure 37.

Figure 36: Icon-array, Mastectomy study



Note: Icon-array depicting a 10% (100 in 1,000) chance that your original cancer returns or spreads beyond your breasts some time in the next ten years.

Figure 37: Example choice set, Base frame, Mastectomy study

	Routine monitoring only	Surgery to remove unaffected breast plus monitoring
The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:	10 in 1,000 	3 in 1,000 
The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:	200 in 1,000 	200 in 1,000 
In addition to your regular self-checks, you will need to have the following tests:	Mammogram	MRI of your breast area
You are scheduled to have your follow-up tests every:	Year	Second year
The cost to you each year for monitoring is:	$300 per year	$300 per year
The cost to you for surgery associated with managing your ongoing risk of cancer recurrence is:	This does not apply	$15,000
The chance you will experience ongoing pain is:	This does not apply	200 in 1,000 
The chance you will experience an ongoing loss of sensitivity in your breast area is:	This does not apply	500 in 1,000 
Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence:	Always	Not very often

Which option would you choose?

☐ Surgery to remove unaffected breast plus monitoring

☐ Routine monitoring only

6.2.5 Model specification

The general approach to the modelling of discrete choice has been described in Chapter 4. The analysis was restricted to main effects only, using the following attributes; the meta-health effects for mode (*Ultrasound: Mammogram*, and *MRI:Mammogram)* and frequency of monitoring (*Annual:Biannual Checks,* and *Biennial:Biannual Checks*), and involvement in decision-making (*Involved Always:Rarely)*; health effects (*CBC Risk, Other BC Risk, Pain Risk,* and *Sensitivity Risk*); and OOP costs (*Monitoring OOP* and *Surgical OOP*). In addition, as this is a labelled experiment, the analysis includes an alternative specific constant (ASC) which reflects the preferences of respondents for routine monitoring over CPM (*Monitoring: CPM)* based on the

information carried in the label only. While the label does not vary between options or over choice sets, it captures individuals' persistent preferences for an option in each set that depend on unobserved attributes, and can therefore be included as a covariate for exploring choice behaviour.[239]

The four risk variables and two OOP cost variables were coded as being continuous (the linearity of the risk variables is subsequently tested using an alternative variable specification in which they are coded as categorical). These variables are expected to be negatively associated with the probability of choosing a breast cancer recurrence management option. The type and frequency of monitoring are categorical variables and are expected to be positively associated with the choice of management option as they indicate an improvement in convenience. Involvement in decision-making is also a categorical variable and is expected to influence choices positively.

For the analyses presented in this chapter, data were analysed using conditional logit regression, mixed logit regression and latent class analysis (equation 13). Latent class analysis was considered because the prior qualitative research suggested the potential for women to be grouped according to those who might hold strong prior notions of how they would choose to manage their ongoing risk of breast cancer recurrence risk (either CPM or routine monitoring), and those who might be influenced to choose between the two options. Latent class analyses can predict, for each individual in the analysis, the probability that they belong to each of the *G* classes. Within this research, these predicted class memberships were used to allocate women to the groups and to describe the demographic profile of each of the resulting classes.

6.2.6 The survey data and analysis

6.2.6.1 Data collection

Respondents for the DCE surveys were recruited using the Pureprofile online panel.[198] This panel has over 60,000 active members Australia-wide monthly. Availability of the

survey was posted to the portals of all women over the age of 16 years on the Pureprofile panel, without further restriction. The pilot survey was completed on 19th August 2015, with a target of 45 respondents. The final survey was available for completion during October 2015, with a target of 450 respondents. During that time, it was monitored to ensure a balanced distribution of respondents across each of the survey blocks.

6.2.6.2 Analysis plan

The pilot data were analysed using a conditional logit regression and mixed logit regression analyses, focusing on the results from the combined, or 'pooled' data. Exploratory analyses were conducted of the effects of demographic covariates using an alternative specific conditional logit analysis.

Analysis of the main survey used mixed logit regression and pooled the data across the three frames. Differences between frames were tested to identify whether subsequent analyses should be conducted using the pooled or unpooled data. Results from the mixed logit regression were used to estimate mWTP that were ranked in terms of their size relative to the largest value in each framing set. Differences in the mWTP for reductions in cancer risk were also investigated using two subgroup analyses: whether or not women traded between choice sets; and according to the degree of concern regarding cancer recurrence women reported thinking about when answering the choice scenarios. For the latter analysis, women were classified as 'cancer concerned' if they reported being very much or extremely concerned about the cancer coming back, and 'not cancer concerned' if they reported being concerned or less than concerned.

The number of classes for the latent class analysis was chosen by first estimating multiple iterations of the analysis in which g was allowed to vary from 2 to 8. The number of classes for the final analysis was based on the level of g which produced the most efficient model in terms of the lowest AIC and BIC values[188,238], and that was

consistent with what was already known about potential differentiators between respondents that would form specific choice groups.

6.2.6.2.1 Maximisation procedure

Estimation commands and maximisation procedures for the conditional logit and mixed logit regression were as described in Chapter 5. All analyses were conducted in STATA 12, and copies of the STATA code are available upon request. Use of the latent class involved the STATA user written command *lclogit*.[238] Model fit criteria such as the overall log-likelihood, model significance, pseudo-R^2 and information criteria (AIC/BIC) were reported for each estimation procedure. While STATA does not report the pseudo-R^2 for mixed logit regression and latent class analyses, these could be estimated using the method described by Greene (2008)[56]; that is 1 – (Log likelihood 1/ Log likelihood Null), where the Null model is described by the intercept term only (the ASC), and model 1 is the relevant model of interest.

6.2.6.3 Test of poolability

The impact of framing effects on the coefficient estimates can be tested globally by assessing whether the data from the three versions of the survey ('Base', 'No-Efficacy Difference' and 'Attribute') can be pooled. The methods for testing the suitability for data pooling were applied as previously described in Chapter 5.

6.2.6.4 Estimating mWTP

Estimates of the mWTP were formed for all health effect and meta-health effect attributes using *Monitoring OOP* as the numeraire. *Surgical OOP* has not been used as the numeraire since this applies to the CPM only option. mWTP values were tabulated for those coefficients that were statistically significant, but all mWTP results are presented graphically. Estimates of mWTP and corresponding confidence intervals were produced in the same manner as described in Chapter 5.

As previously noted, one of the specific preference measures of interest in this study is the value associated with reducing the fear of cancer recurrence. This cannot be measured directly since 'fear' was not included as an attribute in the DCE. However, the extent to which women differ in the value they place on achieving reductions in *CBC Risk*, or *Other BC Risk* can be assessed based on the underlying subgroups of women who reported being concerned about cancer recurrence while answering the choice scenarios. mWTP estimates health effects and meta-health effects were thus formed according to the subgroups specified by the degree of cancer concern.

6.3 Results

6.3.1 The pilot

Results for the pilot survey are discussed in Appendix 13. A total of 87 women completed all aspects of the survey. Feedback on the survey language and design indicated that it was well understood by respondents; the language and the tasks were clear. A key finding from the pilot survey was the extent of non-trading evident in responses. Among the 87 women 44 (50.6%) always chose one option over the other; 34 (39.1%) always chose the routine monitoring only option, and 10 (11.5%) always chose the CPM only option. The significant factors influencing choice in both the conditional logit and mixed logit regressions were the health effects of *CBC Risk*, *Other BC Risk* and costs (*Monitoring OOP* and *Surgical OOP*). The significance of the meta-health effects depended on framing effects; both *Involved* and *Ultrasound: Mammogram* were significant, but only in the 'No-Efficacy Difference' frame. Given the large proportion of non-traders, the ASC, *Monitoring: CPM*, explained much of the choice behaviour. Comparisons of framing effects for the mixed logit analysis showed that there were some differences apparent across the three framing sets, with the 'Base' frame generally producing larger coefficients across all attributes than the other two frames.

Exploratory analysis of the influence of demographic variables using an ASC logit analysis of the pooled data supported the findings of the mixed logit regression analysis with respect to attribute significance (including those for the meta-health effects)(see Appendix 9). The only demographic variable of influence was whether or not women were classified as being cancer concerned. Being cancer concerned statistically significantly reduced the likelihood of choosing routine monitoring only compared with CPM only (as expressed via the ASC).

6.3.2 The final survey – participants and choices

6.3.2.1 Demographics

A total of 464 women completed the online survey. Selected demographic characteristics for respondents are provided in Table 21. In general, women who responded to the survey were skewed towards the middle age groups and tended to be better educated than the general population of Australia. They were similar to the general population in terms of income and geographic location.[321-323] Data from the AIHW indicate that 36% of cancers among women in 2007 were breast cancers, but these occurred in 55,537 women across Australia.[324] This would suggest a low prevalence rate when adjusting for the population of women in that year (over 10 million), and lower than that reported in this sample. However, the AIHW figure is restricted to the last five years, and would be higher if adjusted for the relevant age adjusted population.

Table 37: Final survey - demographics

	Pooled n=464	Base n=158	No-Efficacy Difference n=161	Goal n=145	Australian Popn.
Age					
16-24	19 (4.09)	6 (3.77)	10 (6.29)	3 (1.89)	(15.62)
25-44	168 (36.21)	56 (35.22)	62 (38.99)	50 (31.45)	(34.52)
45-64	175 (37.72)	62 (38.99)	55 (34.59)	58 (36.48)	(30.37)
65-74	81 (17.46)	28 (17.61)	25 (15.72)	28 (17.61)	(10.45)
75 or Over	13 (2.8)	1 (0.63)	7 (4.4)	5 (3.14)	(9.05)
Unknown	8 (1.72)	5 (3.14)	2 (1.26)	1 (0.63)	
			p<0.001		
Household Income					
Median Category, $	1,150-1,529	1,150-1,529	1,150-1,529	1,150-1,529	1,234
			p=0.003		
Education					
School Only	136 (29.31)	48 (30.38)	51 (31.68)	37 (25.52)	(42.62)
University	141 (30.39)	42 (26.58)	42 (26.09)	57 (39.31)	(27.54)
Vocational	178 (38.36)	63 (39.87)	65 (40.37)	50 (34.48)	(28.14)
Unknown	9 (1.94)	5 (3.16)	3 (1.86)	1 (0.69)	n.a.
			p<0.001		
Residence					
Major City	307 (66.16)	100 (63.29)	107 (66.46)	100 (68.97)	(71.30)
Inner Regional	89 (19.18)	33 (20.89)	31 (19.25)	25 (17.24)	(18.29)
Outer Regional	32 (6.9)	10 (6.33)	10 (6.21)	12 (8.28)	(8.68)
Remote	3 (0.65)	1 (0.63)	2 (1.24)	0 (0)	(1.73)
Unknown	33 (7.11)	14 (8.86)	11 (6.83)	8 (5.52)	
			p<0.001		
At Least One Chronic Health Issue	267 (57.54)	92 (58.23)	92 (57.14)	83 (57.24)	n.a.
			p=0.973		
Prior Breast Cancer	16 (3.51)	4 (2.61)	4 (2.52)	8 (5.56)	n.a.
			p<0.001		

Note: Population distribution for Australia is based on all non-indigenous persons.
Abbreviation: n.a., not applicable.

The survey sample was monitored for balance in terms of the allocation of respondents to the respective blocks. Potential differences between the frames in terms of the demographic composition of respondents were tested using one-way tests of analysis of variance (Bartlett's test of equivalence of variance), with the demographic factor as the response variable and the frame as the factor or explanatory variable. The results of these tests showed that there were statistically significant differences across the frames with respect to age $\chi^2(2)= 52.69$ ($p<0.001$); income $\chi^2(2)= 11.82$ ($p=0.003$); education $\chi^2(2)= 1.8e+0.3$ ($p<0.001$); location of residence $\chi^2(2)= 94.84$ ($p<0.001$); and

prior BC status $\chi^2(2)= 700.08$ ($p<0.001$). There were no statistically significant differences for chronic health status $\chi^2(2)= 0.05$ ($p=0.973$).

6.3.2.2 Underlying risk behaviour

The results with respect to women's participation in screening are presented by age group (one of the principal defining characteristics for screening participation), and according to whether women had ever participated in cancer screening (Figure 38) or had participated in screening in the last two years (Figure 39). Screening participation, across all cancer types, in all age groups was less than 50% without limiting the screening period; this declined to less than 40% when restricted to the last two year period – the interval of interest for both the National Cervical Screening and Breast Cancer Screening programmes. Data from the AIHW indicate that in the last two years 54% of women aged 50-69 years had a mammogram for breast cancer screening, 57% of those aged 20-69 years had a Pap test for cervical cancer screening, and 38.5% of those aged 50-65 years participated in screening for bowel cancer.[325,326] Participation in screening among women in the Australian population is thus higher than among those in the survey. This suggests that women in the survey might be willing to accept more risk as it pertains to cancer, or are less concerned about cancer occurrences, than the general population of Australian women.

Figure 38: Do respondents participate in cancer screening?

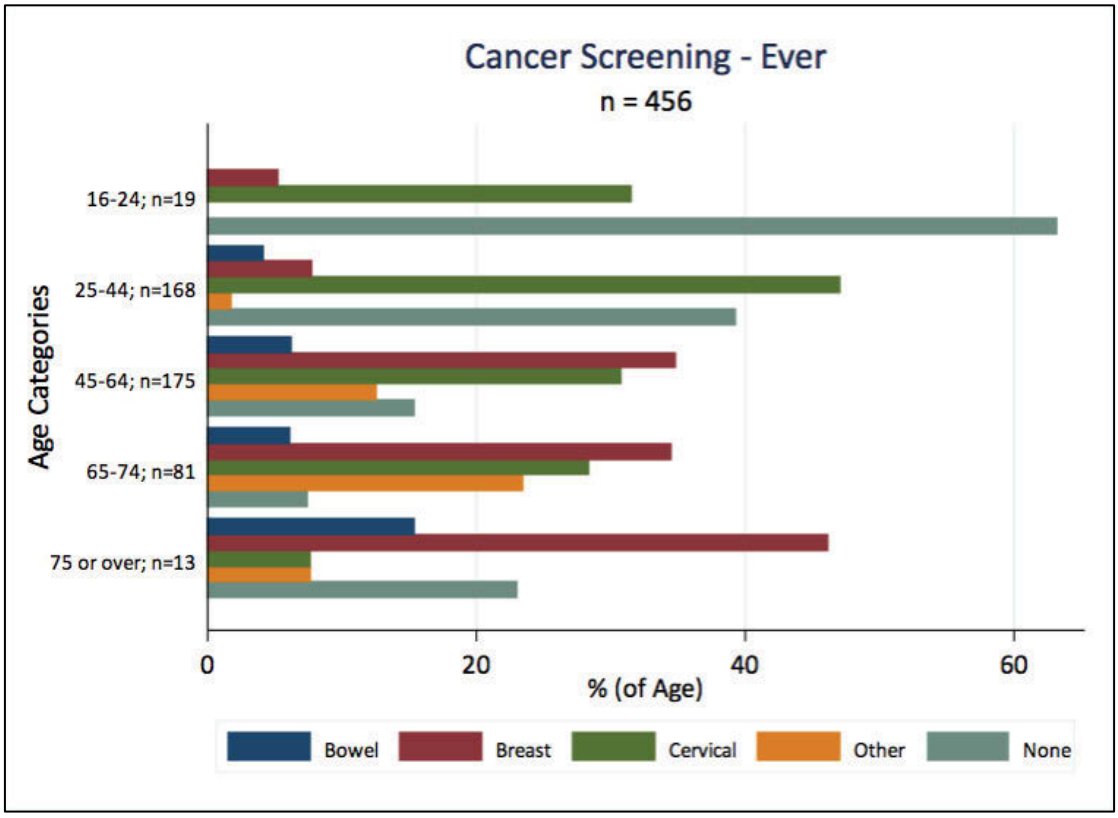
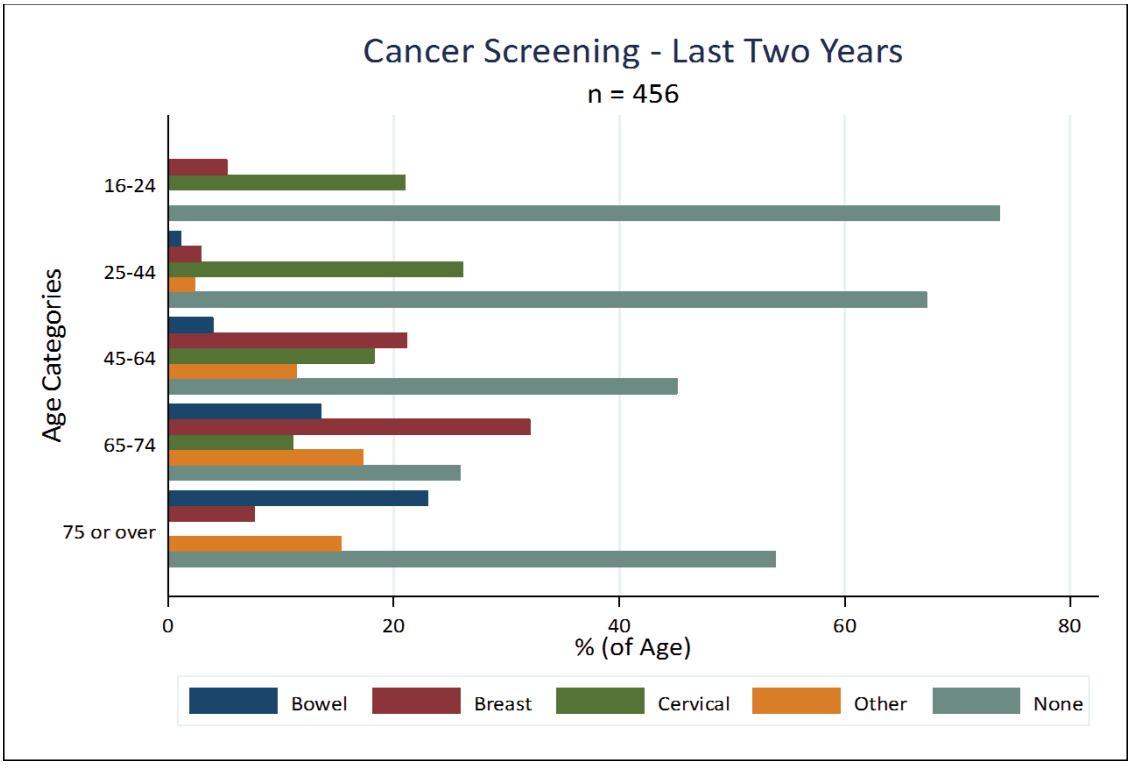


Figure 39: Did respondents participate in cancer screening in the last two years?

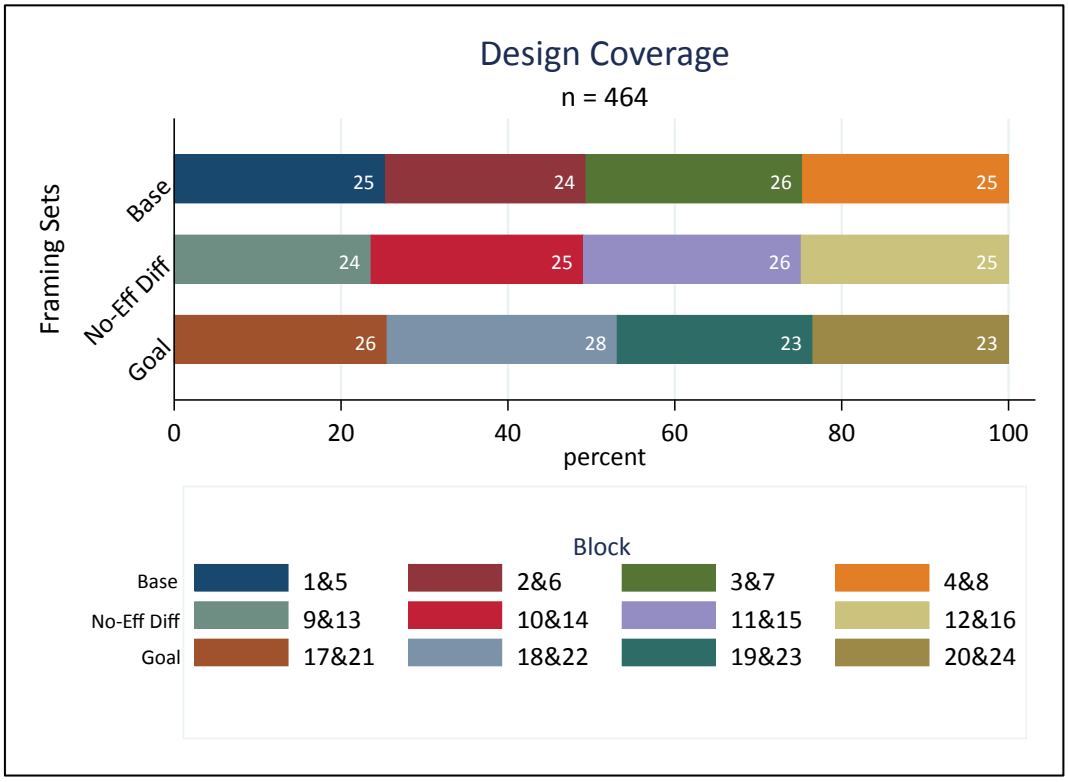


6.3.2.3 Survey coverage

The DCE survey design consisted of 48 choice tasks, allocated into four blocks of 12 choice sets, and replicated over the three different framing sets ('Base', 'Goal' and 'No-Efficacy Difference'). Women were randomised to which block they saw, but within blocks also to whether or not they saw each option on the left or right of the screen. This meant that the total number of blocks per survey version presented in Qualtrics was doubled (but retained the same design of 48 choice sets). Each frame was thus allocated eight block numbers in the randomisation process (allowing for randomisation of labeled choice options to appear on the left or right), giving 24 block numbers in total.

The results Figure 20 show that within each framing set, the coverage of respondents across the blocks was reasonably balanced; the proportion of total respondents allocated to each block was essentially the same within and across the three frames. Some variation arose with respect to the 'Goal' framing set in that there was a difference of six respondents (there were a total of 145 allocated to this framing set) between blocks 18&22 (40 respondents) and the last two blocks in that frame (19&23 and 20&24, each with 34). A test of the significance of this difference (using an alternative specific conditional logit) showed that block allocation was not significant, either for the overall sample or when restricted to the 'Goal' framing set (see Appendix 14).

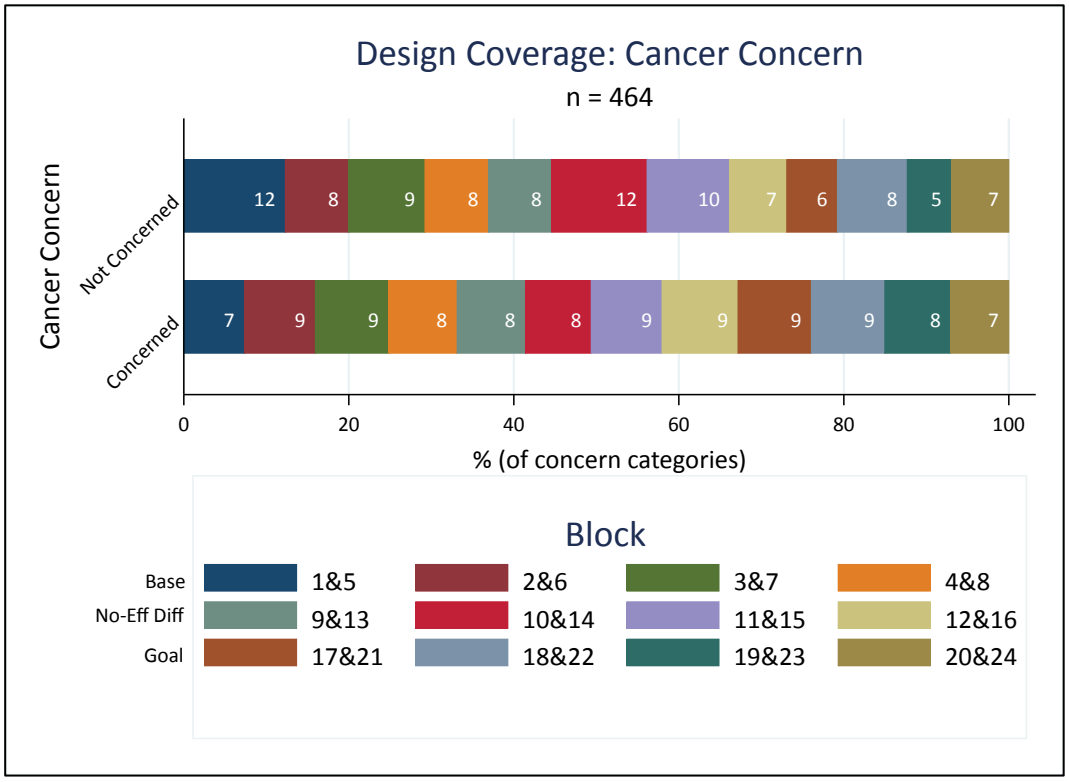
Figure 40: Design coverage by frame



Note: Percentages shown in bars are of the total survey sample.

One potential stratification of respondents when investigating choices is their degree of cancer concern, noting that this would be an ex-post subgroup. Design coverage stratified for cancer concern is shown in Figure 21. The results indicate some imbalance across the framing sets in terms of which blocks appeared for the not-cancer concerned group: a higher proportion of respondents in block 1&5 for the 'Base' frame than the remaining blocks in that frame; more in block 10&14 of the 'No-Efficacy Difference' frame than its remaining blocks; and fewer in block 19&23 of the 'Goal' frame than its other blocks. A test of association between cancer concern and block allocation was statistically significant; Bartletts's $\chi^2(11)= 62.53$ ($p<0.001$). Despite the apparent link between block allocation and the subsequent classification of women according to cancer concern, the latter was used to form an exploratory subgroup analysis to investigate choice behaviour.

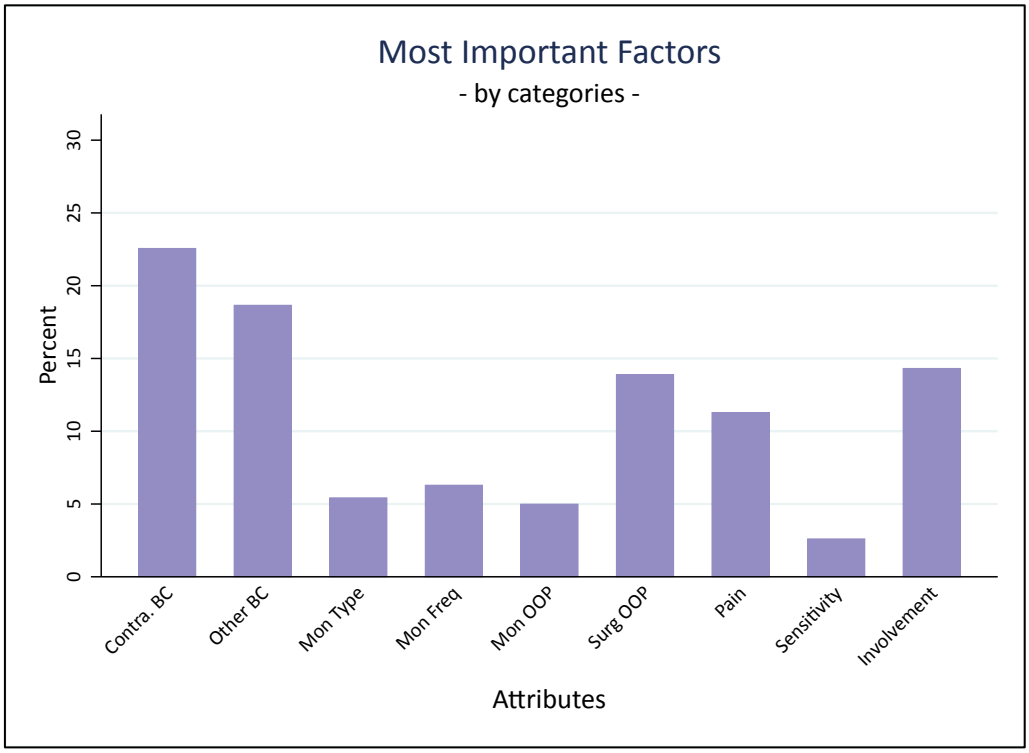
Figure 41: Design coverage by cancer concern



6.3.2.4 Attribute rankings in decisions

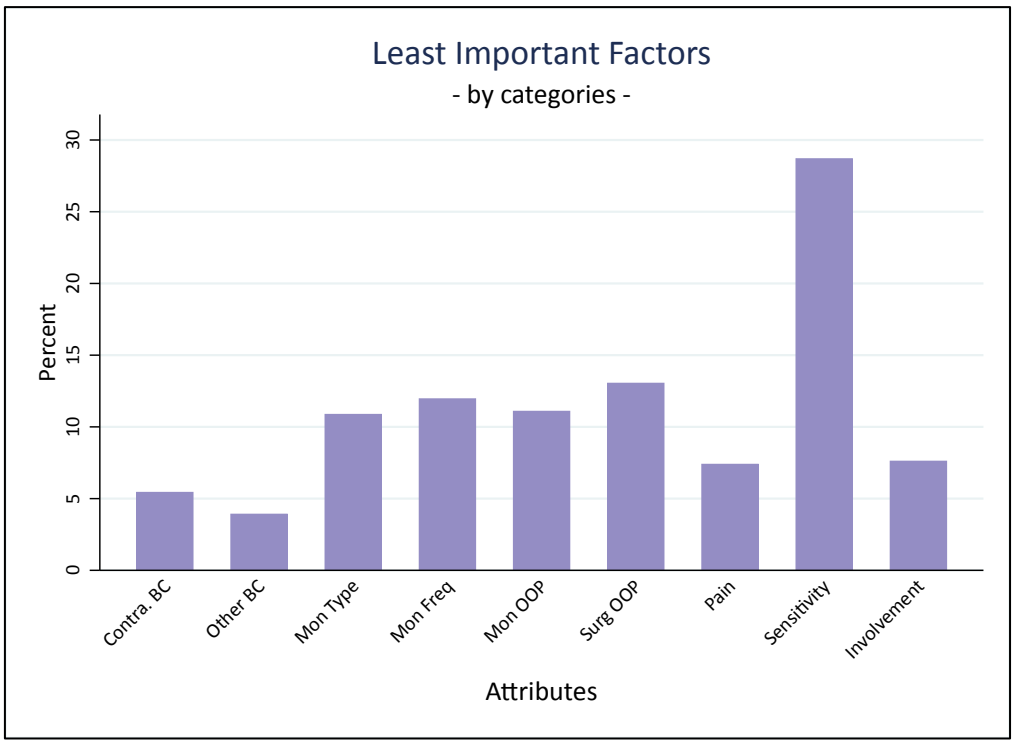
After completing the 12 choice sets, women were asked to report which of the attributes they considered to be the most important when making their choices, and which the least important. As can be observed from the results in Figure 42, the majority of women considered *CBC Risk* and *Other BC Risk* to be the most important factors in their decision-making, followed by the meta-health effect of *Involvement* (extent of involvement in decision-making) and *Surgical OOP*. The risk of losing breast *Sensitivity* was considered by women to be least most important factor as shown in Figure 43.

Figure 42: Attributes considered most important in choice task



Abbreviations: BC, breast cancer; OOP, out-of-pocket.

Figure 43: Attributes considered least important in choice task

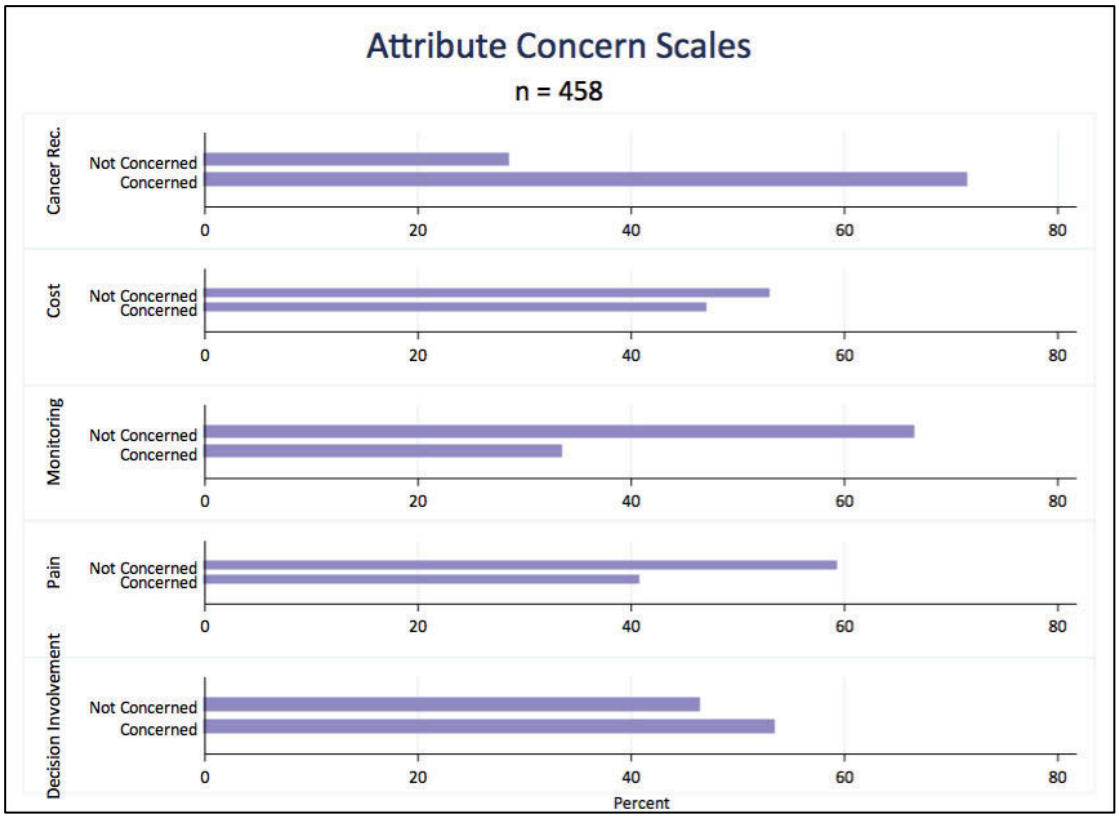


Abbreviations: BC, breast cancer; OOP, out-of-pocket.

Attribute Concern

The results for how women rated the extent to which they were concerned about the attributes while they were completing the choices tasks are provided in Figure 44. From these results it can be observed that the attribute with which most women were concerned was that of cancer recurrence; 71.5% of women were classified as being concerned with cancer recurrence. Of the remaining four attributes for which concern was assessed, there was either a reasonable balance between women who were concerned or not concerned (*Cost* of care and *Involvement*) or the majority of women were not concerned (*Pain Risk* and *Monitoring* type and frequency).

Figure 44: Attribute concern during the choice tasks

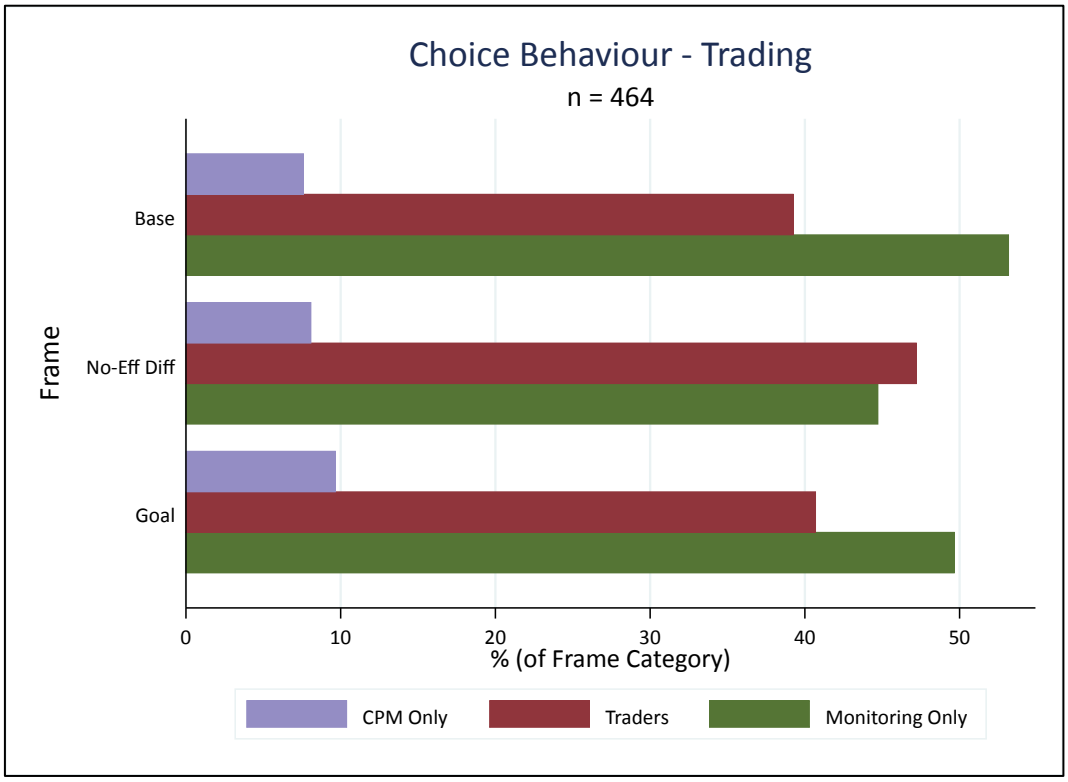


Choice Behaviour

As with the pilot survey, women were categorised into three groups according to their choice behaviour; those who only ever chose the option for surgery (CPM Only), those who only ever chose the option for routine monitoring (Monitoring Only), and those

who chose either option to a greater or lesser extent (Traders). Women in the CPM Only or Monitoring Only categories can be referred to as non-traders, and have been retained in the analysis. The effect of excluding their results from the analysis is tested as part of the investigation of heterogeneity of responses. The results of the count of choice behaviour, reported by framing set, are provided in Figure 45. Overall, 57.5% of women were non-traders; 8.4% always preferring CPM Only and 49.1% always preferring Monitoring Only. On balance, there were always more non-traders than traders, but the proportion of traders was higher for the 'No-Efficacy Difference' (47.2%) frame compared with either the 'Base' (39.2%) or 'Goal' (40.7%) frames.

Figure 45: Trading behaviour, by frame



Abbreviation: CPM, contralateral prophylactic mastectomy.

An initial analysis of women's choice behaviour was conducted by describing traders and non-traders using the underlying demographic characteristics, including whether or not women were classified as being cancer concerned. The results in Table 38 show

that a higher proportion of women were classified as cancer concerned among the traders than the non-traders, and that among the non-traders the majority of those in the CPM Only group were cancer concerned. Non-traders were slightly older, included a higher proportion of women with vocational education, and exhibited more prior screening experience for cervical cancer (not breast cancer) than women who traded. A higher proportion of women in the CPM Only group reported that they knew someone who had been treated for breast cancer (*BC Aware*) compared with the other groups.

Table 38: Respondent demographics by choice behaviour and cancer concern

	Non-Traders				**Traders**	
	CPM Only		***Monitoring Only***			
	Not Cancer Concerned n = 5	*Cancer Concerned* n = 33	*Not Cancer Concerned* n = 90	*Cancer Concerned* n = 132	*Not Cancer Concerned* n = 35	*Cancer Concerned* n = 161
Age						
16-24	0 (0)	0 (0)	5 (5.56)	2 (1.52)	1 (2.86)	11 (6.83)
25-44	2 (40.00)	12 (36.36)	30 (33.33)	42 (31.82)	15 (42.86)	66 (40.99)
45-64	3 (60.00)	14 (42.42)	34 (37.78)	53 (40.15)	13 (37.14)	58 (36.02)
65-74	0 (0)	7 (21.21)	15 (16.67)	29 (21.97)	6 (17.14)	22 (13.66)
75 or Over	0 (0)	0 (0)	5 (5.56)	5 (3.79)		3 (1.86)
Unknown	0 (0)	0 (0)	1 (1.11)	1 (0.76)		1 (0.62)
			p=0.232[a]	p<0.001[b]		
Income						
Under $39,999	1 (20.00)	9 (27.27)	28 (31.11)	30 (22.73)	8 (22.86)	50 (31.06)
$40,000-$79,999	3 (60.00)	11 (33.33)	20 (22.22)	33 (25.00)	8 (22.86)	40 (24.84)
$80,000-$149,999	1 (20.00)	7 (21.21)	20 (22.22)	37 (28.03)	7 (20.00)	35 (21.74)
Over $150,000	0 (0)	1 (3.03)	5 (5.56)	11 (8.33)	3 (8.57)	13 (8.07)
Unknown	0 (0)	5 (15.15)	17 (18.89)	21 (15.91)	9 (25.71)	23 (14.29)
			p=0.001[a]	p<0.001[b]		
Education						
School Only	1 (20.00)	12 (36.36)	25 (27.78)	38 (28.79)	9 (25.71)	50 (31.06)
University	1 (20.00)	11 (33.33)	20 (22.22)	43 (32.58)	13 (37.14)	52 (32.3)
Vocational	3 (60.00)	10 (30.30)	43 (47.78)	50 (37.88)	13 (37.14)	58 (36.02)
Unknown	0 (0)	0 (0)	2 (2.22)	1 (0.76)	0 (0)	1 (0.62)
			p<0.001[a]	p=0.044[b]		
BC Awareness and History						
BC Aware	0 (0)	26 (78.79)	51 (56.67)	89 (67.42)	18 (51.43)	111 (68.94)
BC Not-Aware	4 (80.00)	6 (18.18)	32 (35.56)	35 (26.52)	17 (48.57)	44 (27.33)
Prior BC	1 (20.00)	0 (0)	5 (5.56)	6 (4.55)	0 (0)	4 (2.48)
Unknown	0 (0)	1 (3.03)	2 (2.22)	2 (1.52)	0 (0)	2 (1.24)
			p<0.001[a]	p=0.244[b]		

	Non-Traders				Traders	
	CPM Only		*Monitoring Only*			
	Not Cancer Concerned	*Cancer Concerned*	*Not Cancer Concerned*	*Cancer Concerned*	*Not Cancer Concerned*	*Cancer Concerned*
	n = 5	n = 33	n = 90	n = 132	n = 35	n = 161
Cancer Screening Behaviour						
Bowel	0 (0)	3 (9.09)	5 (5.56)	7 (5.30)	2 (5.71)	8 (4.97)
Breast	1 (20.00)	6 (18.18)	13 (14.44)	34 (25.76)	10 (28.57)	46 (28.57)
Cervical	0 (0)	12 (36.36)	32 (35.56)	53 (40.15)	7 (20.00)	56 (34.78)
Other	0 (0)	6 (18.18)	13 (14.44)	9 (6.82)	3 (8.57)	14 (8.70)
None	4 (80.00)	6 (18.18)	27 (30.00)	28 (21.21)	13 (37.14)	36 (22.36)
Unknown	0 (0)	0 (0)	0 (0)	1 (0.76)	0 (0)	1 (0.62)
			$p<0.001^a$	$p<0.001^b$		

Notes: p values shown are for the significance of the Bartlett's test of equality of variance in a one-way analysis of variance test. Results shown for BC Awareness and History are for BC Awareness only as there were just 16 women with a prior diagnosis of BC.
a shows p values for test with cancer concern as the response variable and each demographic characteristic shown as the relevant factor variable.
b shows p values for test with trading status as the response variable and each demographic characteristic shown as the relevant factor variable.

Abbreviations: BC, breast cancer; CPM, contralateral prophylactic mastectomy.

6.3.3 Determinants of choice

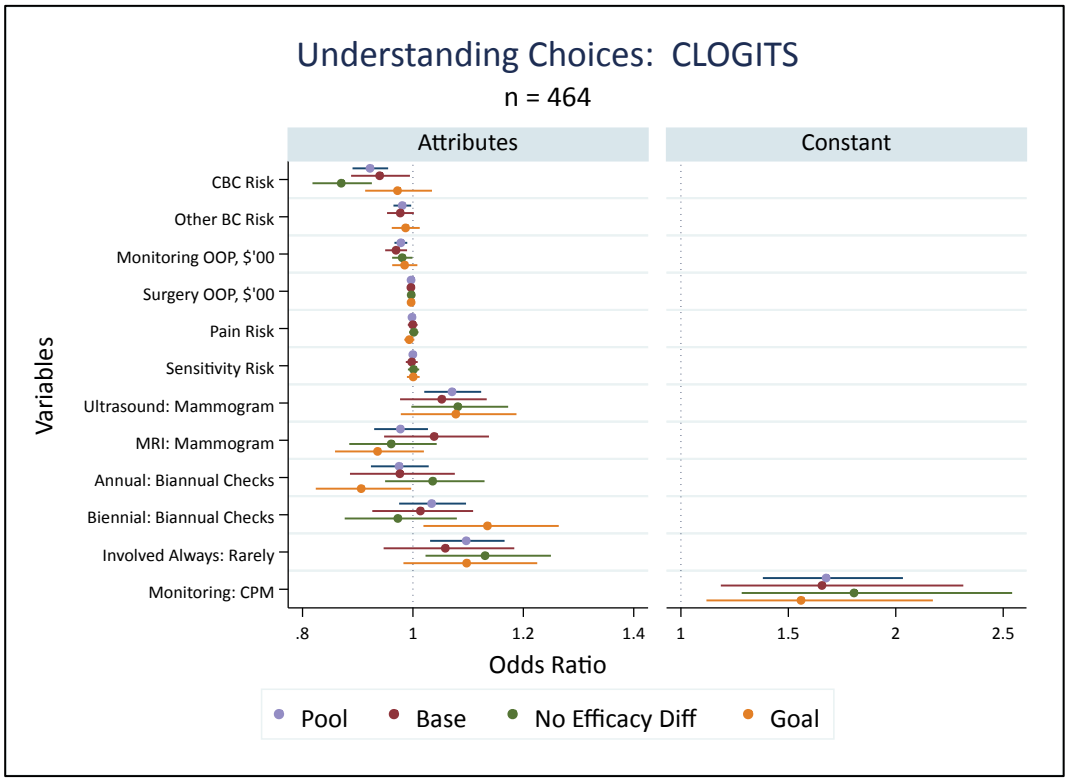
6.3.3.1 Conditional logit regression

Results of the conditional logit analysis are shown graphically in Figure 46 (with the tabulated results presented in Appendix 14). From these results it can be observed that the dominant factor explaining women's choices were the ASCs; across all three framing effects there is a higher likelihood of choosing Monitoring Only compared with CPM Only, regardless of any other information presented. This finding is to be expected given the high proportion of women who always preferred the Monitoring Only option. Of those factors which did influence choice, higher rates of both types of breast cancer risk were associated with a lower likelihood of choice across all frames, except the 'Goal' frame. Similar findings were observed for *Monitoring OOP* and *Surgical OOP*. *Pain* and *Sensitivity Risk*, which applied only to the CPM Only option, were not statistically significant. The results for the meta-health effects were mixed. Mode of monitoring was significant only in the pooled analysis for the comparison of ultrasound with mammogram, resulting in an increase in the likelihood of choice –

women prefer a less intrusive monitoring modality. The frequency of monitoring influenced choice in the 'Goal' frame only, but there appeared to be a preference inversion in these results; annual monitoring reduced the likelihood of choice compared to biannual, while biennial increased the likelihood of choice. Finally, being involved in treatment decisions increased the likelihood of choice in the pooled analysis and the 'No-Efficacy Difference' frame, but was not significant in the remaining frames.

These results were essentially replicated in the ASC logistic regression (see Table A 30). The inclusion of demographic characteristics into alternative estimations of this analysis showed that being cancer concerned reduced the likelihood of choice. No other demographic characteristics statistically significantly influenced choice.

Figure 46: Conditional logits – pooled analysis and framing effects



Abbreviations: BC, breast cancer; CPM, contralateral prophylactic mastectomy; OOP, out-of-pocket.

6.3.3.2 Mixed logit regression

The results of the mixed logit regressions are presented in Table 39. As observed previously, given the large proportion of women who displayed a strict preference for one option over the other (non-trading), the ASC is statistically significant across all four specifications and favours Monitoring Only over CPM Only. Where women were influenced by the attributes describing the options presented, the results are largely as observed for the conditional logit regression. Increasing cancer risks (both *CBC Risk* and *Other BC Risk*) statistically significantly deterred women from choosing a management option. This effect was significant for the pooled data, but varied across framing sets; the effect of *CBC Risk* was only significant for the 'No-Efficacy Difference' frame and *Other BC Risk* was only significant for the 'Base' frame. Similarly, increasing *Pain* and *Sensitivity Risk* both deterred women from choosing CPM (by design these variables did not apply to Monitoring). Once again this varied by frame, with neither being significant in the 'No-Efficacy Difference' frame, both appearing as significant in the 'Goal' frame, and only *Sensitivity Risk* being significant for the 'Base' frame. As would be anticipated, both *Monitoring OOP* and *Surgical OOP* had negative signs, but neither was statistically significant in the 'Goal' frame. Of the meta-health effects, a preference was observed for less invasive types of monitoring with *Ultrasound: Mammogram* being positively signed (increasing the likelihood of choice) and significant in all frames except the 'Goal' frame; *MRI:Mammogram* was not statistically significant. However, women in the 'Goal' frame were the only respondents for whom the frequency of monitoring influenced choice with a shift to biennial compared with biannual screening increasing the likelihood of choice.

The results for the standard deviations indicate heterogeneity across women in the influence of all attributes (although not all levels of all attributes in all frames) in the choice of ongoing breast cancer risk management. Women differed in how they viewed the ASC, justifying the inclusion of this attribute as a random rather than fixed parameter. Women also differed across all three frames with respect to the heterogeneity of the values placed on *CBC Risk*. This is to be expected given the results

of the qualitative research that observed differences across women with experience of BC in terms of how they approached the risk of a contralateral breast cancer.

Table 39: Respondent heterogeneity - mixed logit analysis and framing effects

	Pooled	Base	Goal	No-Eff Diff
Attribute Means				
CBC Risk	-0.134	-0.139	0.003	-0.301
	(0.052)*	(0.106)	(0.128)	(0.080)**
Other BC Risk	-0.052	-0.090	-0.036	n.a.
	(0.019)**	(0.038)*	(0.030)	n.a.
Monitoring OOP, $'00	-0.077	-0.176	-0.069	-0.046
	(0.017)**	(0.043)**	(0.036)	(0.023)*
Surgery OOP, $'00	-0.011	-0.020	-0.014	-0.010
	(0.002)**	(0.005)**	(0.007)	(0.002)**
Pain Risk	-0.011	-0.019	-0.030	-0.004
	(0.006)	(0.014)	(0.012)**	(0.011)
Sensitivity Risk	-0.014	-0.036	-0.068	-0.010
	(0.006)*	(0.017)*	(0.026)**	(0.009)
Ultrasound: *Mammogram*	0.191	0.300	0.196	0.187
	(0.057)**	(0.153)*	(0.153)	(0.090)*
MRI: *Mammogram*	-0.013	0.188	-0.087	-0.069
	(0.059)	(0.132)	(0.121)	(0.099)
Annual: *Biannual Checks*	-0.023	-0.035	-0.262	0.171
	(0.061)	(0.139)	(0.137)	(0.088)
Biennial: *Biannual Checks*	0.018	-0.001	0.329	-0.160
	(0.056)	(0.102)	(0.151)*	(0.090)
Involved Always: *Rarely*	0.177	0.104	0.211	0.222
	(0.055)**	(0.123)	(0.101)*	(0.085)**
Monitoring: *CPM* Constant	1.245	1.637	0.078	1.306
	(0.220)**	(0.501)**	(0.380)	(0.346)**
Attribute Standard Deviation				
CBC Risk	0.479	0.441	0.459	0.507
	(0.097)**	(0.113)**	(0.124)**	(0.119)**
Other BC Risk	0.071	0.105	0.120	n.a.
	(0.041)	(0.070)	(0.058)*	n.a.
Monitoring OOP, $'00	-0.090	0.088	0.171	0.022
	(0.041)*	(0.045)	(0.080)*	(0.034)
Surgery OOP, $'00	0.012	0.025	0.014	0.006
	(0.002)**	(0.005)**	(0.011)	(0.003)
Pain Risk	0.002	0.060	-0.025	0.035
	(0.012)	(0.019)**	(0.009)**	(0.015)*
Sensitivity Risk	-0.033	0.075	0.107	0.026
	(0.011)**	(0.019)**	(0.022)**	(0.005)**
Ultrasound: *Mammogram*	0.040	-0.286	-0.212	-0.174
	(0.047)	(0.170)	(0.187)	(0.202)
MRI: *Mammogram*	-0.308	0.417	-0.376	-0.329
	(0.106)**	(0.205)*	(0.174)*	(0.181)

	Pooled	Base	Goal	No-Eff Diff
Annual: *Biannual Checks*	0.057	-0.550	-0.101	-0.000
	(0.187)	(0.234)*	(0.240)	(0.043)
Biennial: *Biannual Checks*	0.059	0.207	-0.360	-0.025
	(0.076)	(0.158)	(0.222)	(0.044)
Involved Always: *Rarely*	0.313	-0.346	-0.191	0.418
	(0.112)**	(0.271)	(0.175)	(0.150)**
Monitoring: *CPM* Constant	2.287	2.852	-1.308	2.182
	(0.195)**	(0.453)**	(0.213)**	(0.310)**
Observations	11,136	3,792	3,480	3,864
Individuals	464	158	145	161
Wald Chi	125.88	36.05	32.41	71.03
d.f.	12	12	12	11
p-value	<0.001	<0.001	<0.001	<0.001
Log-likelihood	-1,744.54	-548.18	-522.83	-647.21
Pseudo R^2	0.058	0.082	0.058	0.072

Notes: * p<0.05; ** p<0.01
All categorical data are effects coded to minimise the potential for contamination of the base effect in the grand mean.
Base categories are shown in italics.
All analyses were conducted in STATA 12, with robust standard errors (in parentheses) to account for the survey nature of the data.
Analyses conducted with 1,000 replications.

Abbreviations: BC, breast cancer; CBC, contralateral breast cancer; CPM, contralateral prophylactic mastectomy; d.f., degrees of freedom; OOP, out-of-pocket.

6.3.3.2.1 Tests of pooling

The results of the LR tests for the ability to pool the data across the three framing sets are presented in Table 40. Based on the results from the two-way LR tests, estimate comparability is rejected. The LR test comparing the combined LLH from the three samples with that of the pooled regression was also statistically significant, indicating that the data from the three framing sets should be analysed separately.[259,271]

Table 40: LR test of poolability

	LLH from Mixed Logit	**d.f.**	**LR Test X^2 Statistic (d.f.)**		
			No-Efficacy Difference	*Goal*	*Pooled*
Base	-548.18	12	198.05 ***(1)	50.71***(1)	
No-Efficacy Difference	-647.21	11		248.77**(1)	
Goal	-522.83	12			
Summed	-1718.22	35			52.64***(23)
Pooled	-1,744.54	12			

Notes: * p<0.05, ** p<0.01, *** p<0.001
LR Test for the comparison of the framing sets was constructed as 2*(LLHFrame1-LLHFrame2) with d.f. equal to the difference of the d.f. of the respective framing sets. This means that for the Base and Goal framing sets which each have 12 d.f. a X^2 test statistic could not be estimated for d.f.=0, so is approximated with d.f.=1.
Abbreviation: d.f., degrees of freedom; LLH, log-likelihood; LR, likelihood ratio.

The results of the Chow test do not support the findings of the LR tests that the data from the three framing sets are best considered separately. That is, the hypothesis that the frame specific dummies interacted with the attributes are jointly zero could not be rejected at a standard 5% level of significance (χ^2 (23) = 33.18; p = 0.078). Frame specific effects were noted only for the frequency of monitoring attributes, with an apparent preference inversion occurring between *Annual* and *Biennial* monitoring for the 'No-Efficacy Difference' and 'Goal' frames, and for *CBC Risk* in the 'No-Efficacy Difference' frame. Specific results of the Chow test regression are provided in Appendix 14, Table A 32. Given the small number of significant interactions, it is not surprising that the Chow test does not support separability of the framing sets since it is estimated based on the occurrence of frame specific attribute effects within the pooled data set (and hence is comparing attribute specific coefficients), rather than comparing the overall model performance captured by the Log-Likelihood for a given respondent set (as with the LR test).

6.3.3.2.2 Linearity of effects

The results of the mixed logit regressions including the categorically coded risk attributes are provided in Table A 33, in Appendix 14. The linearity of the four risk variables, *CBC Risk, Other BC Risk, Pain Risk and Sensitivity Risk,* in the mixed logit

regression was tested in two ways: using an unscaled and a scaled Wald test. The unscaled Wald test was used to test whether the coefficients for each level (compared with a base level) were the same within a given attribute, within a framing set. For the scaled Wald test, the coefficients for each level within an attribute were scaled by their proportional differences from the base level, prior to testing for equivalence between levels.

The results of the Wald tests of linearity are presented in Table 29, and indicate that differences exist across attributes and framing sets with respect to whether or not the risk variables should be treated as categorical or continuous. For the 'Base' frame, it would be reasonable to treat all variables except *Other BC Risk* as categorical; the coefficients for the individual attribute levels are significantly different from one another. For the 'Goal' frame, both *CBC Risk* and *Pain Risk* can be treated as categorical, while for the 'No-Efficacy Difference' frame the results indicate all attributes should be treated as continuous. This adds further weight to the importance of framing to the results produced in the context of choice experiments. In addition, the variability within framing sets suggests that women were evaluating different types of risk (e.g. *CBC Risk* and *Other BC Risk*) differently in deciding how to manage ongoing breast cancer recurrence.

Table 41: Tests of coefficient linearity

	Base	Goal	No-Efficacy Difference
CBC Risk			
Wald	13.82 (p<0.001)	6.31 (p=0.01)	3.25 (p=0.07)
Scaled Wald	13.48 (p<0.001)	6.06 (p=0.01)	1.93 (p=0.17)
Other BC Risk			
Wald	3.66 (p=0.06)	0.95 (p=0.33)	n.a.
Scaled Wald	1.38 (p=0.24)	0.81 (p=0.37)	n.a.
Pain Risk			
Wald	5.81 (p=0.02)	9.32 (p<0.01)	0.04 (p=0.84)
Scaled Wald	7.71 (p<0.01)	6.40 (p=0.01)	0.00 (p=0.95)
Sensitivity Risk			
Wald	4.57 (p=0.03)	0.29 (p=0.59)	0.07 (p=0.80)
Scaled Wald	11.18 (p<0.001)	0.22 (p=0.64)	0.09 (p=0.76)

Notes: The Wald test was estimated using the following linear combinations:
1. For CBC Risk: CBC 2%:5% - CBC 1%:5% - Efficacy 0.3%:5% = 0
2. For Other BC Risk: Other 15%:20% - Other 10%:20%
3. For Pain Risk: Pain 30%:40% - Pain 20%:40% - Pain 10%:40% = 0

4. For Sensitivity Risk: Sensitivity 30%:60% - Sensitivity 40%:60% - Sensitivity 50%:60% = 0
The Scaled Wald test was estimated using the following linear combinations:
1. For CBC Risk: (CBC 2%:5%)*(4.97/3) – (CBC 1%:5%)*(4.97/4) - Efficacy 0.3%:5% = 0
2. For Other BC Risk: (Other 15%:20%)*2 - Other 10%:20%
3. For Pain Risk: (Pain 30%:40%)*3 – (Pain 20%:40%)*1.5 - Pain 10%:40% = 0
4. For Sensitivity Risk: Sensitivity 30%:60% - (Sensitivity 40%:60%)*1.5 – (Sensitivity 50%:60%)*3 = 0

Abbreviation: BC, breast cancer; CBC, contralateral breast cancer; n.a., not applicable.

6.3.3.2.3 Valuing determinants of choice

Observing the differences across framing sets suggested by the LR tests, the mWTP has been estimated using the coefficient estimates from the mixed logit regressions (Table 39) for all framing sets. Results are presented in Table 42 for those attributes with a statistically significant choice coefficient in at least one framing set. There is a difference across framing sets in terms of the attribute with the highest mWTP. For the 'Base' and 'No-Efficacy Difference' frames, this was given by the ASC – *Monitoring:CPM* (although the mWTP value bounds zero for the 'No-Efficacy Difference' frame). In the 'Goal' frame, women were prepared to pay the most for always being involved in decisions about their care (again the value spans zero). In general, the meta-health effects associated with monitoring type and with being involved in decision-making had large relative effect sizes (as proportions of the highest mWTP in each frame) compared with the health effects. The exception was for *CBC Risk* in the 'Base' frame (ranked third at 8.5%) and the 'No-Efficacy Difference' frame (ranked second at 23.0%). Within both those frames, the mWTP for the remaining health effects were then ranked below those of the meta-health effects. For the 'Goal' frame, the meta-health effects were all ranked above the health effects. This further supports a difference across the framing sets in the results observed, both in terms of the estimated mWTP and in their interpretation for the importance placed on the relative product attributes e.g. health effects and meta-health effects. The latter is in some ways confounded by the strong result for the ASC, since it is not known what unobserved characteristics inherent in the label motivated women's preferences for one over the other.

Table 42: Estimated mWTP – framing effects

	Base		Goal		No-Efficacy Difference	
	mWTP, $ (95% CI)	**Wgt (rnk)**	**mWTP, $ (95% CI)**	**Wgt (rnk)**	**mWTP, $ (95% CI)**	**Wgt (rnk)**
Health Effects						
CBC Risk	-79.04 (-196.9 to 38.82)	0.08 (3)	3.73 (-361.44 to 368.91)	0.01 (7)	-649.69 (-1314.9 to 15.52)	0.23 (2)
Other BC Risk	-50.94 (-90.79 to -11.09)	0.05 (5)	-52.55 (-151.16 to 46.06)	0.17 (5)	0 (0-0)	
Pain Risk	-10.68 (-25.14 to 3.79)	0.01 (7)	-43.85 (-101.83 to 14.12)	0.14 (6)	-8.79 (-53.68 to 36.1)	0 (5)
Sensitivity Risk	-20.21 (-38.38 to -2.03)	0.02 (6)	-98.54 (-224.03 to 26.95)	0.32 (4)	-22.05 (-65.31 to 21.2)	0.01 (4)
Meta-Health Effects						
Ultrasound: *Mammogram*	170.79 (22.58 to 319.01)	0.18 (2)	284.85 (-155.9 to 725.61)	0.93 (2)	403.44 (-124.75 to 931.63)	0.14 (4)
Involved Always: *Rarely*	58.94 (-70.96 to 188.83)	0.06 (4)	307.46 (-77.36 to 692.28)	1 (1)	480.6 (-131.35 to 1,092.56)	0.17 (3)
General Preference						
Monitoring: *CPM*	931.52 (398.53 to 1,464.5)	1 (1)	112.93 (-938.99 to 1,164.85)	0.37 (3)	2,824.21 (-317.98 to 5,966.4)	1 (1)

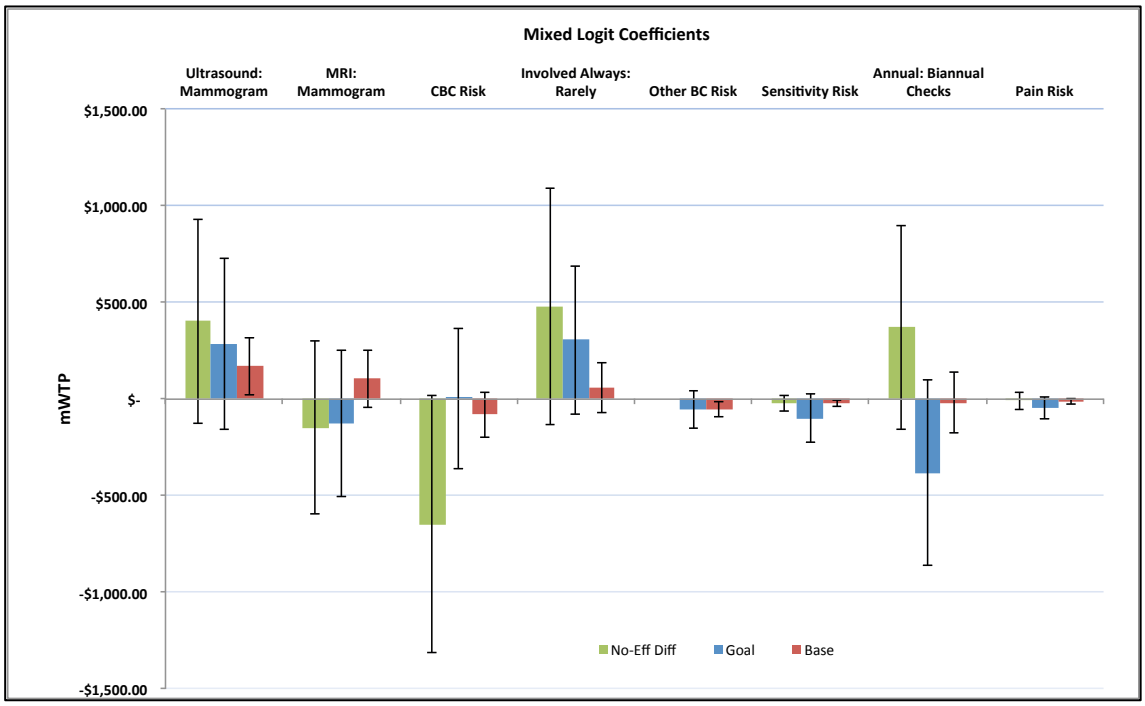
Note: All coefficient ranks are based on the absolute values of the mWTP compared with that of the largest mWTP of those shown for each frame.

Abbreviations: BC, breast cancer; CI, confidence interval; CBC, contralateral breast cancer; CPM, contralateral prophylactic mastectomy; mWTP, marginal willingness to pay; rnk, rank; Wgt, weight.

The ranking of the attributes in terms of their relative value, as well as differences in those rankings across the framing sets, can be observed from the graphical presentation of the mWTP in Figure 47. While the point estimates varied across the three framing sets, it can be observed that there is significant overlap in terms of their associated 95% confidence intervals. Overall these results show that the meta-health effects are ranked higher than the health effects; taking into account that the ASC value is not shown (which dominates all other effects in both the 'Base' and 'No-Efficacy Difference' frames). In addition, the health effects have been treated as continuous in the analysis used for these estimates, yet for the 'Base' frame at least there is evidence that they could have been analysed as categorical. Evidence from Chapter 5 suggests that this would have resulted in higher mWTP values for those attributes within that frame. However, since the health effects were not uniformly non-linear across the

framing sets, an analysis of mWTP using categorically coded health effects for the 'Base' frame alone has not been pursued.

Figure 47: mWTP – main results



Note: Estimates formed by dividing the coefficient for each attribute by the coefficient on *Monitoring OOP,* invoking the STATA *wtp* command. All confidence intervals produced using the Delta method.

Abbreviations: BC, breast cancer; CI, confidence interval; CBC, contralateral breast cancer; MRI, magnetic resonance imaging; mWTP, marginal willingness to pay; OOP, out-of-pocket.

6.3.4 Understanding the drivers of choice: predicated probabilities and class segmentation

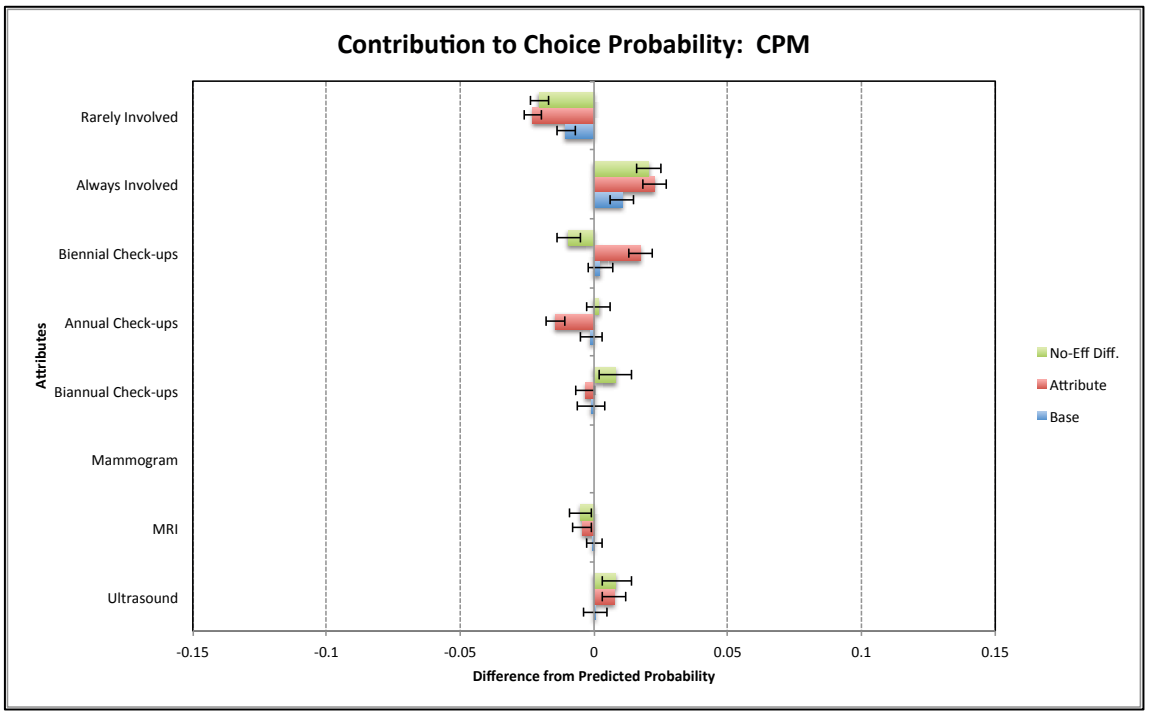
6.3.4.1 Attribute levels and predicted probabilities

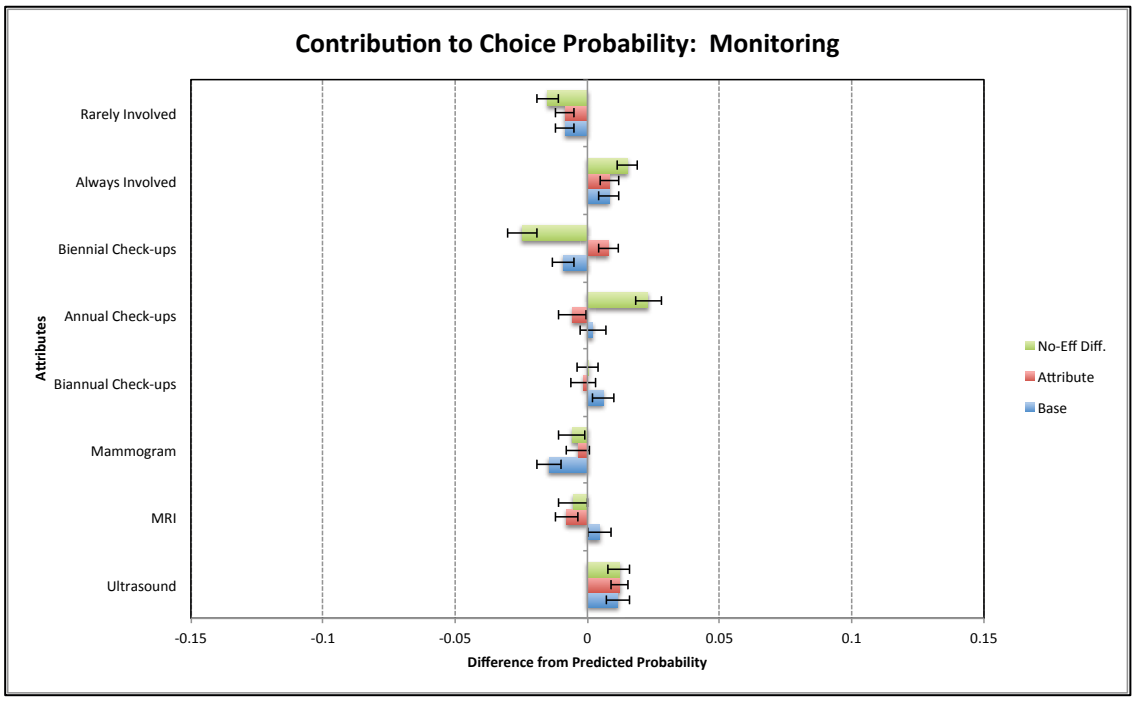
The predicted probabilities of choice from the mixed logit regression are summarised in a descriptive analysis across the choice sets for each of the respective attribute levels. These analyses utilise the mean of the predicted probabilities for each option (CPM Only or Monitoring Only), restricting the choices to those when each respective attribute level appears. The difference between that mean probability, and the overall mean predicted probability per frame, is interpreted as being the influence on the

choice probability of each attribute level, given the occurrence of all the other attribute levels in the design. All predicted probabilities were generated by invoking the STATA command *mixlpred* (1,000 replications).

The impact of each attribute on the probability of choosing either CPM or Monitoring is depicted graphically in Figure 48 to Figure 50. *CBC Risk* was the most influential health effect for both choice options, noting that *Pain Risk* and *Sensitivity Risk* did not appear for the Monitoring Only option. Probabilities behaved in the manner that would be expected relative to costs, with larger changes in the probability of choice as the cost increased, being greatest for *Surgical OOP* for CPM Only. Overall, the influence of meta-health effects varied across the framing sets, but was reasonably consistent between the choice options. That is, being involved in decision-making was generally influential across frames, although to a slightly greater extent for the CPM option. Similar results were observed for the type of monitoring. The influence of the frequency of monitoring differed widely between the framing sets and the choice options.

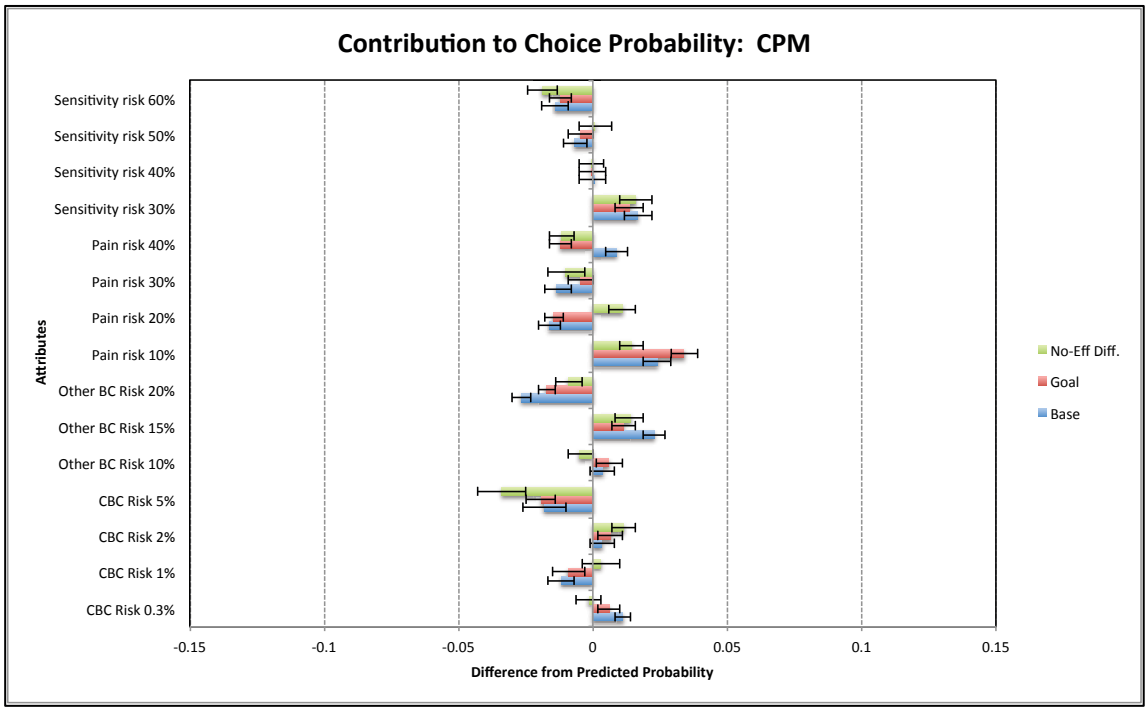
Figure 48: Impact on choice probability – meta-health effects

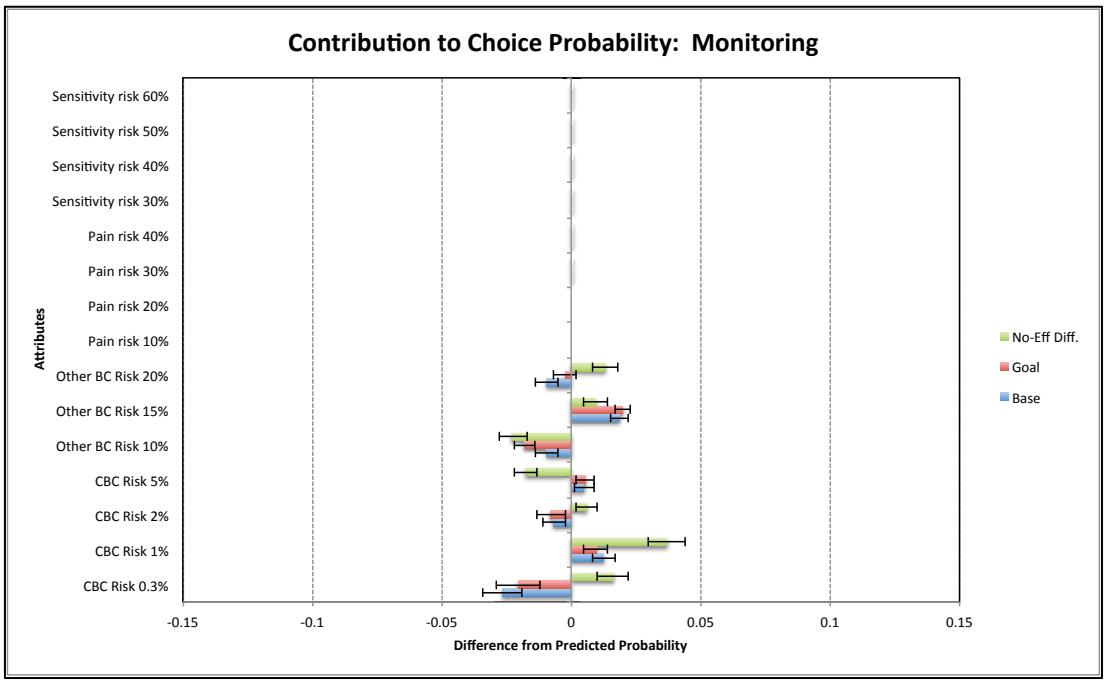




Abbreviation: CPM, contralateral prophylactic mastectomy; MRI, magnetic resonance imaging.

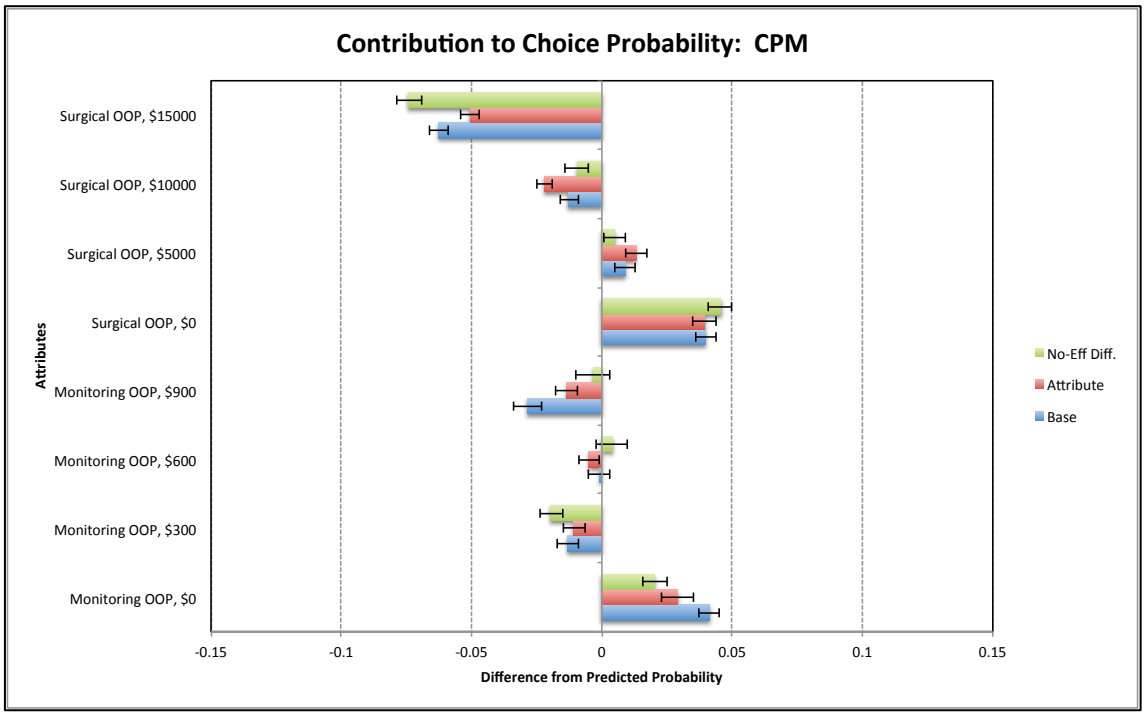
Figure 49: Impact on choice probability – health effects

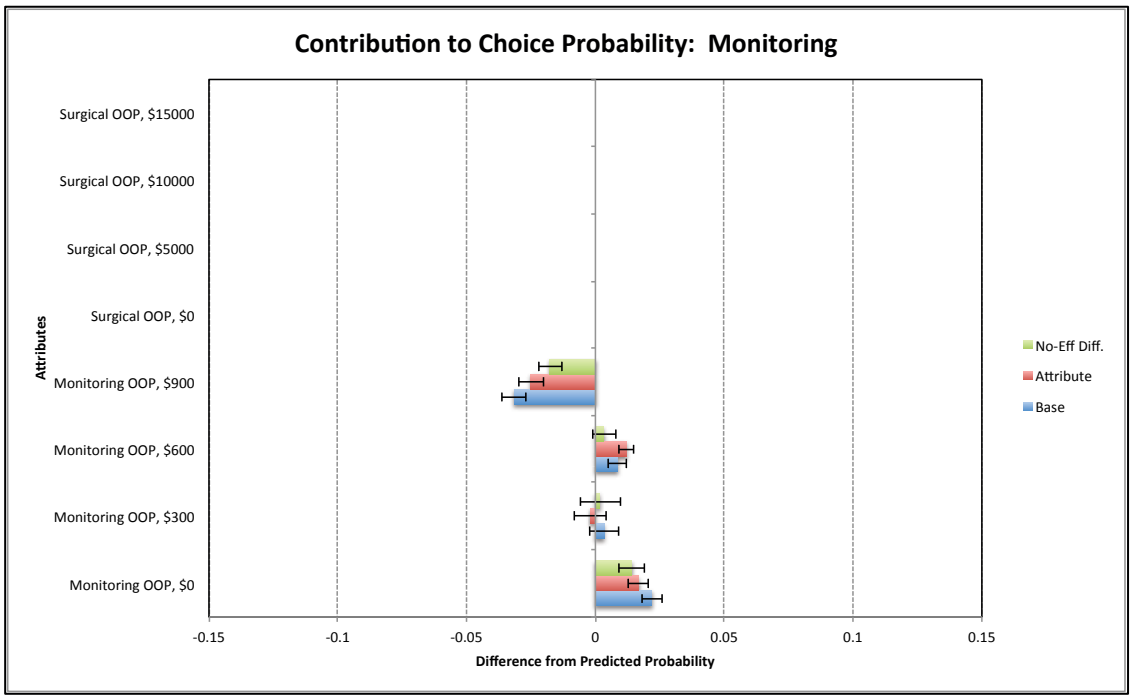




Abbreviation: BC, breast cancer; CBC, contralateral breast cancer; CPM, contralateral prophylactic mastectomy.

Figure 50: Impact on choice probability – costs





Abbreviation: CPM, contralateral prophylactic mastectomy; OOP, out-of-pocket.

6.3.4.2 Trading and cancer concern: subgroup analyses

The mixed logit regression was re-estimated using the pooled dataset to investigate potential subgroup effects in women who exhibited trading behaviour, and according to whether women could be classified as cancer concerned or not. The results are presented in Table 43.

Table 43: Respondent heterogeneity - responses by influencers and trading

	All Responses	Traders	Not Cancer Concerned	Cancer Concerned
Attribute Means				
CBC Risk	-0.134	-0.179	0.314	-0.203
	(0.052)*	(0.050)**	(0.221)	(0.055)**
Other BC Risk	-0.052	-0.056	0.065	-0.070
	(0.019)**	(0.018)**	(0.075)	(0.020)**
Monitoring OOP, $'00	-0.077	-0.077	-0.288	-0.073
	(0.017)**	(0.017)**	(0.149)	(0.018)**
Surgery OOP, $'00	-0.011	-0.011	-0.041	-0.010
	(0.002)**	(0.002)**	(0.016)*	(0.002)**
Pain Risk	-0.011	-0.010	-0.044	-0.012
	(0.006)	(0.006)	(0.036)	(0.007)
Sensitivity Risk	-0.014	-0.009	-0.148	-0.009
	(0.006)*	(0.007)	(0.066)*	(0.007)
Ultrasound: *Mammogram*	0.191	0.172	0.080	0.203
	(0.057)**	(0.054)**	(0.286)	(0.063)**

	All Responses	Traders	Not Cancer Concerned	Cancer Concerned
MRI: *Mammogram*	-0.013	0.003	-0.008	0.011
	(0.059)	(0.058)	(0.218)	(0.066)
Annual: *Biannual Checks*	-0.023	-0.016	-0.966	0.037
	(0.061)	(0.060)	(0.669)	(0.068)
Biennial: *Biannual Checks*	0.018	0.010	0.417	-0.012
	(0.056)	(0.053)	(0.409)	(0.064)
Involved Always: *Rarely*	0.177	0.172	-0.037	0.206
	(0.055)**	(0.050)**	(0.211)	(0.060)**
Monitoring: *CPM* Constant	1.245	-0.065	2.078	0.949
	(0.220)**	(0.172)	(0.984)*	(0.260)**
Attribute Standard Deviation				
CBC Risk	0.479	0.400	-1.207	0.440
	(0.097)**	(0.076)**	(0.570)*	(0.085)**
Other BC Risk	0.071	0.083	0.024	0.086
	(0.041)	(0.030)**	(0.068)	(0.032)**
Monitoring OOP, $'00	-0.090	0.067	0.268	-0.054
	(0.041)*	(0.037)	(0.098)**	(0.070)
Surgery OOP, $'00	0.012	0.010	0.031	-0.012
	(0.002)**	(0.003)**	(0.014)*	(0.003)**
Pain Risk	0.002	0.021	-0.080	0.025
	(0.012)	(0.020)	(0.037)*	(0.022)
Sensitivity Risk	-0.033	0.020	0.181	0.034
	(0.011)**	(0.015)	(0.074)*	(0.008)**
Ultrasound: *Mammogram*	0.040	-0.057	-0.767	0.013
	(0.047)	(0.097)	(0.535)	(0.087)
MRI: *Mammogram*	-0.308	-0.201	0.668	0.353
	(0.106)**	(0.138)	(0.296)*	(0.200)
Annual: *Biannual Checks*	0.057	-0.018	0.028	-0.020
	(0.187)	(0.051)	(0.280)	(0.169)
Biennial: *Biannual Checks*	0.059	-0.004	0.694	0.146
	(0.076)	(0.057)	(0.439)	(0.122)
Involved Always: *Rarely*	0.313	0.290	0.646	0.360
	(0.112)**	(0.089)**	(0.449)	(0.094)**
Monitoring: *CPM* Constant	2.287	0.328	-4.412	-2.004
	(0.195)**	(0.566)	(1.866)*	(0.246)**
Observations	11,136	4,728	3,120	7,824
Individuals	464	197	130	326
Wald Chi	125.88	79.06	30.08	91.61
d.f.	12	12	12	12
p-value	0.00	0.00	0.00	0.00
Log-likelihood	-1,744.54	-1,336.41	-286.84	-1,406.65
Pseudo R^2	0.058	0.072	0.115	0.060

Notes: * $p<0.05$; ** $p<0.01$
All categorical data are effects coded to minimise the potential for contamination of the base effect in the grand mean.
Base categories are shown in italics.
All analyses were conducted in STATA 12, with robust standard errors (in parentheses) to account for the survey nature of the data.
Analyses conducted using 1,000 replications.

Abbreviations: BC, breast cancer; CBC, contralateral breast cancer; CPM, contralateral prophylactic mastectomy; d.f., degrees of freedom; OOP, out-of-pocket.

6.3.4.2.1 Analysis by trading behaviour

Restricting the analysis to the 197 women who displayed some trading behaviour, the results were largely as observed in the pooled analysis, with some exceptions. The first is that the ASC is no longer significant and now favours CPM Only amongst women who trade. Of the remaining attributes, the signs and magnitude of the coefficients are consistent with those observed in the pooled analysis with all respondents. This is not surprising since the variability observed within that analysis was due to the effect of the choices made by traders. Only one coefficient that was previously significant is now insignificant, *Sensitivity Risk*, which might be due to the change in respondent numbers.

6.3.4.2.2 Analysis by cancer concern

Larger differences in the determinants of choice are observed by analysing the data according to whether women could be classified as cancer concerned. Both types of cancer risk were statistically significant influences on choice for women who were cancer concerned; they were less likely to choose a management option associated with higher recurrence risk. However, cancer risk was not a significant determinant of choice for women who were not cancer concerned. Differences also arose with respect to the characteristics of ongoing monitoring and involvement in decision-making; women in the cancer concerned group were deterred by higher monitoring costs but in favour of less invasive methods of monitoring, and of being involved in decision-making. None of these attributes significantly influenced the choices of women who were not cancer concerned. Choices by both groups were negatively influenced by increasing *Surgical OOP*, and those by women who were not cancer concerned were also negatively influenced by the risk of losing breast sensitivity. The ASC favouring monitoring was statistically significant and positive for both groups, although twice as large for the not cancer concerned group as for the cancer concerned group.

6.3.4.3 Understanding subgroup structures – latent classes

6.3.4.3.1 Forming the latent classes

The three groups included in the latent class analysis are described as those who Prefer CPM, Prefer Monitoring and Traders. These group descriptions were derived by comparing the posterior predicted probability of class membership with the choices made by women. The groups were formed based on the frequency with which women deviated from either always choosing CPM Only or Monitoring Only over the 12 choice scenarios. Cross-tabs were formed between the posterior predicted probability of class membership and choice counts for CPM Only and Monitoring Only. The results indicated that 56 (class share 12%) women belonged to the Prefer CPM group. Of those, 69% only ever chose CPM Only, a further 16% made one choice out of 12 for Monitoring Only and the remainder at most chose routine monitoring in three instances. Respondents in the Prefer Monitoring group were more consistently non-traders; 274 (class share 59%) women were in this group, with the majority 83% always choosing Monitoring Only, a further 15% making only one choice for CPM Only. The remaining 134 (class share 29%) women were in the Traders group, with a slight preference for choosing Monitoring Only more frequently than CPM Only. Posterior predicted probabilities of class membership were produced by invoking the *lclogitpr* command in Stata.

6.3.4.3.2 Results of latent class analysis

The results for the latent class analysis are provided in Table 44. For the Prefer CPM group, the ASC was negative (and of borderline significance, p=0.073) showing a preference for CPM Only over Monitoring Only. No other attributes significantly influenced choice. The Traders group was not influenced by the ASC (p=0.291), but was influenced by a number of attributes in terms of choice. Women were negatively influenced in their choices by increasing breast cancer risk and *Pain Risk*, higher *Monitoring* and *Surgical OOP*, but were favourably disposed to options in which they were always involved in decisions about their care. Finally, the Prefer Monitoring

group showed a strong preference for monitoring as evidenced by the positive and significant ASC ($p<0.01$). Despite these strong preferences for monitoring, some women in this group were deterred by rising *Monitoring OOP* and *Surgical OOP* (which applied to CPM only).

Table 44: Understanding subgroups through latent class analysis

	Prefer CPM	Traders	Prefer Monitoring
CBC Risk	0.119	-0.216	-0.031
	(0.186)	(0.036)**	(0.132)
Other BC Risk	-0.020	-0.036	-0.015
	(0.057)	(0.015)*	(0.043)
Monitoring OOP, $'00	0.005	-0.049	-0.094
	(0.052)	(0.013)**	(0.044)*
Surgery OOP, $'00	-0.005	-0.007	-0.013
	(0.005)	(0.001)**	(0.005)**
Pain Risk	0.031	-0.012	0.002
	(0.024)	(0.005)*	(0.016)
Sensitivity Risk	0.001	-0.009	-0.009
	(0.025)	(0.006)	(0.020)
Ultrasound: *Mammogram*	0.242	0.080	0.232
	(0.215)	(0.055)	(0.157)
MRI: *Mammogram*	0.094	0.041	-0.112
	(0.220)	(0.055)	(0.162)
Annual: *Biannual Checks*	0.255	-0.067	-0.135
	(0.245)	(0.058)	(0.190)
Biennial: *Biannual Checks*	-0.040	0.058	0.078
	(0.218)	(0.051)	(0.145)
Involved Always: *Rarely*	0.106	0.132	-0.039
	(0.152)	(0.040)**	(0.118)
Monitoring: *CPM* Constant	-1.415	-0.182	1.659
	(0.789)	(0.173)	(0.435)**
Class Shares	0.121	0.285	0.594
Membership Share Contributions (Relative to Group – Prefer Monitoring)			
Age 16: *25*	-1.289	1.513	
	(0.000)	(0.463)**	
Age 45: *25*	1.402	-0.173	
	(0.219)**	(0.286)	
Age 65: *25*	0.731	-0.386	
	(0.297)*	(0.344)	
Age 75: *25*	-2.043	-1.201	
	(0.000)	(0.751)	
Income $149K: *$39K*	-0.334	-0.267	
	(0.353)	(0.233)	
Income $79K: *$39K*	0.448	-0.003	
	(0.300)	(0.230)	
Income Over $150K: *$39K*	-0.433	0.577	
	(0.674)	(0.340)	
Income Unknown: *$39,000*	0.206	-0.445	

	Prefer CPM	Traders	Prefer Monitoring
	(0.366)	(0.282)	
University: *School*	0.058	0.175	
	(0.227)	(0.174)	
Vocational: *School*	-0.287	-0.163	
	(0.236)	(0.176)	
Cancer Concern: *Not Concerned*	0.681	0.799	
	(0.212)**	(0.159)**	
Efficacy Frame: *Base*	0.155	0.197	
	(0.219)	(0.167)	
'Goal' frame: *Base*	-0.109	-0.114	
	(0.229)	(0.174)	
Cancer Screening > 2 yr: *Never Screened*	-0.254	0.199	
	(0.264)	(0.183)	
Cancer Screening <2 yr: *Never Screened*	0.027	-0.212	
	(0.220)	(0.170)	
Constant	-3.152	-1.071	
	(0.283)**	(0.258)**	
Total Observations	11,136		
Total N	464		
Log-Likelihood	-1,742.73		
AIC	3617.46		
BIC	4100.44		
Pseudo R^2	0.047		

Notes: * $p<0.05$; ** $p<0.01$
All categorical data are effects coded to minimise the potential for contamination of the base effect in the grand mean.
Base categories are shown in italics.
All analyses were conducted in STATA 12, with robust standard errors (in parentheses) to account for the survey nature of the data.
Analyses conducted using 1,000 replications.
Results for the membership share contributions are equivalent to the estimation of a separate multi-nomial logit regression in which trading status (Prefer CPM, Prefer Monitoring, Traders) is the dependent variable, regressed on the same demographic characteristics included in class membership status.

Abbreviations: BC denotes breast cancer; CBC contralateral breast cancer; CPM contralateral prophylactic mastectomy; d.f. degrees of freedom; OOP, out-of-pocket.

6.3.4.3.3 Determinants of class membership

The latent class analyses were estimated allowing for the influence of demographic factors in determining membership to the resulting class shares. This not only adjusts for the effect of those variables in determining class shares but also in estimating the individual choice coefficients. There are two results of note with respect to the determinants of class membership; the first is that the negative and significant constant on the analysis of the share determinants for the Prefer CPM and Traders groups

(relative to Prefer Monitoring) suggests that by and large women are less likely to be in either of those two groups than the Prefer Monitoring group. However, younger women (16-25 years) are more likely than those aged 25-45 years to be Traders, while those who are aged 45-65 years are more likely than 25-45 year olds to be in the Prefer CPM group, relative to those in the Prefer Monitoring group. There is consistency between the groups in that women in both groups are more likely to be cancer concerned than not cancer concerned compared with those in the Prefer Monitoring group.

Further evidence of the composition of each of the groups is provided in Table 45. These data confirm the results of the analysis of latent class membership. In addition to the differences noted with respect to age and the influence of cancer concern, it can be noted that those in Prefer CPM group differ with respect to educational attainment (a higher proportion of women with school education only) while a higher proportion of Traders has a university degree. Differences are also noted with respect to prior screening behaviour, with a higher proportion of women in the Prefer CPM group having attended for cancer screening within the last two years compared with the other groups. While there appear to be some differences in the distribution of income across groups, this is confounded by the relatively high proportion of women for whom income was unknown in the Prefer CPM and Prefer Monitoring groups. Finally, class membership was not affected by the framing set to which women were randomised.

Table 45: Demographics of class membership

	Prefer CPM, n=56 n (%)	**Traders**, n=134 n (%)	**Prefer Monitoring,** n=274 n (%)
Age			
16-24	0 (0)	12 (8.96)	7 (2.55)
25-44	19 (33.93)	57 (42.54)	92 (33.58)
45-64	27 (48.21)	44 (32.84)	104 (37.96)
65-74	9 (16.07)	18 (13.43)	54 (19.71)
75 or Over	0 (0)	1 (0.75)	12 (4.38)
Unknown	1 (1.79)	2 (1.49)	5 (1.82)

	Prefer CPM, n=56 n (%)	**Traders**, n=134 n (%)	**Prefer Monitoring,** n=274 n (%)
		p<0.001	
Income			
Under $39,999	15 (26.79)	40 (29.85)	73 (26.64)
$40,000-$79,999	19 (33.93)	34 (25.37)	63 (22.99)
$80,000-$149,999	10 (17.86)	29 (21.64)	67 (24.45)
Over $150,000	2 (3.57)	15 (11.19)	17 (6.2)
Unknown	10 (17.86)	16 (11.94)	54 (19.71)
		p<0.001	
Education			
School Only	20 (35.71)	36 (26.87)	80 (29.2)
University	18 (32.14)	47 (35.07)	76 (27.74)
Vocational	17 (30.36)	49 (36.57)	112 (40.88)
Unknown	1 (1.79)	2 (1.49)	6 (2.19)
		p<0.001	
Cancer Concern			
Not Concerned	8 (14.55)	16 (12.03)	106 (39.55)
Concerned	47 (85.45)	117 (87.97)	162 (60.45)
		p<0.001	
Cancer Screening			
Never Screened	14 (25)	36 (26.87)	71 (25.91)
Screened >2 years	12 (21.43)	41 (30.6)	68 (24.82)
Screened < 2 years	30 (53.57)	57 (42.54)	135 (49.27)
		p<0.001	
Frame			
Base	19 (33.93)	43 (32.09)	96 (35.04)
Efficacy	20 (35.71)	52 (38.81)	89 (32.48)
Goal	17 (30.36)	39 (29.1)	89 (32.48)
		p=0.86	

Notes: p values shown are for the significance of the Bartlett's test of equality of variance in an oneway analysis of variance test.

Abbreviations: CPM, contralateral prophylactic mastectomy.

6.3.4.4 Attitudinal differences and mWTP

6.3.4.4.1 mWTP by trading behaviour

The results from the mixed logit regressions examining choice across the trader and cancer concern subgroups were used to estimate mWTP for the various attributes across these groups. The results presented in Table 46 indicate that the mWTP does

not differ between the trader and pooled analyses with the exception of that for *Monitoring: CPM*. For the pooled analysis, this was associated with a high and positive mWTP; women are prepared to pay $1,617 for Monitoring Only rather than CPM Only. In contrast, for Traders, the mWTP is negative, implying that women would need to be paid to shift from CPM to monitoring, but the confidence interval spans zero (but does not overlap with the confidence interval of the pooled group). The results for the mWTP for all attributes (excluding the ASC), regardless of significance, are shown in Figure 51. These results confirm the consistency between the groups, and are to be expected since the variation in choice observed in the pooled analysis, apart from the ASC, is due to traders, hence they will drive the resulting mWTP in both the pooled and trader subgroup analyses.

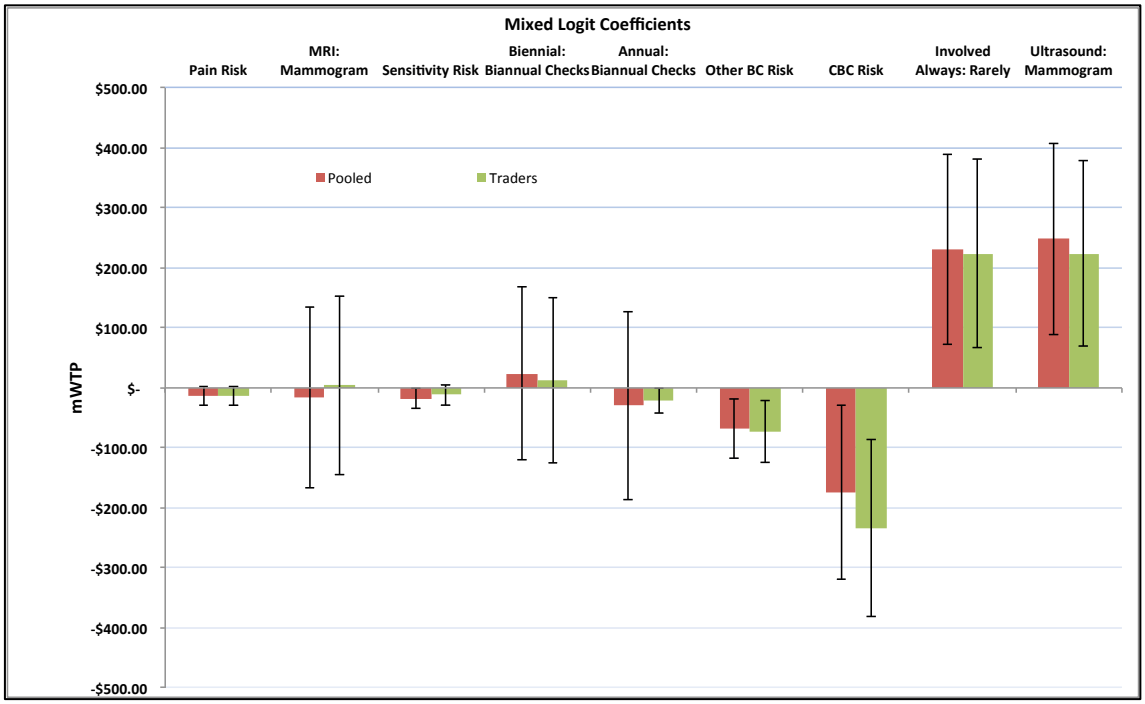
Table 46: Impact of trading on mWTP

	Pooled		**Traders**	
	mWTP, $ (95% CI)	**Wgt (rnk)**	**mWTP, $ (95% CI)**	**Wgt (rnk)**
Health Effects				
CBC Risk	-174.36 (-318.77 to -29.95)	0.11 (4)	-234.03 (-380.87 to -87.18)	1 (1)
Other BC Risk	-67.68 (-117.31 to -18.05)	0.04 (5)	-72.59 (-124.07 to -21.1)	0.31 (5)
Sensitivity Risk	-17.83 (-34.18 to -1.48)	0.01 (6)	-11.84 (-28.65 to 4.98)	0.05 (6)
Meta-Health Effects				
Ultrasound: *Mammogram*	248.02 (88.92 to 407.12)	0.15 (2)	224.04 (69.1 to 378.98)	0.96 (3)
Involved Always: *Rarely*	230.44 (72.68 to 388.2)	0.14 (3)	224.32 (67.15 to 381.49)	0.96 (2)
General Preference				
Monitoring: *CPM*	1616.98 (737.34 to 2496.62)	1 (1)	-84.29 (-526.38 to 357.8)	0.36 (4)

Note: All coefficient ranks are based on the absolute values of the mWTP compared with that of the largest mWTP for each frame.

Abbreviations: BC, breast cancer; CI, confidence interval; CBC, contralateral breast cancer; CPM, contralateral prophylactic mastectomy; mWTP, marginal willingness to pay; rnk, rank; Wgt, weight.

Figure 51: mWTP – impact of trading



Abbreviation: BC, breast cancer; CBC, contralateral breast cancer; mWTP, marginal willingness to pay; MRI, magnetic resonance imaging.

6.3.4.4.2 mWTP by cancer concern

The results of the mWTP for the cancer concern subgroups presented in Table 47 are reversed from those of the trader subgroup analysis; there are differences in the value expressed by the mWTP for the attributes other than the ASC between the subgroups. For example, women who are cancer concerned have a negative mWTP for both cancer risks (implying they would need to be paid to choose a management option that has a higher risk of cancer recurrence); while those who are not cancer concerned have a positive mWTP, for which the confidence intervals span zero. These confidence intervals do not overlap between subgroups for either cancer risk. This is an important finding in terms of being able to discern a value for a reduction in the fear of cancer recurrence.

The same outcome was observed for whether women are involved in decision-making; those who are cancer concerned place a positive mWTP on always being involved

while those who are not cancer concerned have a negative, and statistically different, mWTP for this attribute. Importantly for women in the cancer concerned group, the meta-health effect of being involved in decision-making is given the second highest value relative to the ASC; for those who are not cancer concerned, health effects are given more weight than meta-health effects. For both groups, the ASC was given the highest mWTP, and this did not differ between the cancer concern and not cancer concerned groups. The overall differences between the groups are apparent in the graphical presentation of the mWTP in Figure 52.

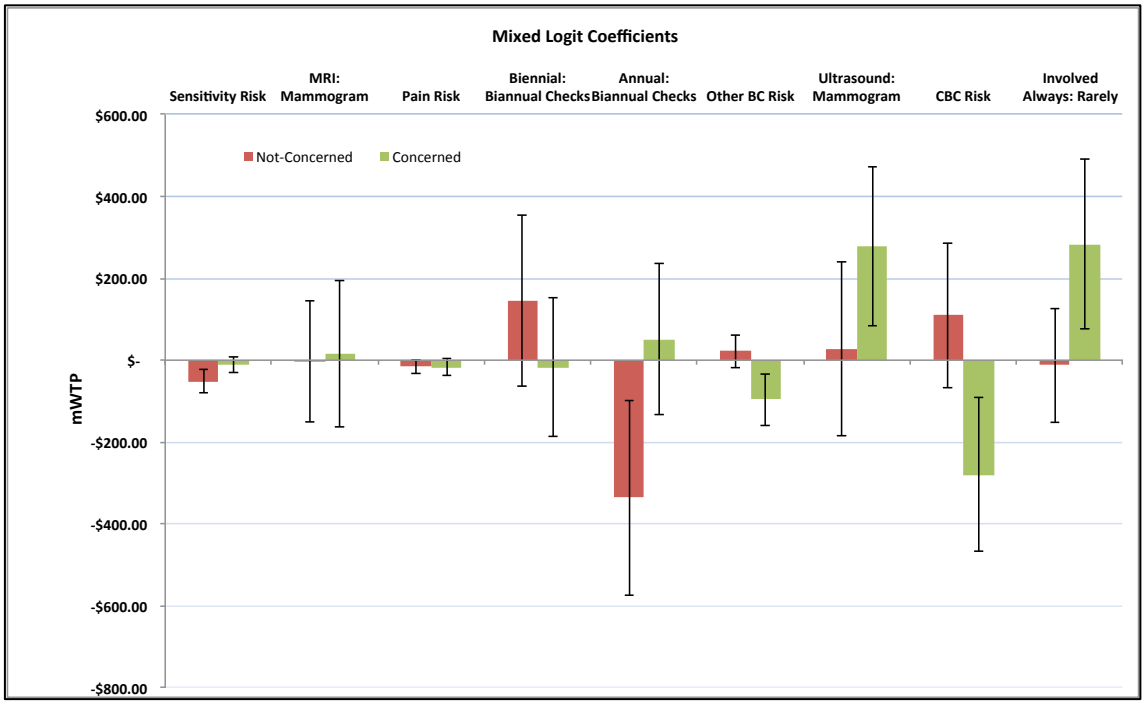
Table 47: Impact of cancer concern on mWTP

	Cancer Concern		Not Cancer Concerned	
	mWTP, $ (95% CI)	**Wgt (rnk)**	**mWTP, $ (95% CI)**	**Wgt (rnk)**
Health Effects				
CBC Risk	-279.26 (-466.17 to -92.35)	0.21 (3)	108.95 (-67.12 to 285.03)	0.15 (2)
Other BC Risk	-96.37 (-159.48 to -33.26)	0.07 (5)	22.44 (-18.19 to 63.08)	0.03 (5)
Sensitivity Risk	-12.17 (-31.17 to 6.82)	0.01 (6)	-51.38 (-79.92 to -22.84)	0.07 (3)
Meta-Health Effects				
Ultrasound: *Mammogram*	278.56 (83.91 to 473.2)	0.21 (4)	27.93 (-184.36 to 240.21)	0.04 (4)
Involved Always: *Rarely*	283.54 (76.54 to 490.54)	0.22 (2)	-12.84 (-151.71 to 126.03)	0.02 (6)
General Preference				
Monitoring: *CPM*	1,304.80 (452.69 to 2,156.91)	1 (1)	721.48 (102.15 to 1,340.80)	1 (1)

Note: All coefficient ranks are based on the absolute values of the mWTP compared with that of the largest mWTP for each frame.

Abbreviations: BC, breast cancer; CI, confidence interval; CBC, contralateral breast cancer; CPM, contralateral prophylactic mastectomy; mWTP, marginal willingness to pay; rnk, rank; Wgt, weight.

Figure 52: mWTP – impact of cancer concern



Abbreviation: BC, breast cancer; CBC, contralateral breast cancer; mWTP, marginal willingness to pay; MRI, magnetic resonance imaging.

6.4 Discussion

There is often much discussion in the popular and academic press of the importance of early cancer detection and treatment as a means of reducing the longer-term effects of breast cancer. This is no less the case for women who have already experienced breast cancer; how women choose to manage their ongoing risk of recurrence has implications not only for the potential breast cancer recurrence risk, but also for the ongoing experience of care. This was evident in the qualitative research underpinning the analysis in this chapter. Women were motivated by both health and meta-health effects in how they chose to manage their ongoing risks of cancer recurrence. For some women, those choices centred on reducing the likelihood of future cancer recurrences, for others it was focussed on a loss of trust in the standard methods of monitoring, or avoidance of the intrusiveness of ongoing monitoring and negative effects of surgical intervention. These findings were consistent with the results of previous research

regarding the factors motivating decisions by women when thinking about CPM.[284,285,327]

What was also apparent from those initial discussions with women is that they express strong preferences for one form of ongoing management over another. This was observed in the quantitative DCE analyses in that over half of the women who responded to the pilot and final surveys respectively always chose one form of management over the other, typically routine monitoring. It is likely that it reflects both the use of a labelled design and the existence of strong preferences[328] regarding the topic women were asked to consider: prophylactic mastectomy to reduce the risk of a cancer recurrence. Testing that this is indeed the case might require further surveys incorporating a similar design, using a larger sample potentially, and with follow-up qualitative work with women who are non-traders.

Where women were persuaded to trade between management options this was most often motivated by a desire to avoid higher risks of breast cancer, or in favour of options in which they are involved in decision-making. There was a strong differentiator between women in terms of those who would trade and those who would not: the extent to which they were concerned about cancer recurrence when answering the choice questions. Women who could be classified as concerned about cancer recurrence were more likely to trade, and generally were more likely to prefer the CPM option over monitoring.

In the current research into the management of ongoing breast cancer risks, the health effects were by and large found to dominate the meta-health effects, with some exceptions depending on the type and manner of information presented. That is, where the risk of other breast cancer recurrence did not differ between the management options, both the type of monitoring and involvement in decision-making influenced choices. This was not observed for the other information conditions. However, these meta-health effects were more important to women who were

concerned about cancer recurrence compared with those who were not. This suggests that the choices these women made were influenced by the experience of care as well as the outcomes of that care.

It is possible that the dominance of health effects was a result of respondents adopting simplifying heuristics in their decision-making. Each choice option contained nine attributes, which might have led some respondents to focus only on those attributes they considered important.[xxxvi] Constructing DCE designs with attributes in overlapping subsets has been used to test the impact of up to 30 attributes in a choice context.[303] This was not used due to its potential implications for the required sample size, particularly in light of objective in this research to also test framing effects. Evidence of the potential impact of the number of attributes on the results observed was available from the pilot DCE (see Appendix 13). These results indicate that 28% of respondents adopted such a heuristic when faced with choice options varying across 10 attributes. A further 13% of respondents reported their were too many attributes, but that they did not affect how they responded, while the remaining respondents felt the number of attributes was appropriate. Taking this into account, and that three attributes applied to CPM option only (*Surgical OOP, Sensitivity Risk, Pain Risk)*, it is unlikely that the number of attributes adversely affected the results observed.

DCEs have been used elsewhere to understand women's preferences for care with respect to the experience of care in breast cancer and its outcomes. Damen et al. (2011)[299] assessed women's preferences regarding breast reconstruction, and found the use of autologous material (as opposed to an implant), and an excellent aesthetic outcome were the key drivers of utility; while complication rates (short or long) reduced utility. In a DCE among 331 women who had undergone treatment for breast cancer patients, Kimman et al. (2010)[224] examined preferences for alternative follow-up programmes described by varying attendance at an educational programme, the

[xxxvi] Reviews of the literature indicate that a range of four to eight attributes is common in DCE studies, but studies have been conducted with up to 30 attributes.[44,55,303]

frequency of visits, the waiting time at each appointment, the mode of contact for the follow-up, and the health care provider. Health care provider and the mode of contact were the most important characteristics, while waiting time was the least important in affecting the choice of follow-up method.[224] Similarly, Bessen et al. (2014)[225] found that the choice of breast cancer follow-up services among 722 Australian women was highly influenced by the provider of that care, with a preference for specialist led over general practitioner led care. Osoba et al. (2006)[300] used a DCE (159 women) to demonstrate that the impact of the disease and its care on physical functioning had the greatest influence on women when choosing between hypothetical treatments for breast cancer. Finally, in a DCE of 87 women, Gerard et al. (2003)[130] conclude that while screening accuracy was the most influential determinant of women's choices, they were also influenced by factors such as the time involved for screening, distance to travel to be screened, the information provided about screening and the privacy available to women. These factors constitute what is considered in the current research to be meta-health effects, but they were not given a monetary value (the experiment by Gerard et al. (2003) did not include a cost attribute) or a duration based QoL value (consistent with estimation of a QALY).

Within this chapter the preferences of interest were those women expressed for the meta-health effects compared with health effects. The information provided to women was varied in an attempt to evaluate how this influenced the resulting preferences, both in terms of the direct comparison of attributes (through the exclusion of *Other BC Risk*) and in the background information provided (in the 'Goal' frame). Drawing on Levin et al.'s (1998)[89] taxonomy of framing (risk, attribute and goal), the risk and 'Goal' frame were tested: the former by removing in one frame ('No-Efficacy Difference') the likelihood of a gain or loss in efficacy associated with choosing one option over the other; and the latter by providing more information about the process associated with a diagnosis of breast cancer, and thereby allowing for the possibility for this to influence the perceived trade-off between attributes.

Overall, the results for the evidence of framing effects were somewhat mixed. Comparisons between the framing sets suggest there is evidence of a framing effect. For example, the strong preference over the labelled options observed in both the 'Base' and 'No-Efficacy Difference' frames is not present for the 'Goal' frame. This suggests that in the case of the latter, providing women with more information about the process and background of their breast cancer diagnosis influenced the value they placed on the observed relative to the unobserved attributes of the (labelled) choice options. The use of such background information in affecting preferences reflects the use of comprehensive accounts (a form of decision heuristic) in which individuals consider not only the comparison of attributes of choices, but other external information pertaining to that choice.[88] That women in the 'Goal' frame did not demonstrate a preference for monitoring over CPM is consistent with the feedback obtained during the qualitative research. From that research it was observed that one motivating factor for women who had decided to undergo a CPM was that they no longer trusted routine monitoring because it had not detected their cancer. The health scenario in the 'Goal' frame potentially described such a situation; a breast cancer had been detected despite a recent mammogram that was described as clear. It is possible that some women responding to the survey interpreted this as a sign that the initial mammogram had failed to detect the cancer (even though those words were not used).

However, the evidence for framing effects was not consistently supported; the tests for the presence of such effects were equivocal. While framing effects were evident for both health effects and meta-health effects when comparing between framing sets, they were not present when assessing framing across all three sets. This is likely due to the degree of non-trading among women responding to this survey. Testing such an effect would have required restricting the analysis of framing to traders alone, and then comparing across framing sets. However, the reduction in sample size to traders within framing sets required to conduct such an analysis meant that this was not possible. It is possible that the impact of non-trading on the analysis of framing effects might have been captured by respecifying the latent class analyses to control for scale

effects.[240] This could be explored in future analyses of these data. Nonetheless, the understanding that the results could be assessed on a pooled basis without necessarily being confounded by framing allowed the exploration of differences in the factors influencing women who were cancer concerned compared with those who are not cancer concerned.

Within the current research project, women who were concerned about cancer recurrence had a higher mWTP to avoid a higher risk of cancer than those who were not. This is an important finding in the context of this research if this difference is interpreted as being due to a fear of recurrence; that women who express a concern for cancer recurrence would also exhibit a fear of recurrence. Such an interpretation is reasonable given that the terms 'fear of recurrence' and 'concern about recurrence' are used to measure the same psychological constructs in women with cancer.[319] It is also possible that the difference in the mWTP for avoiding additional cancer risk as observed for the subgroups of women based on cancer concern reflects differences in risk aversion. That is, if women who are cancer concerned are perceived as risk averse, the higher and significant mWTP might reflect loss aversion. The results suggest that women responded differently depending on the source of that risk: the different types of risk (e.g. *CBC Risk* and *Other BC Risk*) impacted differently on decisions of how to manage ongoing breast cancer recurrence. This bears further consideration in future research.

Previous research has found high levels of health care utilisation among patients who exhibit excess levels of concern about disease symptoms or illness (as a diagnostic component of somatic disorder; the presence of medically unexplained symptoms).[329-332] Taking into account patients' underlying conditions and comorbidities, increased health care use among somatising patients has been shown across all levels of care including primary care, specialist care, ambulatory and hospital inpatient services.[329,330] It is plausible that similar patterns of health care use might be exhibited among women

who are concerned about cancer recurrence following breast cancer treatment. This is a topic worth exploring in future research.

6.4.1 Limitations

The intention in the initial design of this research was to compare the determinants of preferences for the ongoing management of breast cancer risk among women who had an experience of breast cancer with those of women from a general community sample. Ultimately, this was not possible as the DCE survey was conducted among a general community sample. Thus the results of this research are derived from, and applicable to, a population for whom a future choice regarding the ongoing management of breast cancer risk might be possible. Moreover, since women who had experience of breast cancer were part of the qualitative research and the DCE development, the lived patient experience has been included, enhancing the generalisability of this research.

As with the research in Chapter 5, the research in this chapter was limited in terms of the sample size available within each of the framing sets. Although approximately 150 women completed each of the three framing sets, the sample size at which coefficient estimate precision begins[47], larger within frame sample sizes would have increased the precision of the estimates. In particular, it is possible that the marginal difference noted for the pooled Chow test might have achieved significance with a larger sample size.

The results show a high degree of non-trading, largely in favour of routine monitoring over CPM. This was observed through the magnitude and significance of the alternative specific coefficient. The source of such strong preferences appears to lie in information embedded in the choice option labels. However, it is possible that some women were motivated to always choose monitoring because there were some attributes (*Surgical OOP, Pain Risk and Sensitivity Risk*) that did not apply to that option. This could not be tested for the non-traders. An alternative choice structure that might have induced more trading would have been to allow these attribute levels to vary

over both options; but this would give rise to implausible combinations, which could compromise the external validity of the survey. The other alternative would have been to use an unlabelled design, and to include an attribute for removal of the contralateral breast. This might not have reduced non-trading since the previously described restrictions would have continued to apply (essentially forming a quasi-labelled design).

6.4.2 Conclusion

The research presented in this chapter focuses on choices women make regarding the ongoing management of breast cancer recurrence risk. Qualitative research was used to develop a stated preference survey used to elicit women's preferences regarding the choice of having CPM or routine monitoring only as a means of managing ongoing recurrence risk following treatment for early stage breast cancer. Results from 464 women show a strong preference for the mode of monitoring, regardless of any of the underlying choice characteristics; over 50% of women always chose one option, typically routine monitoring, with fewer women always choosing CPM.

Typically, where women were willing to choose between options, those choices were most highly influenced by the associated health-effects of monitoring, notably the impact on subsequent cancer risks. There was some evidence that women valued meta-health effects, notably being able to access less intrusive methods of monitoring (such as ultrasound), and being involved in decisions about their care. The extent to which those meta-health effects influenced women's choices, and the value placed on them by women, differed depending on the amount and type of information included in the health vignettes and choice scenarios.

Finally, it was possible to categorise women according to whether or not they were concerned about cancer recurrence when completing the choice scenarios. Women who were concerned about cancer recurrence, potentially associated with a fear of recurrence, were more likely to choose CPM over routine monitoring, and attached a

statistically significantly higher value to reducing the risk of cancer recurrence than did women who were not cancer concerned.

These results add to the literature in a number of ways. First, the question of women's preferences for managing ongoing breast cancer risk has not been explored previously using a DCE. This research is the first to enumerate the factors influencing the decisions women make, including the extent to which they hold to a particular preference in the face of shifting stimuli. Second, it was possible to identify women as being cancer concerned or not, and thereby infer the value those women might place on achieving reductions in cancer risk (and potentially increased reassurance). Third, patients were integral to both the development and conduct of this research to enhance its relevance and interpretation. Finally, it explores the impact of framing on the value of meta-health effects. Together with the research in Chapter 5, it adds to the evidence that individuals' preferences for meta-health effects relative to health effects are affected by the manner in which information is presented. Overall, the findings of this research are important in terms of how options for the ongoing management of breast cancer are communicated to women; understanding that some women will have a strong preference for one form of management over another is an important aspect of patient centred care.

7 What Role for Meta-Health Effects? A Discussion

Chapter Summary

The results of the three case studies in this thesis show that meta-health effects matter to individuals in a number of decision contexts. The use of preference measures to explicitly value those meta-health effects means that they can be used as inputs for public sector decision-making. The finding that attitudes can also be used to assess the importance of meta-health effects suggests that they can also be considered in such decision-making.

The results in this research suggest that meta-health effects can be valued in a way that is consistent with their inclusion in CBA. This creates a tension for public sector decision-making in health care in Australia that to date has focused on the use of CUA. This tension might be overcome through the explicit inclusion of the values derived for meta-health effects into prices for technologies when evaluated using CUA. Alternatively, the results of health-outcomes focused CUAs could be supplemented by analyses focused on comparing the value of meta-health effects with the costs associated with their production.

The results of the studies assessing the value of meta-health effects show that these values are health care context specific and influenced by framing. There is evidence of reference bias, focus bias and priming effects that all favour the relative value of meta-health effects compared with health effects. The recommendation is that preference studies for meta-health effects include as much detail as possible, particularly where those values are to be used as inputs for public sector decision-making.

The research in this thesis also raises the question of whether governments should pay for gains in meta-health effects simply because they are valued by individuals. This is an area for further study. Similarly, further study is suggested into the influence on the value of meta-health effects of the numeraire used in valuation studies.

7.1 Introduction

Decision-making in health care often involves comparisons of how much treatments cost and their effects on health outcomes. But the outcomes produced by engaging with the health care system go beyond health and include meta-health effects; the effects other than health outcomes that arise from the experience of using health care. These meta-health effects, such as convenience, autonomy and choice, and reassurance can all influence the decisions individuals and governments make regarding health care.

This thesis addressed the question "*What is the influence of meta-health effects in health care decision-making?"* through empirical studies in four settings. The first examined the existing evidence on decisions made by the PBAC for the consideration of meta-health effects. This was supported by a review of the published literature for evidence of how meta-health effects have been valued previously for inclusion in economic evaluations. The three other case studies focused on individual decision making in the primary care setting (choices for patient-GP loyalty relationships), the chronic care setting (choice of treatment for RA), and an acute intervention, with long-term implications (whether to undergo CPM to manage breast cancer recurrence risk).

A number of key sub-questions were also addressed in this thesis: how does the influence of meta-health effects compare with that of health effects?; what impact do meta-health effects have on values for use in economic evaluations?; how does framing affect values derived for meta-health effects?; and, who should pay for meta-health effects? The findings on each of these questions from the four case studies and their implications for incorporating meta-health effects into health care decision-making and future research are discussed in this chapter. Suggestions for further research in this area are also offered.

7.2 The Influence of Meta-Health Effects

The results from this thesis demonstrate that meta-health effects do influence decision-making in the health care setting, as evidenced by the results from all four case studies. The evidence review demonstrated that while health outcomes are the main driver of reimbursement decisions made by the PBAC, meta-health effects are also influential in decisions about which drugs are funded. As supported by the published literature, convenience was the meta-health effect most often investigated as a source of value in economic evaluations. Thus convenience was a meta-health effect investigated in the three other empirical studies. The results of the GP Loyalty study showed that individuals' attitudes to having a choice of providers influenced their relationships with their GP; convenience factors did not. In contrast, the DCE in the RA Therapy study showed that convenience in terms of the mode and frequency of administration were important to individuals' choice of therapy, but health effects dominated. Similarly, health effects dominated women's preferences for how to manage the ongoing risk of breast cancer recurrence in the DCE in the Mastectomy study. In that setting, the meta-health effect of autonomy influenced choices, as did women's attitudes to cancer recurrence.

Overall, the research expands on the existing evidence regarding the value for meta-health effects in a number of ways. First, it used DCEs to explore the trade-offs between health and meta-health effects, including in the previously unresearched area of women's preferences for managing ongoing breast cancer risk. Second, both DCE studies investigated the importance of framing effects and data analysis in producing values for meta-health effects, demonstrating that these values are influenced by the amount and type of information presented. Third, as described in Chapter 1 and 3, it used attitudes as well as preferences to capture the influence of meta-health effects, demonstrating that attitudes can be used to identify meta-health effects of importance (using the community survey based ratings data in the GP Loyalty study) and to assess differences in the values of meta-health effects between sub-groups (utilising the cancer concern ratings in the Mastectomy study). Finally, the RA Therapy study

explored the influence of who pays on the value of meta-health effects, finding that Government costs only influence choices where individuals are presented with more information on the convenience aspects of treatment.

7.2.1 Meta-health effects or health effects?

A key focus of this research was to investigate the trade-offs between meta-health effects and health effects in influencing health care decision-making. All three empirical studies found evidence that meta-health effects influence health care decision-making; the extent of that influence relative to health effects varied. The largest consistent influence for meta-health effects was observed for treatment choices in RA, while the smallest was observed for choices with respect to the management of ongoing breast cancer risk. While these studies were conducted in different groups, and were subject to different experimental conditions (in terms of the DCE designs and information provided), a difference in the relative value of meta-health effects observed for choices in RA therapy and the management of ongoing breast cancer risk suggests that the importance of the health implications might influence the extent to which individuals are willing to consider meta-health effects.

The difference in the value of meta-health effects relative to health effects in those two decision contexts is consistent with prospect theory; the impact of a loss is greater than a gain of the same magnitude, and in making choices associated with risk, individuals overweight the likelihood of rare events occurring.[80,81] That is, for women making decisions regarding the management of breast cancer risk, the potential loss in well-being associated with a second diagnosis of cancer is large compared with the losses that might be incurred by a patient being treated for RA. In the presence of a larger loss, individuals place more weight (value) on the choices they make in avoiding the potential loss as opposed to improving the experience of care. Therefore, while meta-health effects are still relevant, in this context more weight is placed on health outcomes. For those making decisions about the ongoing treatment of RA, the focus was on the chronic nature of an existing condition. While treatment will potentially

improve that condition (a gain), more frequent and invasive treatment administration was perceived as a greater loss, shifting individuals' focus to place value on the convenience factors of treatment. This resulted in a higher value for the experience of care relative to health effects.

7.2.2 Valuing meta-health effects for economic evaluations

The results presented in Chapter 2 show that the PBAC has made recommendations on the basis of evidence that included meta-health effects. The majority of cases in which claims were made on the basis of meta-health effects involved a gain in convenience. This is consistent with the published literature that shows convenience is the most widely investigated meta-health effect.

The majority of published studies assessed the value of meta-health effects using WTP, either through CV or DCEs. In keeping with this result, the research in this thesis used DCEs to assess the value of meta-health effects for treatment choice in RA and the ongoing management of breast cancer risk. In both studies, it was possible to estimate the value of the meta-health effects of importance; convenience in the case of treatment for RA, and autonomy (involvement in decision-making) for the ongoing management of breast cancer recurrence. For the latter, additional subgroup analyses were used to infer a value for reassurance based on differences between women who were concerned about cancer recurrence and those who were less concerned on reducing the risk of cancer recurrence.

The finding that individuals derive value from these meta-health effects suggests that such values are relevant for consumers of health care and therefore have a role in economic evaluations in these indications. Whether such values should be incorporated into an economic evaluation in part depends on the perspective adopted for the analysis. If a societal perspective is taken, it seems reasonable that the full value of the meta-health effects be included in the assessment of benefit. However, if a health care system perspective is adopted it might be reasonable to reduce the benefit

attributed to the meta-health effect to apportion some responsibility for its acquisition to the patient (although this might mean that some individuals would not be able to afford treatment from which they derive significant benefits and therefore have implications for equity).

Whether governments should pay for meta-health effects, or whether their cost should fall on those to whom they accrue, is still in question.[72] As a society, we might be willing to pay for individuals to have access to treatments or health care services that provide them with more meta-health effects. This proposition has not been tested in this thesis, but can be tested in much the same way as previous studies have tested whether individuals believe governments should pay for health gains or QoL gains, and whether this depends on the population to which those benefits accrue.[333-335] The implication is that such further research will reveal a hierarchy of outcomes (health and meta-health effects) for which society is willing to pay.

Whether the inclination to pay for meta-health effects differs depending on the population who benefits is of particular interest in the context of the current research; women who were concerned about cancer recurrence had a higher mWTP to avoid a higher risk of cancer than those who were not. This has implications for a funder that agrees to reimburse a product or service on the basis of a demonstrated value shown using mWTP since, in the case of managing the ongoing breast cancer recurrence risk, it implies that those who demonstrate greater cancer concern have the ability to 'attract' more funding than those who do not. Choosing to reimburse interventions that are more highly valued by some groups than others has potential equity implications, particularly if access to those interventions is also limited by subgroup.

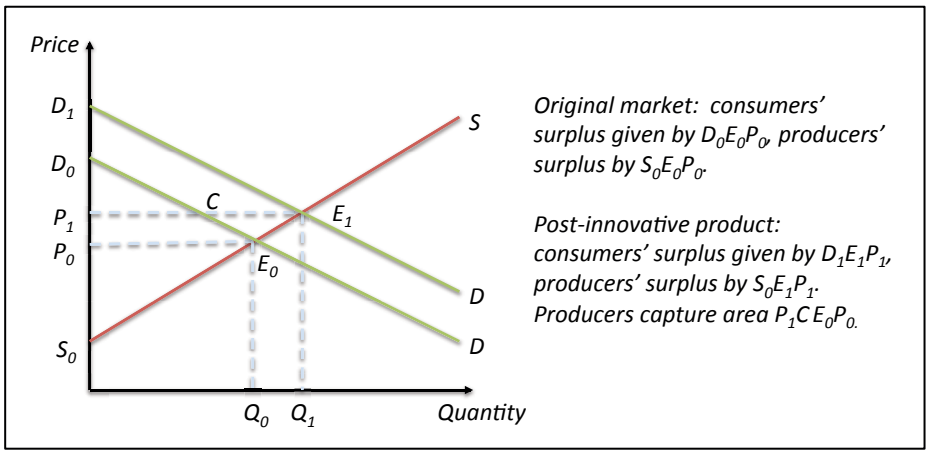
7.2.2.1 Capturing the consumers' surplus

The values derived for meta-health effects have been incorporated in decision-making by using them as the basis of a claim for a higher price for a new product or service that delivers that effect (witness the claims in submissions made to the PBAC). The

claim is that since those effects are valuable to individuals, the production of those effects should attract higher prices. This implies that producers of goods or services seek to capture some of the gain in consumers' surplus that individuals would otherwise derive from the meta-health effects associated with their products.

In consumer theory, the consumers' surplus is reflected by the difference in the maximum amount individuals would pay to avoid going without a given quantity of a good or service, and that which they actually pay for that same quantity.[336,337] This is typically represented by the area lying to the left (below) of the individual's demand curve for a good or service, and above the price paid at the quantity consumed; e.g. the area $D_0E_0P_0$ in Figure 53. Where a new health care product or service is introduced that provides more of some desirable attribute, say convenience, with all else being at least equal, individuals might be expected to demand more at all existing prices; the demand curve shifts to the right (D_1D in Figure 53) so that for a given supply curve (S_0S), at given quantities, the price within a competitive market increases.[336,337] In this case, the difference between the price paid for the initial quantity demanded, and the price at the new quantity demanded, reflects a capture of the gain in consumers' surplus arising from the introduction of the new product ($P_1CE_0P_0$ in Figure 53).

Figure 53: Consumers' surplus with innovative products



If producers can estimate the value individuals place on meta-health effects associated with their product prior to its introduction, it allows them to increase their price to reflect the associated rise in consumers' value. Walzer et al. (2013) note that while evidence exists of producers attempting to capture the consumers' surplus arising from gains in meta-health effects, such as convenience, their ability to do so is limited where purchasers possess strong price negotiating power.[338] This is of particular interest in negotiated product markets, such as the publicly reimbursed PBS in Australia within which the Government acts as a monopsony buyer from largely monopoly sellers.

The proposition that producers are able to capture the consumer surplus as part of the price they negotiate rests in part on the assumption that it is possible to measure the demand function for such goods, across all individuals. Within this thesis, DCEs have been used to investigate the impact on choice of various meta-health effects. Train (2002)[339] shows how the consumers' surplus can be estimated based on choice coefficients derived using such DCEs. While such calculations are possible, it is unclear how their results should be interpreted. The use of DCE results to estimate gains in consumer surplus implies that DCEs establish a relationship between the quantity demanded and price. This would require testing of choices across different combinations of price and quantity for the good or service in question. However, in most instances DCEs assess the probability of a product or service being demanded at various prices for a given quantity. Knowing the probability that a given quantity will be demanded at different prices does not establish the relationship between price and quantity required to specify the associated demand curve, and hence to estimate the consumers' surplus.

If a relationship can be formed between expected quantities of use and price, Whynes et al. (2005)[340] show that the method by which WTP values are elicited in stated preference research impacts on the functional form of the subsequent demand equations that can be elicited. This in turn impacts on the size of the consumers' surplus that can be estimated using those demand functions.[340] This is supported by

Gyrd-Hansen (2013)[278] who cautions that while stated preference methods are valuable in being able to derive values for a concept of interest, health effects or meta-health effects, the values derived will be influenced by the nature of the payment vehicle used in the elicitation task, e.g. via OOP or through links to income tax. How those monetary values are framed therefore depends on the system in which those preferences will ultimately be used. Thus, if OOP costs are used to elicit mWTP, such as in the RA Study in this thesis, and these are subsequently used to support a claimed price differential for a publicly reimbursed product, it is not clear that the initial preferences have been derived in a manner that best reflects their ultimate intended use.

The extent to which DCEs reflect changes in consumers' surplus, let alone those for meta-health effects, thus represents an opportunity for further research. This could help to further inform the question of whether governments should be paying for products and services that result in increases in meta-health effects.

7.2.3 Framing and meta-health effects or health effects

The results from the research in this thesis show that how information is presented, and the health experiences it describes, can influence the values for meta-health effects by shifting the focus of respondents, influencing the reference points they use as comparators and priming respondents with different information before they commence a valuation task. This finding is not universal. Recent results by Rowan et al. (2016)[335] show that framing effects, tested in terms of the amount and type of information presented to respondents, did not influence preferences regarding different burdens of illness (combinations of QoL and survival).

The results from both DCEs in this thesis indicate that when told that choice profiles do not differ in their efficacy (they do not carry a risk of a loss or gain), respondents focus on the attributes that do differ, and place more focus on the meta-health effects – notably convenience in both DCEs. Similarly, providing respondents with more

information about an attribute alters the manner in which it is perceived, both within the attribute and relative to other attributes. This was observed for the 'Attribute' frame within the RA Therapy study; providing individuals with more information about the mode of treatment administration increased its relative value across all three levels of that attribute. This is consistent with what would be predicted by prospect theory[80]; providing more information shifted the reference (or anchor) against which the attribute levels were assessed (e.g. providing more information clarified the potential loss in convenience from shifting from an oral tablet to IV treatment administration). Finally, the results of the Mastectomy study indicate that providing women with more information about the process and background of their breast cancer diagnosis resulted in their placing different values on the labelled choice options; there was no longer a preference for routine monitoring over CPM. In this case, priming respondents by providing additional information[87] as part of the health scenario influenced not only the comparisons between attributes, but women's overall preferences between the labels, potentially by altering the reference from which comparisons were made.

These findings are important for three reasons. First, because they show that the way tasks are framed and presented influences individuals' preferences. Understanding the potential influence of the elicitation task on preferences is an important aspect in terms of examining the validity of the resulting values.[77,78] Second, understanding framing in the context of valuing meta-health effects matters because these effects are being used as a source of difference in economic evaluations for products seeking public sector reimbursement. Third, as demonstrated in the RA Therapy study, the presence of framing effects also influences how choices are analysed; whether health effects behave as continuous (linear) or categorical variables is affected by framing, which in turn influences the estimated trade-offs between health effects and meta-health effects.

The results in this thesis suggest that, where possible, it is preferable to provide respondents with as much information as is practical without making the task cognitively burdensome. While using simple descriptions of attributes, their levels and the overall health scenario might lessen the cognitive requirements of choice experiments, they might also result in a misrepresentation of the values placed on meta-health effects compared with health effects. In addition, the results show the importance of being aware of the potential for bias due to the manner in which information is presented and to consider this in the design and subsequent analysis of DCEs purporting to assess values for meta-health effects.

7.2.3.1 Attitudes and preferences

Two approaches have been used in this research to reflect the role of meta-health effects: as attitudes (GP Loyalty study) and preferences (RA Therapy and Mastectomy studies). As described in Chapter 1, the principal distinction between attitudes and preferences is that attitudes reflect an unconstrained (non-comparative) rating over an attribute, while preferences are comparative and reflect trade-offs between attributes.[76,97] Kahneman and Sugden (2005)[76] extend this distinction to claim that preferences are formed in relation to experienced events. This would imply that stated preference surveys are eliciting experienced preferences. However, the hypothetical nature of DCEs and other stated preference techniques means that it could be argued that respondents form their preferences at the time of completing such tasks, albeit informed by prior experiences.[77,78] This implies that preferences would differ between respondents with and without experience of the event of interest. Some evidence of such an effect was observed in the RA Therapy study which showed a difference in the distribution of attribute importance between those respondents taking chronic medicines and those on no medications.

Another difference between attitudes and preferences is their capacity to convey a quantum of value for the effect of interest e.g. choice in patient-GP loyalty. While it was possible to demonstrate a role for meta-health effects using attitudes, their use did

not permit the derivation of a 'value' associated with the meta-health effect in that research. Nonetheless, as noted in Chapter 3, the finding that choice influences the patient-GP loyalty relationship is of relevance when considering policy changes that might impact on the ability of individuals to choose between GP practices, such as requiring patients to nominate or enrol with specified GP in order to qualify for publicly funded care. In addition, the results from the Mastectomy study indicate that attitudes (in this case to cancer concern) can be used to gain a richer understanding of preference information; how preferences differed based on whether or not women are concerned about cancer recurrence. This has direct policy relevance in not only understanding what influences the decisions women make, but how the messages provided to women might need to differ based on their attitude to cancer recurrence.

7.3 Applying these Findings

The results of this thesis show that meta-health effects matter to individuals, supporting the inclusion of such effects in policy analysis and public sector decision-making. Consideration of meta-health effects for this purpose might be implicit or explicit. Implicit inclusion of these effects refers to their consideration where they are known to be important, but a measure of their value is not available e.g. attitude ratings of the importance of choice. Explicit inclusion of such values would include the use of mWTP, or other measures of value, as inputs to relevant economic evaluations to inform public sector decision-making.

Adopting an explicit approach to the inclusion of meta-health effects suggests a broader role for CBA in decision-making for public funding of health care in Australia. This would align public allocation decisions for health care with recommendations for the evaluation of resource allocation in other sectors of public expenditure.[341] The general recommendation in Australia is that government agencies should conduct a CBA when seeking assess whether the benefits of a new programme justify the costs.[341,342] As discussed in Chapter 1, and outlined in the Australian Cost-Benefit Handbook[342], one of the appeals of CBA is its scope to present a broad range of

benefits. Adopting a CBA approach in health care would be consistent with expanding the scope of benefits to include meta-health effects.

Smith and Sach (2009) note the adoption of CBA in health care has been limited, potentially by the absence of a "road-map" or clear guidance on how it is to be applied in decision-making in the health care setting.[343] The research in this thesis helps to inform that road-map by showing that monetary values for both meta-health effects and health effects can be derived in the form of mWTP. But, the use of CBA is typically applied to the introduction of entire programmes of expenditure (e.g. the PBS), rather than parts of an existing programme (e.g. funding for a specific drug).[342,343] This is not typically the manner in which assessments of health care programmes are enacted; comparisons are based on the introduction of a new technology, or service, how it relates to existing services, and the changes that may occur within the programme. Recommendations from the Commonwealth Government in Australia, and Smith and Sach (2009), are that where the analysis is focused on such within programme redistributions, the use of cost-effectiveness analysis, utilising health outcomes or QALYs as the relevant endpoint, is likely to be more informative.[342,343]

Such a recommendation suggests that a shift to CBA for health care decision-making would not be justified on the basis of aligning decision-making in that sector with other Commonwealth Government sectors. Rather, it would rely on the premise that its use would enable the evaluation of a broader range of benefits than those captured by QALYs. This once again raises the question of whether Government should pay for such benefits, meta-health effects.

7.3.1.1 Hybrid evaluation approaches

Monincx (2000)[3] suggests that the value of meta-health effects can be readily combined into existing CUA decision-making frameworks by using the mWTP for those effects as the basis on which to value resource use when estimating costs for inclusion in a CUA.[3] However, including the value derived via mWTP to form the costs in a CUA might

constitute double counting since it is possible that the outcome measure (the QALY) might also capture some aspect of the utility impact of those meta-health effects.

Mortimer (2006)[344] suggests that in many instances, such as where there are distributional concerns, CBA might play a role in providing decision-makers with a second metric of 'value for money' where questions remain as to the acceptability of an initial cost per QALY. While Mortimer's suggestion was couched in terms of overcoming QALY ratios subject to some 'distributional' deficiency, it bears further consideration in the context of assessing meta-health effects. It might be the case that the valuation of a programme incorporating meta-health effects can be formed, for instance using a DCE, that allows inputs to both QALYs and CBA to be estimated. Thus, the analysis would consider the results using a CUA approach, and a CBA approach with its potentially broader assessment of value. It is unclear which result would be given primacy if the CUA and CBA were to produce different conclusions.

7.4 Some Final Thoughts

7.4.1 Where to from here?

That DCEs have been used to assess preferences of meta-health effects is clear, both from the existing literature and from the research in this thesis. The evidence to date is that the bulk of such valuations are in the form of mWTP. However, there have been studies that have considered trade-offs between survival and meta-health effects; notably, Harrison et al. (2015)[254] produced utility values on a 0 to 1 scale for the mode and frequency of delivery of RA treatment using a DCE, while a number of TTO and SG experiments have been used to investigate meta-health effects.[63-68,99,113,345] The use of DCEs incorporating a duration (survival) attribute to value meta-health effects warrants further investigation.

Such research might also explore two additional questions regarding the assessment of value for meta-health effects. The first is whether it is possible to include duration and

cost as attributes within the same choice profiles in a DCE; how would individuals trade-off in the presence of these potentially competing sources of value? The second, is that if individuals are indeed willing to trade-off survival for meta-health effects what are the implications for government funding of products that influence meta-health effects?

Finally, this research has addressed the question of how individuals alter their choices in the face of changing costs to governments for treatment for RA. Further research is warranted on whether we as a society are willing to allocate public sector resources on the basis of gains in meta-health effects; should governments pay for meta-health effects? Based on the existing research, this remains an important area of inquiry, both in terms of the propensity for gains in meta-health effects to be used to motivate claims for government funding, and given the methodological challenges of eliciting values for such effects in a robust manner.

7.4.2 Concluding comments

The research in this thesis has focused on how meta-health effects influence decision-making in health care. It has shown that meta-health effects do have a role in influencing health care decision-making, from the choice of patient-GP loyalty relationships to managing the ongoing risk of breast cancer. The degree of influence demonstrated varied depending on the context of that decision; meta-health effects were of lesser influence where there were greater potential losses/gains associated with the health effects of that decision. Qualitative and quantitative research methods were used to show that it is possible to estimate values for meta-health effects in a manner that can be used as inputs for subsequent economic evaluations. Those values are subject to framing effects, both to the type of information presented and the manner in which it is analysed.

Overall, these findings result in a cautionary tale for governments and funding agencies alike. At a fundamental level is the question of whether or not governments

should fund treatments on the basis of gains in meta-health effects. At a more pragmatic level, faced with contemporary requests to fund new treatments that affect the experience of care, and which purport to capture the value of meta-health effects, careful attention needs to be paid to the manner in which such values have been derived lest they misrepresent the resulting value to society.

Appendix 1: Chapter 2 – Studies Included in Systematic Review

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
Augustovski et al 2013[90]	Rheumatoid arthritis	Drug	Convenience	Research focus – preferences	Commercial funding	DCE – forced choice, unlabelled. Two blocks of choice sets: one with 12 choice sets, the other 13. Choice question: Which option do you prefer?	Inclusion of a cost attribute allows calculation of WTP – which is done; both for the overall profile and to test for attribute importance (including NHO).
Benning et al 2014[132]	CRC	Diagnostics – screening	Convenience/ Acceptability	Inform policy.	Not discernible	DCE – labelled, with an opt-out (no test) option. Design of 72 choice sets blocked in 12 per respondent. Analysed with MNL (clogit). Choice question: I prefer	No, the design did not include a cost attribute.
Bijlenga et al 2009[150]	Prenatal care (induction at term due to complications)	Obstetrics	Process	Research focus – methods	Research grant funding	Comparison of DCE ("choice of best") and TTO (included VAS). The number of choice sets or vignettes (for TTO/VAS) was not clear from the publication. DCE was forced choice (does not state if labelled or unlabelled). Preference question was not provided, but the participants were asked to choose the best vignetter out of the two on offer.	While the TTO could have produced values for use in an economic evaluation, the actual utility values were not reported. TTO rated harder to complete than DCE. DCE did not include cost or time so cannot be used for economic evaluation. VAS converted to "utility" by dividing by 100.
Birch et al 2003[65]	Abnormal pap test result (watchful waiting versus colposcopy).	Screening	Process	Trial focus	Research grant funding	SG used to derive utility scores.	Find that where pathology indicated immediate treatment is not needed, patients prefer to wait. Different utility values for aggressive and conservative follow-up in face of the same clinical outcomes shows value of process utility. Difference in utility values reflect process utility value.
Borghi and Jan 2008[71]	Pre-maternal education	Health promotion	Information Confidence Community/	Research focus – preferences	Not discernible	WTP (contingent valuation).	WTP values reported for overall program, not broken down by contributing factors.

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
			Altruism				Note that NHO were valued by 84% of respondents.
Boye et al 2011[91]	Diabetes	Drug	Convenience	Innovation	Commercial funding	SG assessment of nine health states. Participants also valued own health.	Separate utility values are presented for each of the health states reporting on the various different injection/flexibility regimens. Utility changes due to NHO are small (0.02) – while statistically significant, unlikely to be meaningful and dwarfed by health changes. Discuss importance of context.
Brett Hauber et al 2011[117]	Hepatitis C	Drug	Convenience	Research focus – preferences	Commercial funding	DCE – forced choice. Respondents completed nine choice sets (two alternatives in each). Preference question: which medicine would you choose if these were the only options available?	DCE excluded both cost or survival attributes so could not be used to derive utility values for use in economic evaluation.
Brown et al 2011[116]	Haemophilia	Drug	Convenience	Research focus - preferences	Commercial funding	DCE. Each respondent completed 12 choice tasks. Task was unlabelled. Factorial design not described. Three alternatives presented. Choice task: Which treatment are you most likely to use? Which treatment are you least likely to use?	Preferences presented as part-worth utilities, but not scaled by survival. WTP values not presented for attribute levels.
Bryan and Jowett 2010[126]	Warfarin anticoagulation	Drug	Convenience	Trial focus	Research grant funding	WTP. Respondents asked how much they would be willing to pay per month for self-monitoring (vignette was non-comparative). Other methods of monitoring were not valued.	Limited in that it only provides an absolute value for that one health state (no comparative values are provided).
Bunge et al 2010[101]	Scoliosis	Device	Convenience/ Acceptability	Trial focus	Research grant	DCE with three alternatives (including no treatment).	The DCE does not include a cost attribute, nor a survival attribute.

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
					funding	Participants completed 16 choice sets. Question elicited as: "Which brace do you prefer"?	
Burton et al 2014[155]	Stroke	Community services/population health	Process Convenience	Inform policy.	Research grant funding.	DCE – split groups, patients and carers. Forced choice, unlabelled. Choice question: Please tick one box.	No, the design did not include a cost attribute.
Cairns et al 1996[106]	Antenatal Screening	Screening	Information	Research focus – methods	Public funding	Interviewer led SG.	Utility values are reported, but the manner in which they are elicited would suggest that they more closely align to the outcome (likelihood of CF) than the value of information.
Carroll et al 2013[152]	Obstetrics	Diagnostic/screening	Anxiety/process Information	Research focus - preferences	Research grant funding	DCE forced choice, unlabelled. Respondents saw 8 choice pairs. Analysis by clogit and latent class. Choice question: Which test would you prefer?	Yes – present WTP values.
Chan et al 2009[129]	Prenatal screening	Screening	Information Access/ anxiety	Research focus – preferences	Not discernible	DCE with an opt-out choice (no-test). Participants completed 8 choice sets each (two blocks of 8 out of a design of 16). Choice question was not provided.	Reports WTP amounts for the different test types that are used to proxy the levels of information.
Chancellor et al 2008[107]	Diabetes	Drug	Convenience	Innovation	Commercial funding	Chained TTO. Also valued using EQ-5D.	Interesting result that the EQ-5D undervalued compared with TTO for the health states, but not for own health.
Cook et al 1994[108]	Cholelithiasis – lap chole-cystectomy	Surgery	Process	Trial focus	Public funding	Chained TTO used.	While utility values are presented for each of the surgery types, the corresponding vignettes are not. It is therefore not clear what patients were asked to consider in their valuations –

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
							were process effects mentioned?
Davison et al 2010[135]	Chronic kidney disease	Multiple	Information Autonomy Process Access	Inform policy	Research grant funding	Forced choice DCE (no opt out). Was not labelled. Respondents completed 12 choice sets each. The choice question was: "which program do you prefer".	Neither cost nor survival were included as attributes in the DCE.
de Bekker-Grob et al 2013[137]	Prostate cancer	Multiple	Convenience	Resarch focus	Research grant funding.	DCE forced choice, labelled design (three labels); presented as paired choices only. 24 choice sets blocked into 12. Mixed logit analysis. Choice question: Which alternative do you prefer?	No, the design did not include a cost attribute.
Dixon and Shackley 1999[70]	Public health – water fluoridation	Population health	Community/reassurance	Research focus – methods	Not discernible	Contingent valuation. Payment expressed as additional taxation. Used payment cards. Those not in favour were asked how much they would have to be compensated to accept fluoride in water.	WTP are not stated according to NHO or individual components of programme.
Donaldson and Shackley et al 1997[146]	Cholelithiasis	Surgery	Process	Research focus – preferences	Public funding	WTP. Respondents asked how much additional tax per year they were WTP to receive laparoscopy instead of laparotomy.	Report WTP values for scenario with and without process effects.
Dwight Johnson et al 2010[139]	Depression	Multiple	Convenience Access Process	Research focus - preferences	Research grant funding	Described as conjoint analysis (method consistent with DCE). Preferences were forced choice (no opt-out). Factorial design not stated. Each respondent completed six choice sets. Choice question not shown, but described as being "the one preferred".	Results were analysed as OR only. WTP values were not calculated for any of the attributes.

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
Dwight Johnson et al 2013[140]	Depression	Multiple	Convenience Access Process	Research focus – preferences	Research grant funding	Described as conjoint analysis (method consistent with DCE). Factorial design not stated. Each respondent completed six choice sets. Choice question not shown, but described as being "the one preferred". Preferences method not stated but appears to be forced choice (no opt out).	Results were analysed as OR only. WTP values were not calculated for any of the attributes.
Fiebig et al 2011[115]	Female contraception	Drug	Autonomy Convenience	Research focus – preferences	Research grant funding	Labelled DCE. Respondents answered 32 choice sets (three alternatives in each); four blocks, total of 128 sets. Forced choice (no-opt out option). Alternatives valued using best-worst: Of the options presented above, which do you like the most? Of the options presented above, which do you like the least?	Marginal WTP values are estimated for all the attributes, including the NHO.
Franken et al 2013[136]	Social preference	Multiple	Process	Research focus – preferences	Not discernible	DCE – forced choice, unlabelled. Choice question: not stated	No, the design did not include a cost attribute.
Gerard et al 2003[130]	Breast cancer screening	Screening	Information Convenience Process	Inform policy	Research grant funding	Labelled DCE. Overall design produced 32 choices, blocked into two sets of 16. Each compared with no screening. Participants each answered 16 options (shown as one profile in each case that participants chose yes or no). Choice question was: Imagine that your next invitation to be screened is approaching, would you choose to	Cost is not included as an attribute so marginal WTP can't be estimated. While time is included as a an attribute, there are four time variables so it is unclear how it might be used to estimate trade offs.

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
						attend this particular screening service or not? (yes/no).	
Gidengil et al 2012[111]	Paediatric immunisation	Combination vaccine	Convenience	Inform policy	Research grant funding.	DCE: 17 choice sets per respondent. Unlabelled, forced choice. Preference question: If you had to choose one scenario, which would you prefer? Also asked TTO and WTP (CV) (randomised). to avoid child receiving injections.	WTP values reported for the DCE and the individual CV exercise. For the TTO, the traded amounts are reported but the denominator is not. Therefore, utility values cannot be estimated.
Guimaraes et al 2009[120]	Diabetes – Type I and II	Drug	Convenience	Innovation	Research grant funding	DCE. Participants completed 17 choice sets. Choice task was not shown.	WTP values estimated for each of the attributes and levels (including the NHO).
Haas 2005[157]	GP Choice	Primary care	Process Information Autonomy Reassurance	Research focus – preferences	Not discernible	Unlabelled DCE. Respondents completed 24 choice sets – single profile. Choice task: If you need to go to the doctor again for a check-up, would you choose: your own GP; The GP described above; Another GP.	Neither costs nor survival were included in the attribute list so the results can't be used in an economic evaluation.
Hauber et al 2011[128]	Plaque psoriasis	Drug	Convenience	Research focus – preferences	Commercial funding	Unlabelled DCE, with follow-up contingent valuation in a subsample of patients. Forced choice. Choice task: Which treatment would you choose if these were your only options?	WTP values are not presented for the NHO, only for the health profiles which contain the NHO.
Hechmati et al 2015[121]	Bone metastases	Drug	Convenience	Research focus - preferences	Corporate funding.	DCE – not clear if forced choice or labelled. 40 choice sets, blocked into 10. Random parameters logit. Choice question: not shown	No, the design did not include a cost attribute.

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
Hendrix et al 2010[148]	Labour decision (obstetric)	Obstetrics	Autonomy Process	Inform policy	Not discernible	Unlabelled DCE. Each respondent completed seven choice sets. Forced choice (no-opt out). The preference question: Which profile do you prefer?	While cost was included as an attribute, WTP was not estimated from the results.
Howard et al 2008[104]	HPV/Cervical screening	Screening	Process	Research focus – methods	Research grant funding	Two stage SG. Utility scores estimated for paired scenarios essential comparing alternative means of screening/management.	Mean utility values presented (but are for overall health state).
Johnson et al 1996[62]	Cytomegalovirus	Drug	Convenience	Inform policy	Commercial funding	TTO Interviewer guided. Three health states: oral, IV and current health.	Values are reported for the treatment modality specific health states.
Kaambwa et al 2015[138]	Aged care	Multiple	Process Convenience Autonomy	Inform policy.	Research grant funding	DCE, forced choice, unlabelled. 18 choice questions blocked into 6 per respondent. Analysed clogit, mlogit and gmnl. Choice question: Which package would you choose?	No, the design did not include a cost attribute.
Kan et al 2015[153]	Liver disease	Diagnostic	Process – comfort but based on third party. Anxiety – after test explained.	Research focus - preference	Not discernible	All patients completed preference information, and provided WTP for FS (as well as willingness to self-pay).	Total WTP recorded, but not in terms of difference due to MHE, as there is no WTP for LB.
Kauf et al 2008[38]	HIV	Drug	Convenience	Research focus - methods	Commercial funding	Conversion of SF-36 to SF-6D. Contributions of each factor to utility are then estimated.	Results from a regression analysis are presented which show the contribution (only significant and negative for dosing flexibility) of NHOs. However, these are not utility values per se. Overall

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
							utility values for given CD4 counts are presented but it is not clear what combination of NHO attributes is used in those profiles.
Kauf et al 2012[125]	Hepatitis C	Drug	Convenience	Research focus – preferences	Commercial funding	Described as choice format conjoint analysis (DCE). Respondents answered 20 choice sets each. Unlabelled experiment. Choice question: Which treatment would you choose if these were your only options?	Neither cost nor survival were included as attributes. Resulting coefficients cannot be interpreted as utility values for use in CUA, and no WTP values are produced for CBA.
Lakdawalla et al 2012[73]	Solid tumours (melanoma, breast cancer)	Drug	Hope	Inform policy	Commercial funding.	Combined rating and WTP. Participants shown a profile sure bet and hopeful therapy (differ in that the latter offers hope of longer survival, but chance of earlier death), then WTP for the preferred profile.	WTP values are presented. However it is unclear if these truly reflect hope (NHO) or an anticipated survival gain.
Marsidi et al 2014[141]	Cosmesis	Multiple	Convenience Information Reassurance	Research focus - preferences	Not discernible	CA based task; respondents saw 18 scenarios. Choice task: Indicate on the scale below how likely it is that you would visit this clinic (1-7).	No, the design did not include a cost attribute.
Marti[154]	Population health – smoking cessation	Population health	Altruism Autonomy	Inform policy	Research grant funding	Case 1 (attributes only in profile, no levels) best worst scaling DCE. Each respondent completed 16 choice sets. Respondents asked to choose which of the attributes (in each profile) would be the "Most deterrent" and "Least deterrent" with respect to the AE of smoking.	Classifies which factors are the most important to respondents (with respect to potential smoking deterrents) but does not produce values that can be used in an economic evaluation.

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
Matza et al 2015[127]	Hepatitis C	Drug	Convenience	Inform policy	Corporate funding	Community based TTO across 14 health states. Considered utilities using 10 and 1 year time horizon.	Presents disutility of different modes of administration.
Nafees et al 2015[156]	Mixed	Device	Convenience Process	Research focus preferences.	Commercial funding.	DCE – not stated if unlabelled or forced choice. Contained 18 choice sets. Analysis by clogit.	Yes – report WTP for gains in all levels of attributes.
Naik-Panvelkar et al 2012[134]	Asthma	Multiple	Convenience Process Information	Research focus - preferences	Not discernible	Unlabelled, forced choice DCE. Participants self-completed 9 choice sets each (8 blocks in total, 72 possible choice sets). Choice task: If your pharmacy were to offer one of the services above- either A or B, which one would you choose.	Marginal WTP is reported for the attributes, including the NHO.
Neumann et al 2012[72]	Predictive screening (Alzheimers/ Arthritis/ Prostate Cancer/ Breast Cancer)	Screening	Information	Inform policy	Research grant funding	Double-bounded binary bidding WTP. Each respondent answered two questions. Total of 16 possible scenarios – four diseases, with two different risks of disease (10% and 25%), and two levels of test accuracy (perfect and imperfect).	Report WTP for each of the diseases and disease risks. However, unclear how much of that value is attributable to the NHO, or to the health component. Is it anticipatory health effects driving the WTP?
Osborne et al 2007[64]	Haemoglobanopathies (iron overload)	Drug	Convenience	Innovation	Commercial funding	TTO (not chained). Interviewer guided.	Values are specific to the health states described and focus on the difference due to mode of treatment administration. The values are very specific to the NHO. Demonstrate lower value for SC compared with anchor (suggesting that "information" leads to revision of expectation).
Osborne et al 2012[63]	Schizophrenia (anti-psychotic)	Drug	Convenience	Innovation	Commercial funding	TTO (not chained). Interviewer guided.	Values are specific to the health states described and focus on the difference due to frequency of treatment

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
							administration. The values are very specific to the NHO.
Oteng et al 2011[131]	Cervical cancer prevention (HPV vaccine or Cervical Cancer Screening)	Screening	Convenience	Research focus - prefences	Not discernable.	Unlabelled force choice DCE (included "Neither" in choice set as option). Participants completed 12 choice sets. Choice task: Which one would you prefer. Question is not framed as a societal problem, and no apparent preamble, yet both men and women are included.	WTP values are reported for each of the attributes included. Reported utility values are not useable for an economic evaluation. WTP needs to be interpreted with caution given inclusion of males and females (in a question on cervical cancer not framed by a societal elicitation).
Palumbo et al 2011[142]	Assisted Reproductive Techniques	Multiple	Information/ Autonomy Process Convenience	Innovation	Commercial funding	Interviewer guided completion. Described as conjoint analysis (DCE). Each participant ranked eight product profiles (most to least preferred). Profiles were not labelled. No specific choice as they were ranked (no within profile comparisons). Contingent valuation WTP. If stated yes to amount noted, asked maximal amount, ditto if no. Participants also asked WTP for treatment characteristics (comfort, tolerance, effectiveness).	Produces overall WTP for ART and for the component NHO attributes. It appears this is only reported for CV. It is likely that difference in information shown in the two methods would result in them not eliciting the same thing. Utility values are not useable for EE as they are not survival based. Missed opportunity here not to compare these two methods. Includes RP data on costs of prior cycles. Suggest NHO of convenience of pen loses relevance once other attributes are included (yes – but does this mean it carries no value?).
Park et al 2011[158]	Diabetes telemedicine care	Telemedicine	Convenience Process Access Assurance	Research focus - preferences	Research grant funding	Conjoint analysis. Total of 16 profiles developed. Participants ranked profiles (single product) in groups of four only. Appears that this was a ranking task across all	Marginal WTP values are reported. The method of elicitation is however unclear. While analysed using an RUT model it is not clear how this is interpreted in the case of a four profile

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
						four profiles (not a choice task between two profiles). Profiles were not labelled.	ranking (where the comparisons were not pair-wise). Marginal WTP presented for all NHO.
Philips et al 2006[105]	Cervical cancer screening	Screening	Reassurance/ access Information	Inform policy	Not discernible	Self-complete CV (WTP) questionnaires. Three stages: Screening in general. Additional testing. Additional testing with HPV explained.	WTP values reported for each of the stages, including the additional scenarios (showing value of information).
Phillips et al 2002[112]	HIV	Diagnostics	Convenience Process	Research focus - preferences	Research Grant.	Unlabelled force choice (no opt-out). Participants completed 11 choice sets; six versions of the survey. Choice question not provided.	WTP values are reported for each of the attributes. Coefficients are reported but these are not on a survival scale.
Polster et al 2010[109]	T2DM	Drug	Convenience	Research focus - preferences	Commercial funding.	Patient preferences elicited using TTO and conjoint analysis (DCE). TTO was not standard – direct statement of trade (no ping-pong – self complete, and only whole years were possible). Respondents completed eight choice sets. Not clear Commercial funding. whether these were labelled, or the specific choice question.	TTO utility values are produced (which contain NHO). DCE values do not appear to have been scaled by survival (NHO given lowest relative importance).
Porzsolt et al 2010[124]	Diabetes (Type I or II)	Drug	Convenience Process	Inform policy	Not discernible	Conjoint analysis; unlabelled (forced choice). Number of choice tasks adaptive depending on answer to prior ranking questions. Choice task to comparative profiles – respondents had to state which option they preferred by stating: Strongly prefer A, prefer A, prefer B,	Cost is not included as an attribute, nor is survival.

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
						strongly prefer B.	
Prosser et al 2003[113]	Relapsing remitting multiple sclerosis	Drug	Convenience	Research focus – preferences	Research grant funding	SG. Respondents only saw one treatment health state (each health state only evaluated by approximately 20 patients/community members).	Presents results for treatment methods as utility values.
Protiere et al 2004[20]	Cardiac surgery Breast cancer treatment Helicopter ambulance service	Multiple	Information	Research focus – methods	Research grant funding	CV (WTP) – interviewer guided. Open ended elicitation (not bounded). Each health programme evaluated, but additional information only evaluated for cardiac surgery.	Provides WTP values broken down by information group.
Ryan 1999[149]	Assisted reproductive technique	Obstetrics	Process Reassurance	Research focus - methods	Research grant funding	Conjoint analysis (DCE), forced choice, unlabelled. Participants completed 12 pairwise choices (total of 26 profiles in total, split into two groups of 13 – one profile arbitrarily set as "base" for each group). Choice task: Which clinic would you prefer.	Marginal WTP values produced for each attribute.
Salkeld et al 2004[99]	Falls prevention	Health promotion	Reassurance Elation Regret Convenience	Trial based.	Research grant funding.	TTO – interviewer guided. Used named states (each health state had a woman's name).	Utility values are reported for each of the NHO
Schmier et al 2002[110]	Chronic pain	Drug	Convenience	Research focus – preferences	Commercial funding	Adaptive conjoint analysis with rankings and pairwise comparison. The number of scenarios completed is not provided (presumably because of the adaptive design). Choice task: assuming all other factors are equal, which of the	Cost is not included so WTP is not estimated. Similarly survival is not included so utilities for use in an economic evaluation can't be estimated.

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
						following do you prefer?	
Scotland et al 2011[147]	Labour (birthing) decisions	Obstetrics	Process	Trial focus	Research grant funding	DCE. Unlabelled, forced choice. Respondents completed 16 choice sets. Choice task: Which labour scenario would you choose?	Exclusion of costs means WTP was not estimated. Similarly, survival was not included so coefficients are not estimated as utility values (estimate willingness to wait for each attribute, based on MRS for time on labour ward relative to other attributes).
Shackley et al 2001[143]	Peripheral vascular surgery	Surgery	Convenience Access Process	Research focus – methods	Not discernible	Conjoint analysis (DCE). Participants completed eight choice sets. Unlabelled, not a forced choice (could be indifferent – see below). Choice task: Please indicate which vascular service you prefer by ticking box A or box B below. If you prefer both equally, tick both boxes.	Neither cost nor survival is included as an attribute.
Snoek et al 2008[144]	Tetraplegia (paraplegia with upper limb involvement)	Surgery	Process	Innovation	Research grant funding	Conjoint analysis (DCE). Compared profile of treatments (unlabelled). Respondents completed 17 choice sets. Choice task: Select the most preferred treatment scenario with a mouse click.	As neither cost nor survival time are included as attributes it is not possible to estimate WTP or utility values for use in an economic evaluation.
Sung et al 2012[119]	Febrile neutropenia (treatment)	Drug	Convenience	Research focus – preferences	Research grant funding	Labelled DCE – route of administration was the labels. Forced choice experiment. 12 choice sets per respondent. Choice task: would you choose home mouth or hospital IV?	A cost attribute was not included so WTP can't be estimated. Similarly, survival is not included so coefficients can't be interpreted as utility values for use in an EE.
Swan et al 2000[66]	Peripheral vascular disease	Diagnostics	Convenience Process	Innovation	Research grant funding	Modified TTO in which the waiting time for results rather than survival is used as the trade-off. The trade	Present results as quality adjusted days, but could be back calculated to derive disutility values.

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
	imaging (MRI or X-Ray angiogram)					here is time waiting for results compared with test invasiveness/intrusiveness.	
Swan et al 2003[68]	Cerebrovascular imaging (MRI or X-Ray angiogram)	Diagnostics	Convenience Process	Innovation	Research grant funding	Modified TTO in which the waiting time for results rather than survival is used as the trade-off. The trade here is time waiting for results compared with test invasiveness/intrusiveness.	Present results as quality adjusted days, but could be back calculated to derive disutility values.
Swan et al 2006[67]	Breast biopsy (breast cancer diagnosis)	Diagnostics	Anxiety Process	Research focus – methods	Not discernible	Modified TTO in which the waiting time for results rather than survival is used as the trade-off. The trade here is time waiting for results compared with test invasiveness/intrusiveness. Include traditional TTO for assessment of anxiety.	Produce quality adjusted days for each of the biopsy methods (could be reverse estimated to extract the utility value). Also report utility value for anxiety associated with testing.
Torbica et al 2014[122]	Psoriasis	Drug	Convenience	Research focus - preferences	Commercial funding	DCE: unlabelled, force choice. Four choices per participant (27 scenarios blocked into fours – odd!!). Survey administered at clinic by physicians. Choice question not shown. Analysed by mixed logit. WTP by ratio of coeffs (assume fixed cost).	Yes – report WTP for shift from oral to other forms of administration.
Van der Pol and Cairns 1998[145]	Blood transfusions (pre-transfusion test, or transfusion)	Surgery	Convenience	Research focus	Research grant	Conjoint analysis. Conducted one for the pre-transfusion test (18 scenarios – 17 choice sets given one fixed), and one for the transfusion test (16 scenarios – 15 choice sets with one fixed). Not forced choice. Choice task: Definitely prefer A;	WTA values not reported because the coefficient was not significant. In theory could have been calculated.

Authors	Indication	Intervention	Meta-Health Effect	Reason	Funding	Valuation Approach	Useful for Economic Evaluation
						probably prefer A; have no preference; probably prefer B; definitely prefer B.	
Waschbusch et al 2011[118]	ADHD (parent preferences)	Drug	Convenience	Inform policy	Research grant funding	DCE, forced choice (unlabelled). Each respondent completed 27 choice sets. Three options in each choice set. Choice task: If these were your only options, click on the treatment you would prefer for your child with ADHD.	WTP is not estimated as part of this study.
Wilson et al 2014[123]	RRMS	Drug	Convenience	Research focus – preference	Commercial funding.	DCE – forced choice (call it conjoint). Unlabelled, with follow-up intent. Choice question: If these were your only medication options, which would you choose?	No, there is no cost attribute included.
Yasunaga et al 2006[133]	Prostate cancer screening	Screening	Information (authors suggest reassurance as the meta-health effect).	Research focus	Research grant funding	CV (WTP). Men randomly allocated to see scenarios with and without description of PSA (the informed and non-informed groups). Asked WTP in closed form (choice of values presented).	Present mean WTP for the two groups.
Yee et al 2015[151]	Delivery decision – vaginal or caesarean.	Obstetric	Process	Research focus - preferenes	Research grant - funding	TTO and SG.	Produces multiple utility values for the different health states. However utility values are anchored on the preferred delivery mode, not full health.

Abbreviations: ADHD, attention deficit hyperactivity disorder; AE, adverse event; ART, assisted reproductive technology; BC, breast cancer; BSA, body surface area; CA conjoint analysis; CBA, cost-benefit analysis; CF, cystic fibrosis; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; CUA, cost-utility analysis; CV, contingent valuation; CVD, cardio-vascular disease; DCE, discrete choice experiment; EDSS, expanded disability status scale; EQ-5D, EURO-QoL 5 Dimensions; FS, fibroscan; GP, general practitioner; HAART, highly active antiretroviral therapy; HbA1C, glycosylated haemoglobin; HIV, human immunodeficiency virus; HPV, human papillomavirus; ICU, intensive care unit; IUD, intra-uterine device; IV, intravenous; IVF, in-vitro fertilisation; LB, liver biopsy; MRI, magnetic resonance imaging; MRS, marginal rates of substitution; MS, multiple sclerosis; ONJ, osteonecrosis of the jaws; OR, odds

ratios; PC, primary care; PSA, prostate specific antigen; RA, rheumatoid arthritis; RUT, random utility theory; SC, subcutaneous; SF-36, Short Form 36; SF-6D, Short Form 6 Dimensions; SG, standard gamble; SRE, skeletal related event; t.i.d, three times daily; T2DM, type 2 diabetes mellitus; TTO, time trade-off; UK, United Kingdom; USA, United States of America; VAS, visual analogue scale; WTP, willingness to pay.

Appendix 2: Chapter 3 – Copy of CHERE Survey

GP User Survey

Q1 About this Survey
Thank you for agreeing to participate in this online survey. This survey is being conducted by researchers at the Centre for Health Economics Research and Evaluation at the University of Technology Sydney, NSW.

Through this national survey, we are asking people about their experiences when they go to the GP, how they use other health care services, and what factors are important to them when choosing between GPs. Using this information, we want to build a picture of how Australians are using GP services currently and what those services might look like in the future.

Completing the survey should take between 10 and 15 minutes.

The responses that you provide to this survey will be strictly confidential. The information you provide is anonymous and can't be linked back to you. This study is part of a research program approved by the University of Technology Sydney Research Ethics Committee (UTS HREC 2009-143P).

If you have any questions or concerns about this survey, please contact Richard De Abreu Lourenço at: Richard.DeAbreuLourenco@chere.uts.edu.au or the UTS Ethics Secretariat: Research.Ethics@uts.edu.au.

Q2 The following questions are about your health.

Q3 In general, would you say your health is:
- ❍ Excellent
- ❍ Very good
- ❍ Good
- ❍ Fair
- ❍ Poor

Q4 Do you have any of the following ongoing medical conditions (choose as many as apply)?
- ❑ Arthritis
- ❑ Asthma
- ❑ Cancer
- ❑ Chronic obstructive pulmonary disease (COPD) eg. chronic bronchitis or emphysema
- ❑ Chronic pain
- ❑ Depression or another mood disorder
- ❑ Diabetes
- ❑ Heart disease
- ❑ High blood pressure or hypertension
- ❑ Other (please specify) _____
- ❑ No ongoing medical conditions

Q5 The following questions are about the practice you went to for your most recent GP visit.

Q6 Approximately how many GPs work in that practice?
- ❍ Don't know
- ❍ 1 - 2
- ❍ 3 - 5
- ❍ 6 - 10
- ❍ More than 10

Q7 Which of the following other services are also available in that practice or within the same building as the practice (choose as many as apply)?
- ❑ Advice from a dietician/nutritionist
- ❑ Counselling
- ❑ Imaging services (eg. X-rays)
- ❑ Nursing services (eg. vaccinations, diabetes care, wound care, well baby checks)
- ❑ Pharmacy
- ❑ Physiotherapy
- ❑ Podiatry
- ❑ Psychology
- ❑ Pathology/specimen collection (eg.blood tests, urine samples)
- ❑ Other _____
- ❑ No other services
- ❑ Don't know

Q8 Does the practice bulk bill?
- ❍ Yes - sometimes/some patients
- ❍ Yes - always
- ❍ No
- ❍ Don't know

Q9 Is the practice accredited
- ❍ Yes
- ❍ No
- ❍ Don't know

Q10 The following questions all refer to visits to the GP about your own health care.

Q11 Do you usually go to the same practice each time you need to see a GP?
- ❍ Yes
- ❍ No

Q12 In the last 12 months, have you been to more than one practice?
- ❍ Yes
- ❍ No

If No Is Selected, Then Skip To Thinking about the doctor's practice ...

Q13 What were the reasons you went to more than one practice (choose as many as apply)?

- ❑ Availability of the doctors
- ❑ Costs of the visits
- ❑ I was away from home
- ❑ Location (eg. close to where I work)
- ❑ Opening hours were convenient
- ❑ Services provided (eg. I go to different practices for different health problems)
- ❑ To be able to use bulk billing
- ❑ Other _____

Q14 Thinking about the doctor's practice you go to most often, do you usually see the same GP in that practice?

- ❍ Yes
- ❍ No

If Yes Is Selected, Then Skip To How long have you been seeing this GP?If No Is Selected, Then Skip To How many times in the last 12 months ...

Q15 How long have you been seeing this GP?

- ❍ Less than 1 year
- ❍ At least 1 but less than 2 years
- ❍ At least 2 but less than 5 years
- ❍ At least 5 but less than 10 years
- ❍ More than 10 years

Q16 How many times in the last 12 months did you go to a GP?

- ❍ None
- ❍ Once
- ❍ 2 - 3 times
- ❍ 4 - 11 times
- ❍ 12 times or more

Q17 Over the last 12 months, how many home visits did you have from a GP?

- ❍ None
- ❍ One
- ❍ 2 - 3
- ❍ 4 - 11
- ❍ 12 or more

If None Is Selected, Then Skip To Thinking about the last time you were...

Q18 The last time you had a home visit from a GP, were you seen by:

- ❍ Your usual GP
- ❍ Another GP

Q19 Thinking about the last time you were sick or needed care, were you seen by:<?xml:namespace prefix = o ns = "urn:schemas-microsoft-com:office:office" /><o:p></o:p>

- ❍ Your usual GP
- ❍ Another GP in your usual practice
- ❍ A GP at a different practice
- ❍ A doctor at the emergency department
- ❍ Other (please specify) _____

Display This Question:
If How many times in the last 12 months did you go to a GP? None Is Not Selected

Q21 Did you experience any of the following in the last 12 months?

	Yes	Not sure/ Don't remember	No
Test results were not available at the time of your scheduled GP visit.	❍	❍	❍
Medical records were not available at the time of your scheduled GP visit.	❍	❍	❍
You received conflicting information from different GPs or health care professionals.	❍	❍	❍
The GP ordered a medical test that you felt was unnecessary because the test had already been done.	❍	❍	❍
The GP ordered a medical test that you could not afford to have.	❍	❍	❍
The GP ordered a medical test that you decided not to have due to reasons other than cost.	❍	❍	❍
The GP prescribed medication that you could not afford to buy.	❍	❍	❍
The GP pescribed medication that you did not obtain for reasons other than cost.	❍	❍	❍
The GP referred you to a specialist that you could not afford to see.	❍	❍	❍
The GP referred you to a specialist that you decided not to see for reasons other than cost.	❍	❍	❍
The GP did not perform a physical examination even though you felt one was needed.	❍	❍	❍

Q20 How easy or difficult is it to get care in the evenings, on weekends, or public holidays without going to a hospital emergency department?

- ❍ Very easy
- ❍ Somewhat easy
- ❍ Somewhat difficult
- ❍ Very difficult
- ❍ Never needed care in the evenings, weekends, or public holidays
- ❍ Not sure

Q22 The following questions all refer to your most recent visit to the GP.

Q23 Did you have an appointment for your last visit to the GP

- ❍ Yes
- ❍ No

If No Is Selected, Then Skip To How long did it take you to travel to...

Q24 How quickly were you seen by the GP?

- ❍ The same day
- ❍ The next day
- ❍ 2 - 4 days
- ❍ 5 – 7 days
- ❍ More than 7 days

Q25 Was this the appointment time you were after?

- ❍ Yes
- ❍ No

Q26 How long did it take you to travel to the practice?

- ❍ Less than 15 minutes
- ❍ 15 – 29 minutes
- ❍ 30 – 44 minutes
- ❍ 45 – 59 minutes
- ❍ 60 minutes or more

Q27 How did you travel to the practice?

- ☐ Walked
- ☐ Public transport
- ☐ Taxi
- ☐ Family/friend lift
- ☐ Drove myself
- ☐ Other (please specify) _____

Q28 Once you arrived at the practice, how long did you have to wait before you saw the GP?

- ❍ Less than 15 minutes
- ❍ 15 – 29 minutes
- ❍ 30 – 44 minutes
- ❍ 45 – 59 minutes
- ❍ 60 minutes or more

Q29 When you arrived at the practice, were you informed of how long you might have to wait?
- ❍ Yes
- ❍ No

Q30 When you were with the GP, did you feel that the GP:

	Yes	No
Listened to your concerns and needs	❍	❍
Explained things about your condition and any treatment you might need in a way you could understand	❍	❍
Knew you and your medical history well enough	❍	❍
Spent enough time with you	❍	❍
Involved you in the decisions about the treatments available	❍	❍

Q31 How long did you spend with the GP?
- ❍ Less than 5 minutes
- ❍ 5 – 19 minutes
- ❍ 20 – 39 minutes
- ❍ 40 minutes or more

Q32 Which of the following did you discuss with the GP during your visit?
- ❑ Cancer screening
- ❑ Drinking alcohol in moderation
- ❑ Eating healthy food or improving diet
- ❑ Family planning
- ❑ General health checks
- ❑ Increasing physical activity
- ❑ Mental health and wellbeing
- ❑ Reaching a healthy weight
- ❑ Reducing or quitting smoking
- ❑ Safe sexual practices
- ❑ None of the above

Q33 Did you receive care or health advice from someone other than the GP?
- ❍ Yes
- ❍ No

Q34 Were you bulk billed for your visit?
- ❍ Yes
- ❍ No

If Yes Is Selected, Then Skip To End of Block

Q35 How did you pay for your most recent visit?
- ❍ I paid the clinic fee then submitted my own claim to Medicare.
- ❍ I paid the clinic fee but the practice staff submitted my Medicare claim.
- ❍ I paid the clinic fee and submitted the claim to my employer/WorkCover.
- ❍ I don't remember.

Q36 How much did you pay for your most recent visit?

- ❍ $5 - $9
- ❍ $10 – $14
- ❍ $15 - $19
- ❍ $20 - $24
- ❍ $25 - $29
- ❍ $30 - $34
- ❍ $35 - $39
- ❍ $40 - $44
- ❍ $45 - $49
- ❍ $50 - $54
- ❍ $55 - $59
- ❍ $60 - $64
- ❍ $65 - $69
- ❍ $70 - $74
- ❍ $75 - $79
- ❍ $80 - $84
- ❍ Other (specify) _____

Q37 Over the last 12 months have there been any times you needed to visit a GP but didn't?

- ❍ Yes
- ❍ No

If No Is Selected, Then Skip To Which of the following services would...

Q38 What were the reasons you didn't go?

- ❑ I couldn't afford the visit or follow-up care
- ❑ I couldn't afford to take the time off work for the visit
- ❑ The distance I needed to travel to get to the practice
- ❑ I couldn't afford the cost of transport to the GP practice
- ❑ An appointment was not available when I needed to see the GP
- ❑ I was too busy (with work, personal or family responsibilities)
- ❑ Other (please specify) _____

Q39 Which of the following services would you use at your GP practice if they were available (choose as many as apply)?

- ❑ Counselling
- ❑ Advice from a dietician/nutritionist
- ❑ Imaging services (eg X-rays)
- ❑ Nursing services (eg. vaccinations, diabetes care, wound care and well baby checks)
- ❑ Pharmacy
- ❑ Physiotherapy
- ❑ Podiatry
- ❑ Psychology
- ❑ Pathology/specimen collection (eg blood tests, urine samples)
- ❑ Other (please specify): _____
- ❑ None

Q40 Over the past 12 months did you go to an emergency department at a hospital for care?
- Yes
- No

If No Is Selected, Then Skip To Have you heard of the health care hel...

Q41 Was this because you couldn't get an appointment to see a GP?
- Yes
- No

If No Is Selected, Then Skip To Have you heard of the health care hel...

Q42 How many times over the past 12 months did you go to an emergency department at a hospital because you couldn't get an appointment to see a GP?
- Once
- 2 - 3 times
- 4 - 11 times
- 12 times or more

Q43 Have you heard of the health care helpline, HealthDirect Australia?
- Yes
- No

If No Is Selected, Then Skip To Have you heard of the national EHealt...

Q44 How many times over the past 12 months did you call HealthDirect Australia?
- None
- Once
- 2 - 3 times
- 4 - 11 times
- 12 times or more

Q45 Have you heard of the national EHealth medical record keeping system?
- Yes
- No

If No Is Selected, Then Skip To End of Block

Q46 Have you registered to be a part of the EHealth system?
- Yes
- No

Q47 If you had to choose between different GPs, how important would the following factors be to you in making that choice?

Q48 How I arrange to see the GP

	Not at all important 1	2	3	4	Extremely important 5
I can make a same day appointment	❍	❍	❍	❍	❍
I can make an appointment on-line	❍	❍	❍	❍	❍
The practice sends appointment reminders by phone/text/e-mail	❍	❍	❍	❍	❍
I can see the GP of my choice	❍	❍	❍	❍	❍
I can make an appointment to see a GP at a time of day that suits me	❍	❍	❍	❍	❍
The practice provides urgent care out-of-hours (i.e. at night, weekends and holidays)	❍	❍	❍	❍	❍
I can make appointments out-of-hours (i.e. at night, weekends and holidays)	❍	❍	❍	❍	❍

Q49 About getting to see the GP

	Not at all important 1	2	3	4	Extremely important 5
The doctor sees me promptly at the appointed time	❍	❍	❍	❍	❍
The practice staff tell me if the doctor is running late for my appointment when I arrive	❍	❍	❍	❍	❍
The practice is nearby	❍	❍	❍	❍	❍
It costs very little to travel to each appointment	❍	❍	❍	❍	❍
The practice offers alternatives to doctor face-to-face consultations (eg. consultation over the phone, e-mail, text, or video conference)	❍	❍	❍	❍	❍

Q50 How I pay to see the GP

	Not at all important 1	2	3	4	Extremely important 5
The practice offers bulk billing	❍	❍	❍	❍	❍
The practice staff deal with the Medicare forms and I don't need to make a separate claim	❍	❍	❍	❍	❍

Q51 About where I see my doctor

	Not at all important 1	2	3	4	Extremely important 5
The practice is clean and the staff are friendly	❍	❍	❍	❍	❍
There is parking nearby	❍	❍	❍	❍	❍
There is public transport nearby	❍	❍	❍	❍	❍
The waiting room is comfortable	❍	❍	❍	❍	❍
The practice has more than one GP	❍	❍	❍	❍	❍
The practice is part of a larger medical group	❍	❍	❍	❍	❍

Q52 About the medical service I receive

	Not at all important 1	2	3	4	Extremely important 5
The practice is accredited	❍	❍	❍	❍	❍
The consultations with the GP are long enough to deal with my needs	❍	❍	❍	❍	❍
The GP has easy access to my computerised medical records	❍	❍	❍	❍	❍
I can see the same GP each time and she/he knows my medical history	❍	❍	❍	❍	❍
The GP listens and explains the diagnosis and treatment clearly	❍	❍	❍	❍	❍
The GP involves me in discussions about my treatment	❍	❍	❍	❍	❍
The GP gives me sufficient information on my condition and treatment options	❍	❍	❍	❍	❍
The GP offers treatments proven to achieve results	❍	❍	❍	❍	❍
The GP will conduct a thorough	❍	❍	❍	❍	❍

	Not at all important 1	2	3	4	Extremely important 5
physical examination when necessary					
The practice offers home GP visits	❍	❍	❍	❍	❍
The practice offers specialist nurses (such as diabetes care)	❍	❍	❍	❍	❍
The practice offers allied health services (such as physiotherapy)	❍	❍	❍	❍	❍
The practice offers complementary and alternative health care (such as acupuncture)	❍	❍	❍	❍	❍
The practice offers on site pharmacy and pathology services	❍	❍	❍	❍	❍
I can choose whether to see the GP, GP assistant or practice nurse	❍	❍	❍	❍	❍
The GPs in the practice have special expertise (eg. women's health, child health, minor surgery)	❍	❍	❍	❍	❍

Q53 The following questions are about you and your household. The information you provide is strictly confidential and will help us to understand the patterns we see from this survey in GP and health care service use.

Q54 What is your gender?
- ❍ Male
- ❍ Female
- ❍ Other

Q55 Which of the following age brackets do you belong to?
- ❍ 16-19
- ❍ 20-24
- ❍ 25-29
- ❍ 30-34
- ❍ 35-39
- ❍ 40-44
- ❍ 45-49
- ❍ 50-54
- ❍ 55-59
- ❍ 60-64
- ❍ 65-69
- ❍ 70-74
- ❍ 75 and over

Q56 Are you of Aboriginal and/or Torres Strait Islander origin?
- ❍ Yes
- ❍ No

Q57 Where were you born?
- ❍ Australia
- ❍ China
- ❍ England
- ❍ Germany
- ❍ Greece
- ❍ Italy
- ❍ Lebanon
- ❍ Netherlands
- ❍ New Zealand
- ❍ Scotland
- ❍ Vietnam
- ❍ Other _____

Q58 Which language do you mainly speak at home?
- ❍ Arabic
- ❍ Cantonese
- ❍ English
- ❍ German
- ❍ Greek
- ❍ Hindi
- ❍ Italian
- ❍ Mandarin
- ❍ Spanish
- ❍ Vietnamese
- ❍ Other _____

Q59 What is the highest level of education you have completed to date?
- ❍ Primary school only
- ❍ Lower level high school (up to year 10) only
- ❍ Upper level high school (up to year 12) only
- ❍ Post-high school qualification (eg. TAFE, Business School)
- ❍ University degree or higher
- ❍ Other (please specify) _____

Q60 Please indicate which of the following cards you use.
- ❍ Health Care Card
- ❍ Pensioner Concession Card
- ❍ Commonwealth Seniors Health Card
- ❍ Department of Veterans' Affairs Gold or White Card
- ❍ None

Q61 In the last financial year, what was your total household income before taxes (include income from wages/salaries, government benefits, pensions, investments and other incomes that might have been received)?

- Negative or zero Income
- $1 - $9,999 per year ($1 - $189 per week)
- $10,000 - $19,999 per year ($190 - $379 per week)
- $20,000 - $29,999 per year ($380 - $579 per week)
- $30,000 - $39,999 per year ($580 - $769 per week)
- $40,000 - $49,999 per year ($770 - $959 per week)
- $50,000 - $59,999 per year ($960 - $1,149 per week)
- $60,000 - $79,999 per year ($1,150 - $1,529 per week)
- $80,000 - $99,999 per year ($1,530 - $1,919 per week)
- $100,000 - $124,999 per year ($1,920 - $2,399 per week)
- $125,000 - $149,999 per year ($2,400 - $2,879 per week)
- $150,000 - $199,999 per year ($2,880 - $3,839 per week)
- $200,000 or more per year ($3,840 or more per week)
- Declined to answer
- Don't know

Q62 Which of the following best describes your relationship status?

- Married
- Living with a partner
- Single, never married
- Divorced
- Separated
- Widowed
- Declined to answer

Q63 Including you, how many people live in your household?

If Including you, how many peo... Is Equal to 1, Then Skip To Do you provide unpaid care, help or a...

Q64 Of those, how many are under 16 years of age?

Q65 Do you provide unpaid care, help or assistance to someone with a disability, long-term illness or health problem related to old age?

- Yes
- No

If Yes Is Selected, Then Skip To Does this person live in your household?If No Is Selected, Then Skip To What is the post-code of the area i...

Q66 Does this person live in your household?

- Yes
- No

Q67 What is the post-code of the area in which you live?

- Enter your postcode _____
- Declined to answer

Appendix 3: Chapter 3 - Data Frequencies and Levels

Table A 1: Variable Levels and Frequencies

Variable			Respondents	
Name	Category Level	Description	Number	%
Bivariate probit – GP loyal	0	No usual practice	201	8.73
	1	Loyal (One usual practice)	2,102	91.27
Bivariate probit – Multiple practice users	0	Not more than one practice	1,626	70.60
	1	More than one practice	677	29.40
MNP – GP Loyalty	1	Loyal	1,566	68.00
	2	Sometimes loyal	536	23.27
	3	Drifter ("unloyal").	141	6.12
	4	Multiple single practice use	60	2.61
Smoking: Smoking status	1	No	1,620	77.11
	2	Sometimes	78	3.71
	3	Yes	403	19.18
Employment: Employment status	1	Employed	1,412	61.31
	2	Not Employed	428	18.58
	3	Retired	379	16.46
	4	Unknown	84	3.65
Area: Respondent area of residence	1	Major City	1,700	77.34
	2	Inner Regional	344	15.65
	3	Outer Regional	143	6.51
	4	Remote	11	0.5
Income: Household income	1	Low ($0-$39,999)	569	28.8
	2	Medium ($40,000-$79,999)	577	29
	3	High ($80,000-$149,999)	638	32.29
	4	V High (>$150,000)	192	9.72
Gender: Respondent gender[1]	1	Male	1,062	46.27
	2	Female	1,233	53.73
PInsurance: Private health insurance status	0	No	1,620	77.11
	1	Yes	481	22.89
Origin: Respondent place of birth	0	Other	605	26.34
	1	Australia	1,692	73.66
Education: Education attained	1	School Only	733	32.08
	2	University	734	32.12
	3	Vocational and Other	818	35.8
Health: Self-reported health	1	Poor	174	7.56
	2	Fair	781	33.91
	3	Good	882	38.3
	4	Very good	356	15.46
	5	Excellent	110	4.78
Health: Chronic-health issues	0	No chronic health issues	856	37.17
	1	At least one chronic health issue	1,447	62.83
Age: Respondent age	1	16 to 34	603	26.24
	2	35 to 54	954	41.51
	3	55 & Over	741	32.25
Concession: Concession card status	0	No	1,216	52.89

Variable			Respondents	
Name	**Category Level**	**Description**	**Number**	**%**
	1	Yes	1,083	47.11
Attitude Variables: Ratings of importance in choice				
Location: GP practice location	1	Unimportant	490	21.28
	2	Important	1,813	78.72
GPGroup: GP works as part of a group practice	1	Unimportant	1,605	69.69
	2	Important	698	30.31
BBilling: Bulk-billing is available	1	Unimportant	392	17.02
	2	Important	1,911	82.98
Hours: After hours visits are available	1	Unimportant	1,076	46.72
	2	Important	1,227	53.28
GPChoice: Choice of GP is possible	1	Unimportant	340	14.76
	2	Important	1,963	85.24

Notes: 1 In responding to the question on gender, three respondents (1.2%) categorised themselves as "other". These individuals were retained in the sample and classified as males (slightly correcting the initial male, female imbalance in the sample).

Table A 2: Visits to the GP in the Previous 12 Months

Visits	Category Level	Respondent Number	%
None	1	174	7.02
One	2	383	15.46
Two to three	3	1,022	41.26
Four to 11	4	748	30.2
12 or more	5	150	6.06

Appendix 4: Chapter 3 – Supplementary Results

Table A 3: Regression Coefficients – Socio-demographics Only

	Multinomial Probit: Four Outcomes			Bivariate Probit Results	
	Loyal vs Multiple practices	*Sometimes Loyal vs Multiple practices*	*Multiple Single Practice Use vs Multiple practices*	*Usual Practice*	*Multiple Practices*
Smoking: *Non-Smoker*					
Sometimes	-0.218	0.196	-0.202	0.019	0.297
	(0.293)	(0.305)	(0.451)	(0.201)	(0.152)
Smokers	-0.203	-0.046	-0.606*	0.049	0.145
	(0.154)	(0.162)	(0.262)	(0.105)	(0.078)
Unknown	-0.505**	-0.606**	-0.694*	-0.220	0.071
	(0.189)	(0.203)	(0.306)	(0.134)	(0.107)
Employment: *Employed*					
Not Employed	0.103	0.105	0.062	0.053	-0.011
	(0.159)	(0.166)	(0.235)	(0.105)	(0.081)
Retired	0.571*	0.404	-0.085	0.445*	-0.154
	(0.266)	(0.281)	(0.444)	(0.178)	(0.111)
Unknown	0.197	0.077	0.747	-0.133	-0.158
	(0.299)	(0.327)	(0.399)	(0.199)	(0.156)
Area: *Major City*					
Inner Regional	0.511*	-0.019	0.309	0.200	-0.396***
	(0.206)	(0.217)	(0.281)	(0.126)	(0.092)
Outer Regional	0.212	-0.208	-0.140	0.100	-0.270*
	(0.244)	(0.270)	(0.406)	(0.169)	(0.128)
Remote	-0.327	-0.068	-4.346***	-0.061	0.235
	(0.751)	(0.819)	(0.604)	(0.545)	(0.382)
Unknown	0.023	0.252	0.712*	-0.191	0.067
	(0.270)	(0.281)	(0.331)	(0.159)	(0.134)
Income: *Low*					
Medium	-0.191	-0.168	0.056	-0.160	0.030
	(0.184)	(0.194)	(0.294)	(0.126)	(0.089)
High	-0.121	-0.023	0.096	-0.098	0.068
	(0.194)	(0.205)	(0.306)	(0.130)	(0.096)
Very High	-0.163	-0.170	0.148	-0.174	0.006
	(0.246)	(0.259)	(0.378)	(0.170)	(0.126)
Unknown	0.090	-0.062	0.387	-0.093	-0.126
	(0.219)	(0.231)	(0.324)	(0.141)	(0.104)
Gender: *Male*					
Female	-0.170	0.013	-0.282	-0.003	0.145*
	(0.120)	(0.126)	(0.177)	(0.080)	(0.060)
Origin: *Overseas*					
Australia	0.070	-0.040	-0.004	0.028	-0.067
	(0.135)	(0.142)	(0.193)	(0.088)	(0.066)
Health: *Poor*					
Fair	0.247	-0.029	-0.014	0.134	-0.187
	(0.219)	(0.228)	(0.296)	(0.142)	(0.110)
Good	0.114	-0.126	-0.076	0.056	-0.144
	(0.217)	(0.227)	(0.297)	(0.142)	(0.110)
Very Good	-0.041	-0.400	-1.047*	0.109	-0.159
	(0.244)	(0.258)	(0.421)	(0.165)	(0.126)
Excellent	-0.292	-0.329	-0.809	-0.054	0.065

	Multinomial Probit: Four Outcomes			Bivariate Probit Results	
	Loyal vs Multiple practices	*Sometimes Loyal vs Multiple practices*	*Multiple Single Practice Use vs Multiple practices*	*Usual Practice*	*Multiple Practices*
	(0.321)	(0.342)	(0.549)	(0.220)	(0.169)
Age: *16-34*					
35-54	0.701***	0.233	0.463*	0.306***	-0.390***
	(0.131)	(0.137)	(0.193)	(0.088)	(0.070)
>55	1.073***	0.217	0.394	0.547***	-0.662***
	(0.184)	(0.197)	(0.281)	(0.122)	(0.093)
Education: *School Only*					
University	0.071	0.242	0.136	0.036	0.087
	(0.152)	(0.161)	(0.220)	(0.100)	(0.074)
Vocational & Other	-0.028	0.271	-0.002	0.038	0.187*
	(0.152)	(0.161)	(0.226)	(0.100)	(0.076)
Concession: *No*					
Yes	0.099	0.183	0.245	0.034	0.018
	(0.139)	(0.147)	(0.202)	(0.094)	(0.070)
Constant	1.108***	0.728*	-0.616	1.001***	-0.153
	(0.326)	(0.343)	(0.495)	(0.220)	(0.165)
ρ					-0.515***
					(0.040)
n	2,264				
LLHood	-1831.068			-1848.541	
Chi-2	1336.18			211.03	
p-values	0.000			0.000	
Pseudo-R^2	0.084			0.075	

Notes: * p<0.05, ** p<0.01, *** p<0.0001
Standard errors are shown in parentheses.
The omitted (base) level for each independent variable is shown in italics alongside the category name.
Pseudo-R^2 estimated as 1-(L1/L0), where L1 is the log-likelihood of the fitted regression, and L0 is the log-likelihood of the corresponding constant only regression.
Values in parentheses for the Pseudo-R^2 are the LR test statistic estimated as 2*(L1-L0).

Table A 4: Marginal Effects, All Variables

	Multinomial Probit				Bivariate Probit Results			
	Loyal	*Sometimes Loyal*	*Multiple Practice User*	*Multiple Single Practice Use*	*Loyal Only*	*Sometimes Loyal*	*Multiple Practice User*	*Multiple Single Practice Use*
Smoking: *Non Smoker*								
Sometimes	-0.097 (0.058)	0.09 (0.055)	0.011 (0.024)	-0.005 (0.013)	-0.099 (0.056)	0.094 (0.052)	0.011 (0.021)	-0.006 (0.009)
Smokers	-0.03 (0.027)	0.025 (0.025)	0.017 (0.013)	-0.012* (0.005)	-0.039 (0.027)	0.04 (0.024)	0.003 (0.01)	-0.004 (0.005)
Unknown	-0.019 (0.037)	-0.028 (0.031)	0.056* (0.023)	-0.009 (0.007)	-0.042 (0.037)	0.01 (0.03)	0.022 (0.015)	0.011 (0.009)
Employment: *Employed*								
Not Employed	0.004 (0.028)	0 (0.025)	-0.003 (0.014)	-0.001 (0.007)	0.004 (0.027)	-0.001 (0.024)	-0.002 (0.01)	-0.001 (0.006)
Retired	0.065 (0.035)	-0.023 (0.034)	-0.032* (0.014)	-0.01 (0.007)	0.065 (0.035)	-0.021 (0.032)	-0.03** (0.011)	-0.014* (0.006)
Unknown	0.013 (0.053)	-0.032 (0.048)	-0.018 (0.02)	0.037 (0.027)	0.037 (0.049)	-0.058 (0.042)	0.005 (0.02)	0.016 (0.018)
Area: *Major City*								
Inner Regional	0.123*** (0.026)	-0.095*** (0.023)	-0.025* (0.012)	-0.003 (0.007)	0.122*** (0.026)	-0.095*** (0.022)	-0.025** (0.008)	-0.002 (0.006)
Outer Regional	0.1** (0.038)	-0.086* (0.034)	-0.006 (0.019)	-0.008 (0.008)	0.091** (0.039)	-0.075* (0.032)	-0.017 (0.012)	0.001 (0.009)
Remote	-0.044 (0.142)	0.036 (0.134)	0.026 (0.085)	-0.018*** (0.004)	-0.058 (0.143)	0.053 (0.123)	0.008 (0.059)	-0.003 (0.024)
Unknown	-0.054 (0.05)	0.038 (0.047)	-0.009 (0.02)	0.025 (0.019)	-0.029 (0.045)	0.002 (0.044)	0.018 (0.018)	0.008 (0.012)
Income: *Low*								
Medium	-0.015 (0.03)	-0.002 (0.028)	0.011 (0.014)	0.006 (0.007)	-0.014 (0.03)	-0.005 (0.027)	0.012 (0.011)	0.007 (0.007)
High	-0.033 (0.033)	0.023 (0.031)	0.005 (0.015)	0.005 (0.007)	-0.03 (0.032)	0.019 (0.029)	0.009 (0.011)	0.002 (0.006)
Very High	-0.023 (0.043)	0.003 (0.04)	0.013 (0.02)	0.006 (0.01)	-0.022 (0.042)	0 (0.038)	0.014 (0.015)	0.008 (0.01)
Unknown	0.027 (0.033)	-0.03 (0.031)	-0.006 (0.015)	0.01 (0.008)	0.032 (0.033)	-0.04 (0.03)	0.001 (0.011)	0.007 (0.008)
Gender: *Male*								
Female	-0.047* (0.02)	0.041* (0.019)	0.011 (0.009)	-0.004* (0.005)	-0.049* (0.02)	0.044* (0.018)	0.007 (0.007)	-0.002 (0.004)
Origin: *Overseas*								
Australia	0.02 (0.023)	-0.018 (0.021)	-0.002 (0.011)	-0.001 (0.005)	0.017 (0.022)	-0.014 (0.02)	-0.003 (0.008)	0 (0.005)
Education: *School Only*								

	Multinomial Probit				Bivariate Probit Results			
	Loyal	*Sometimes Loyal*	*Multiple Practice User*	*Multiple Single Practice Use*	*Loyal Only*	*Sometimes Loyal*	*Multiple Practice User*	*Multiple Single Practice Use*
University	-0.034 (0.024)	0.042 (0.022)	-0.01 (0.012)	0.003 (0.006)	-0.027 (0.024)	0.032 (0.021)	0 (0.009)	-0.005 (0.006)
Vocational & Other	-0.063* (0.026)	0.072** (0.024)	-0.006 (0.012)	-0.003 (0.006)	-0.058* (0.025)	0.067** (0.022)	0 (0.009)	-0.01 (0.006)
Health: *None*								
One or More	-0.012 (0.022)	0.043* (0.019)	-0.004 (0.01)	-0.027*** (0.007)	-0.019 (0.021)	0.057** (0.018)	-0.018* (0.008)	-0.02*** (0.006)
Age: *16-34*								
35-54	0.145*** (0.026)	-0.095*** (0.025)	-0.05*** (0.014)	0 (0.006)	0.148*** (0.025)	-0.102*** (0.023)	-0.043*** (0.011)	-0.003 (0.005)
>55	0.229*** (0.031)	-0.16*** (0.029)	-0.063*** (0.016)	-0.006 (0.007)	0.23*** (0.031)	-0.162*** (0.028)	-0.062*** (0.013)	-0.007 (0.007)
Concession: *No*								
Yes	-0.009 (0.024)	0.012 (0.022)	-0.01 (0.011)	0.007 (0.006)	0.001 (0.024)	-0.001 (0.021)	0 (0.008)	0 (0.005)
Location:								
Unimportant								
Important	0.016 (0.026)	-0.003 (0.023)	-0.01 (0.013)	-0.003 (0.006)	0.015 (0.025)	-0.004 (0.022)	-0.008 (0.01)	-0.003 (0.006)
GPGroup:								
Unimportant								
Important	-0.038 (0.022)	0.052* (0.021)	-0.009 (0.01)	-0.004 (0.005)	-0.032 (0.022)	0.047* (0.02)	-0.006 (0.008)	-0.009* (0.004)
BBilling:								
Unimportant								
Important	-0.046 (0.026)	0.035 (0.024)	0.025* (0.01)	-0.013 (0.008)	-0.062* (0.025)	0.055* (0.022)	0.01 (0.008)	-0.003 (0.006)
Hours:								
Unimportant								
Important	-0.019 (0.021)	0.011 (0.019)	0.005 (0.01)	0.004 (0.005)	-0.017 (0.021)	0.007 (0.018)	0.007 (0.007)	0.003 (0.005)
GPChoice:								
Unimportant								
Important	0.079* (0.031)	-0.053 (0.028)	-0.019 (0.016)	-0.007 (0.007)	0.075* (0.03)	-0.044 (0.027)	-0.025* (0.012)	-0.006 (0.007)

Notes: * p<0.05, ** p<0.01, *** p<0.0001
The omitted (base) level for each independent variable is shown in italics alongside the category name.
Standard errors are shown in parentheses.

	Multinomial Probit				Bivariate Probit Results			
	Loyal	*Sometimes Loyal*	*Multiple Practice User*	*Multiple Single Practice Use*	*Loyal Only*	*Sometimes Loyal*	*Multiple Practice User*	*Multiple Single Practice Use*

Margins are the discrete change from the base levels.

Appendix 5: Chapter 3 – Sensitivity Analyses

As can be observed from the estimated and observed relationship category memberships (see Table 17), the majority of individuals were classified as either loyal (68.0%) or reported being loyal but used multiple practices (23.3%). This means that the proportion of individuals reporting multiple practice use without being loyal to a practice was relatively low (8.7%). Within the main analysis this has been modelled as two categories; multiple practice users (6.1%) and multiple single practice use (2.1%). The low number of individuals in this latter category might explain why a four level MNP does not perform as well as a bivariate probit. It is possible therefore that structuring the data for the MNP analysis as a three level outcome might improve its explanatory power and performance compared with the bivariate probit. This was tested in a sensitivity analysis in which all multiple practice users who did not report being loyal to a practice were classified as a single category. The results of this sensitivity analysis presented in Table A 5.

Table A 5: MNP Four and Three Level Outcomes

	Multinomial Probit: Four Outcomes			**Multinomial Probit: Three Outcomes**	
	Loyal vs Multiple practices	*Sometimes vs Multiple practices*	*Disenfranch. vs Multiple practices*	*Loyal vs Multiple practices*	*Sometimes vs Multiple practices*
Smoking: *Non-Smoker*					
Sometimes	-0.276	0.128	-0.259	-0.186	0.216
	(0.295)	(0.305)	(0.476)	(0.284)	(0.291)
Smokers	-0.225	-0.103	-0.604*	-0.065	0.057
	(0.155)	(0.163)	(0.266)	(0.145)	(0.154)
Unknown	-0.515**	-0.585**	-0.785*	-0.306	-0.380
	(0.189)	(0.203)	(0.309)	(0.181)	(0.196)
Employment: *Employed*					
Not Employed	0.037	0.029	-0.004	0.023	0.014
	(0.160)	(0.167)	(0.238)	(0.147)	(0.156)
Retired	0.531*	0.358	0.029	0.570*	0.397
	(0.268)	(0.284)	(0.446)	(0.241)	(0.259)
Unknown	0.212	0.086	0.787*	-0.103	-0.227
	(0.297)	(0.326)	(0.399)	(0.265)	(0.296)
Area: *Major City*					
Inner Regional	0.511*	-0.019	0.213	0.440*	-0.090
	(0.209)	(0.219)	(0.284)	(0.176)	(0.190)
Outer Regional	0.213	-0.241	-0.242	0.280	-0.175
	(0.248)	(0.271)	(0.419)	(0.229)	(0.256)
Remote	-0.289	-0.122	-8.630***	-0.189	-0.031
	(0.752)	(0.799)	(0.621)	(0.721)	(0.770)
Unknown	0.004	0.196	0.544	-0.237	-0.040
	(0.270)	(0.283)	(0.339)	(0.230)	(0.245)
Income: *Low*					
Medium	-0.148	-0.132	0.095	-0.192	-0.176
	(0.181)	(0.192)	(0.291)	(0.169)	(0.181)
High	-0.105	0.017	0.123	-0.146	-0.026
	(0.193)	(0.204)	(0.303)	(0.178)	(0.190)

	Multinomial Probit: Four Outcomes			Multinomial Probit: Three Outcomes	
	Loyal vs Multiple practices	*Sometimes vs Multiple practices*	*Disenfranch. vs Multiple practices*	*Loyal vs Multiple practices*	*Sometimes vs Multiple practices*
Very High	-0.179	-0.133	0.069	-0.233	-0.189
	(0.244)	(0.256)	(0.369)	(0.228)	(0.242)
Unknown	0.116	-0.025	0.371	-0.029	-0.170
	(0.215)	(0.228)	(0.316)	(0.191)	(0.206)
Gender: *Male*					
Female	-0.191	0.011	-0.244	-0.122	0.080
	(0.117)	(0.124)	(0.181)	(0.110)	(0.117)
Origin: *Overseas*					
Australia	0.050	-0.038	-0.002	0.061	-0.026
	(0.134)	(0.141)	(0.194)	(0.121)	(0.130)
Health: None					
One or More	0.034	0.190	-0.631***	0.254*	0.410***
	(0.124)	(0.130)	(0.184)	(0.112)	(0.120)
Education: *School Only*					
University	0.063	0.258	0.184	0.005	0.201
	(0.151)	(0.161)	(0.223)	(0.137)	(0.148)
Vocational & Other	-0.027	0.302	-0.033	-0.001	0.327*
	(0.149)	(0.158)	(0.227)	(0.137)	(0.148)
Age: *16-34*					
35-54	0.661***	0.171	0.416*	0.542***	0.051
	(0.128)	(0.135)	(0.200)	(0.122)	(0.129)
>55	0.984***	0.127	0.410	0.888***	0.031
	(0.190)	(0.202)	(0.295)	(0.173)	(0.187)
Concession: *No*					
Yes	0.094	0.146	0.303	-0.007	0.046
	(0.140)	(0.147)	(0.209)	(0.130)	(0.137)
Location: Unimportant					
Important	0.127	0.094	0.000	0.130	0.097
	(0.153)	(0.161)	(0.214)	(0.138)	(0.147)
GPGroup: *Unimportant*					
Important	0.051	0.269*	-0.019	0.071	0.289*
	(0.130)	(0.136)	(0.198)	(0.120)	(0.127)
BBilling: *Unimportant*					
Important	-0.388*	-0.203	-0.627**	-0.148	0.036
	(0.178)	(0.184)	(0.230)	(0.148)	(0.157)
Hours: *Unimportant*					
Important	-0.082	-0.020	0.056	-0.104	-0.042
	(0.125)	(0.131)	(0.190)	(0.116)	(0.122)
GPChoice: *Unimportant*					
Important	0.311	0.032	-0.007	0.343*	0.064
	(0.169)	(0.174)	(0.224)	(0.150)	(0.157)
Constant	1.265***	0.528	-0.032	0.792**	0.057
	(0.308)	(0.326)	(0.475)	(0.281)	(0.303)
n	2,264				
LLHood	-1814.439			-1715.972	
Chi-2	3850.210			242.110	
p-values	0.000			0.000	

	Multinomial Probit: Four Outcomes			Multinomial Probit: Three Outcomes	
	Loyal vs Multiple practices	*Sometimes vs Multiple practices*	*Disenfranch. vs Multiple practices*	*Loyal vs Multiple practices*	*Sometimes vs Multiple practices*
Pseudo-R^2	0.092 (367.295)			0.085 (319.172)	
AIC	3796.877			3543.943	
BIC	4277.768			3864.537	
Notes:	* p<0.05, ** p<0.01, *** p<0.0001 Standard errors are shown in parentheses. The omitted (base) level for each independent variable is shown in italics alongside the category name. Pseudo-R^2 estimated as 1-(L1/L0), where L1 is the log-likelihood of the fitted regression, and L0 is the log-likelihood of the corresponding constant only regression. Values in parentheses for the Pseudo-R^2 are the LR test statistic estimated as 2*(L1-L0).				

The multinomial outcome could also have been modelled as an MNL, which would potentially have been more stable. The test of IIA was conducted for the mlogit specification by restricting the dependent variable to exclude *j*=2 (Sometimes loyal). A Hausman test was conducted comparing this specification with the full specification reported previously. The resulting χ^2 (56) = 6.64, p=1.00, indicating that the null supporting IIA could not be rejected. Indeed, running the multinomial analysis as an MNL produced qualitatively the same results as the MNP, without violating IIA (see Table A 6).

Table A 6: MNL Specification

	Multinomial Probit: Four Outcomes			Multinomial Logit: Four Outcomes		
	Loyal vs Multiple practices	*Sometimes Loyal vs Multiple practices*	*Disenfran. vs Multiple practices*	*Loyal vs Multiple practices*	*Sometimes Loyal vs Multiple practices*	*Disenfran. vs Multiple practices*
Smoking: *Non-Smoker*						
Sometimes	-0.276	0.128	-0.259	-0.366	0.129	-0.628
	(0.295)	(0.305)	(0.476)	(0.463)	(0.480)	(0.926)
Smokers	-0.225	-0.103	-0.604*	-0.345	-0.184	-1.080*
	(0.155)	(0.163)	(0.266)	(0.249)	(0.263)	(0.529)
Unknown	-0.515**	-0.585**	-0.785*	-0.806**	-0.905**	-1.276*
	(0.189)	(0.203)	(0.309)	(0.272)	(0.300)	(0.555)
Employment: *Employed*						
Not Employed	0.037	0.029	-0.004	0.076	0.070	0.086
	(0.160)	(0.167)	(0.238)	(0.248)	(0.263)	(0.446)
Retired	0.531*	0.358	0.029	0.962	0.761	-0.023
	(0.268)	(0.284)	(0.446)	(0.503)	(0.530)	(0.939)
Unknown	0.212	0.086	0.787*	0.324	0.138	1.405*
	(0.297)	(0.326)	(0.399)	(0.448)	(0.506)	(0.684)
Area: *Major City*						
Inner Regional	0.511*	-0.019	0.213	0.813*	0.142	0.574
	(0.209)	(0.219)	(0.284)	(0.361)	(0.382)	(0.535)
Outer	0.213	-0.241	-0.242	0.245	-0.328	-0.369

	Multinomial Probit: Four Outcomes			**Multinomial Logit: Four Outcomes**		
	Loyal vs Multiple practices	*Sometimes Loyal vs Multiple practices*	*Disenfran. vs Multiple practices*	*Loyal vs Multiple practices*	*Sometimes Loyal vs Multiple practices*	*Disenfran. vs Multiple practices*
Regional						
	(0.248)	(0.271)	(0.419)	(0.395)	(0.442)	(0.837)
Remote	-0.289	-0.122	-8.630***	-0.434	-0.197	-10.242***
	(0.752)	(0.799)	(0.621)	(1.148)	(1.203)	(1.267)
Unknown	0.004	0.196	0.544	0.035	0.266	0.890
	(0.270)	(0.283)	(0.339)	(0.430)	(0.452)	(0.580)
Income: *Low*						
Medium	-0.148	-0.132	0.095	-0.197	-0.176	0.292
	(0.181)	(0.192)	(0.291)	(0.292)	(0.311)	(0.563)
High	-0.105	0.017	0.123	-0.140	0.008	0.251
	(0.193)	(0.204)	(0.303)	(0.306)	(0.325)	(0.577)
Very High	-0.179	-0.133	0.069	-0.230	-0.172	0.226
	(0.244)	(0.256)	(0.369)	(0.379)	(0.401)	(0.679)
Unknown	0.116	-0.025	0.371	0.162	-0.006	0.736
	(0.215)	(0.228)	(0.316)	(0.342)	(0.367)	(0.578)
Gender: *Male*						
Female	-0.191	0.011	-0.244	-0.272	-0.022	-0.400
	(0.117)	(0.124)	(0.181)	(0.187)	(0.200)	(0.339)
Origin: *Overseas*						
Australia	0.050	-0.038	-0.002	0.054	-0.057	-0.033
	(0.134)	(0.141)	(0.194)	(0.214)	(0.228)	(0.355)
Health: None						
One or More	0.034	0.190	-0.631***	0.039	0.260	-1.222***
	(0.124)	(0.130)	(0.184)	(0.193)	(0.206)	(0.362)
Education: *School Only*						
University	0.063	0.258	0.184	0.126	0.373	0.308
	(0.151)	(0.161)	(0.223)	(0.243)	(0.262)	(0.423)
Vocational & Other	-0.027	0.302	-0.033	0.008	0.425	-0.048
	(0.149)	(0.158)	(0.227)	(0.236)	(0.255)	(0.423)
Age: *16-34*						
35-54	0.661***	0.171	0.416*	0.924***	0.338	0.706
	(0.128)	(0.135)	(0.200)	(0.197)	(0.210)	(0.367)
>55	0.984***	0.127	0.410	1.401***	0.358	0.680
	(0.190)	(0.202)	(0.295)	(0.319)	(0.344)	(0.572)
Concession: *No*						
Yes	0.094	0.146	0.303	0.152	0.217	0.584
	(0.140)	(0.147)	(0.209)	(0.219)	(0.233)	(0.387)
Location: Unimportant						
Important	0.127	0.094	0.000	0.187	0.139	0.075
	(0.153)	(0.161)	(0.214)	(0.244)	(0.258)	(0.400)
GPGroup: *Unimportant*						
Important	0.051	0.269*	-0.019	0.111	0.383	-0.077
	(0.130)	(0.136)	(0.198)	(0.208)	(0.219)	(0.377)
BBilling: *Unimportant*						
Important	-0.388*	-0.203	-0.627**	-0.701*	-0.475	-1.231**
	(0.178)	(0.184)	(0.230)	(0.308)	(0.320)	(0.435)

	Multinomial Probit: Four Outcomes			Multinomial Logit: Four Outcomes		
	Loyal vs Multiple practices	*Sometimes Loyal vs Multiple practices*	*Disenfran. vs Multiple practices*	*Loyal vs Multiple practices*	*Sometimes Loyal vs Multiple practices*	*Disenfran. vs Multiple practices*
Hours: *Unimportant*						
Important	-0.082	-0.020	0.056	-0.100	-0.022	0.129
	(0.125)	(0.131)	(0.190)	(0.200)	(0.211)	(0.363)
GPChoice: *Unimportant*						
Important	0.311	0.032	-0.007	0.445	0.107	-0.035
	(0.169)	(0.174)	(0.224)	(0.261)	(0.270)	(0.399)
Constant	1.265***	0.528	-0.032	1.800***	0.881	-0.106
	(0.308)	(0.326)	(0.475)	(0.473)	(0.508)	(0.858)
n	2,264					
LLHood	-1814.439			-1813.800		
Chi-^{2}	3850.210			855.200		
p-values	0.000			0.000		
	0.092			0.076		
Pseudo-R^2	(367.295)			(298.766)		
AIC	3796.877			3795.6		
BIC	4277.768			4276.491		

Abbreviations: AIC denotes Akaike Information Criterion, BIC Bayesian Information Criterion.

Notes: * p<0.05, ** p<0.01, *** p<0.0001

Standard errors are shown in parentheses.

The omitted (base) level for each independent variable is shown in italics alongside the category name.

Pseudo-R^2 estimated as 1-(L1/L0), where L1 is the log-likelihood of the fitted regression, and L0 is the log-likelihood of the corresponding constant only regression. Values in parentheses for the Pseudo-R^2 are the LR test statistic estimated as 2*(L1-L0).

The results for the sensitivity analysis on the treatment of missing data are presented in Table A 7.

Table A 7: Effect of missing observations on probabilities

	Base	Lower Bounds		Upper Bounds		Complete Case	
Bivariate Probit Probabilities	*Mean (s.d)*	*Mean (s.d)*	*Diff in Means*	*Mean (s.d)*	*Diff in Means*	*Mean (s.d)*	*Diff in Means*
Loyal	0.682 (0.137)	0.682 (0.136)	-0.00007	0.682 (0.136)	-0.000087	0.687 (0.129)	-0.0050009
Sometimes Loyal	0.23 (0.099)	0.23 (0.098)	0.00011	0.23 (0.098)	0.000165	0.235 (0.096)	-0.0044161
Multiple Practice User	0.062 (0.045)	0.062 (0.044)	-0.00009	0.062 (0.045)	-0.000132	0.058 (0.039)	0.0042766
Multiple User of Single Practice	0.026 (0.019)	0.026 (0.018)	0.00005	0.026 (0.017)	0.000054	0.021 (0.014)	0.0051404

Notes: Differences in means are the difference from the values for the base regression.

Number of observations for the complete case was 1,710 compared with 2,264 for all other analyses.

Abbreviation: s.d., standard deviation.

The results of the sensitivity analysis on the effect of limiting the sample to those individuals who visited a GP practice two or more times within the 12 months prior to the CHERE Survey are presented in Table A 8.

Table A 8: Bivariate probit comparing sample based on visit numbers

	Base Result: More than One Visit		**More than Two Visits**	
	Usual Practice	*Multiple Practices*	*Usual Practice*	*Multiple Practices*
Smoking: *Non-Smoker*				
Sometimes	-0.035	0.297	-0.133	0.331
	(0.204)	(0.152)	(0.240)	(0.172)
Smokers	0.009	0.126	-0.080	0.117
	(0.107)	(0.078)	(0.117)	(0.085)
Unknown	-0.209	0.094	-0.347*	0.125
	(0.134)	(0.107)	(0.143)	(0.117)
Employment: *Employed*				
Not Employed	0.018	-0.009	0.021	0.005
	(0.106)	(0.081)	(0.116)	(0.087)
Retired	0.389*	-0.155	0.319	-0.107
	(0.178)	(0.112)	(0.191)	(0.118)
Unknonwn	-0.127	-0.164	0.127	-0.135
	(0.197)	(0.158)	(0.231)	(0.173)
Area: *Major City*				
Inner Regional	0.222	-0.392***	0.241	-0.426***
	(0.126)	(0.092)	(0.143)	(0.098)
Outer Regional	0.120	-0.288*	0.068	-0.219
	(0.167)	(0.129)	(0.185)	(0.140)
Remote	-0.033	0.170	-0.019	0.106
	(0.530)	(0.381)	(0.548)	(0.388)
Unknown	-0.164	0.057	-0.217	0.030
	(0.161)	(0.134)	(0.177)	(0.146)
Household Income: *Low*				
Medium	-0.141	0.020	-0.162	0.014
	(0.126)	(0.089)	(0.138)	(0.097)
High	-0.088	0.082	-0.102	0.084
	(0.131)	(0.096)	(0.145)	(0.105)
Very High	-0.162	0.043	-0.121	0.143
	(0.168)	(0.126)	(0.189)	(0.141)
Unkown	-0.063	-0.122	-0.019	-0.099
	(0.140)	(0.103)	(0.163)	(0.112)
Gender: *Male*				
Female	-0.036	0.155*	0.009	0.097
	(0.080)	(0.060)	(0.091)	(0.066)
Origin: *Overseas*				
Australia	0.024	-0.052	-0.028	-0.061
	(0.088)	(0.067)	(0.102)	(0.072)
Chronic Health Issues: None				
One or More	0.260**	0.118	0.207*	0.015
	(0.080)	(0.064)	(0.091)	(0.071)
Education: *School Only*				

	Base Result: More than One Visit		More than Two Visits	
	Usual Practice	*Multiple Practices*	*Usual Practice*	*Multiple Practices*
University	0.034	0.101	0.052	0.089
	(0.101)	(0.074)	(0.110)	(0.080)
Vocational & Other	0.070	0.203**	0.152	0.182*
	(0.099)	(0.076)	(0.112)	(0.082)
Age: *16-34*				
35-54	0.278**	-0.399***	0.352***	-0.407***
	(0.088)	(0.070)	(0.100)	(0.078)
>55	0.474***	-0.658***	0.617***	-0.693***
	(0.125)	(0.095)	(0.141)	(0.103)
Concession Card: *No*				
Yes	0.005	-0.003	0.008	-0.094
	(0.096)	(0.071)	(0.108)	(0.080)
GP Location: Unimportant				
Important	0.078	-0.034	-0.028	0.003
	(0.101)	(0.075)	(0.119)	(0.082)
GP Group Practice: *Unimportant*				
Important	0.118	0.121	0.150	0.148*
	(0.089)	(0.064)	(0.098)	(0.070)
Bulk-Billing: *Unimportant*				
Important	-0.052	0.203*	-0.072	0.241**
	(0.106)	(0.082)	(0.123)	(0.092)
After Hours Appt.: *Unimportant*				
Important	-0.070	0.043	-0.090	-0.004
	(0.085)	(0.062)	(0.097)	(0.067)
Choice of GP: *Unimportant*				
Important	0.204	-0.198*	0.185	-0.154
	(0.107)	(0.085)	(0.129)	(0.098)
Constant	0.810***	-0.418**	0.940***	-0.268
	(0.205)	(0.159)	(0.235)	(0.178)
ρ	-0.535***		-0.651***	
	(0.040)		(0.040)	
n	2,264		1,884	
LLHood	-1831.381		-1483.626	
Chi-2	251.500		222.14	
p-values	0.000		0.0000	
Pseudo-R^2	0.083 (333.410)		0.092 (302.130)	
AIC	3776.762		3081.253	
BIC	4103.08		3397.099	

Notes: * p<0.05, ** p<0.01, *** p<0.0001
Standard errors are shown in parentheses.
The omitted (base) level for each independent variable is shown in italics alongside the category name.
Pseudo-R2 estimated as 1-(L1/L0), where L1 is the log-likelihood of the fitted regression, and L0 is the log-likelihood of the corresponding constant only regression. Values in parentheses for the Pseudo-R2 are the LR test statistic estimated as 2*(L1-L0).

Appendix 6: Chapter 5 & 6 - Process for Developing the Online Survey

The following procedure was used to generate the survey:

1. The DCE design was imported from Ngene into Excel.
2. If statements were used to generate text versions of the attribute levels. Separate versions were generated for each of the three framing sets within the survey (Base, Attribute, No-Efficacy Difference).
3. A mail merge document was created in Word with the colour format, attribute text, font size and layout as required for the final survey.
4. Separate mail merge documents were initiated for each of the framing sets. Each Word document contained all 48 scenarios, labelled with the scenario number, block (1-12) and frame.
5. The icon-arrays for efficacy and safety were imported into the formatted survey documents to complete the scenario formats.
6. All merged documents with the loaded icon-arrays were individually checked to ensure that the correct icon-arrays had been loaded into the relevant attribute and corresponding level.
7. Using Word for Macintosh, each table (choice scenario) was copied from Word individually and pasted as a picture into a new document. Each was then saved as an individual picture using the "save as picture" function, type ".png".
8. Each picture was saved with the frame and scenario number in the name.
9. The initial pictures were loaded into Qualtrics to test the size and look of the saved scenarios once loaded. This revealed that there were excessive white spaces around the scenarios that should be cropped within Word before being saved as a picture as they affected the size of the pictures within Qualtrics.
10. Subsequently, the scenarios were extracted from the Word documents one by one and saved as pictures, resizing as required to remove any excessive white spaces.
11. The pictures were then loaded into Qualtrics using the software import function – 144 pictures were loaded; being 48 scenarios each for three frames.

Within Qualtrics the survey was divided such that each frame was presented in four blocks (as per the Ngene design) of 12 scenarios. The allocation of scenarios to blocks was as per the design generated within Ngene. The scenario pictures were individually loaded into the relevant blocks, being sure to include the block number, frame and scenario number as a question label to be exported with the survey responses. Once loaded, all questions were checked to ensure that the survey preserved the original design.

The survey flow was as follows:

1. All respondents saw the survey instructions, including the ethics statement.
2. All respondents saw the background health vignette.
3. Respondents were randomised to one of 12 blocks as follows; 'Base' frame, blocks 1-4; 'Attribute' frame, blocks 5-8; No Efficacy-Difference frame, blocks 9-12. Each block consists of the entire 48 choice sets. That is, each framing set replicates the overall design of 48 scenarios so that each scenario occurs three times across the entire survey.

4. Choice options were unlabelled, and were not randomised by presentation to left or right (Option A always appeared on the left, Option B always appeared on the right of the screen).
5. All respondents completed the questions regarding attribute importance and requesting demographic information.

Screenshots of the final survey for one of the blocks are provided as Appendix 15.

Appendix 7: Chapter 5 - DCE Experimental Design and Survey Generation

Overview

The process for the generation of the design for the discrete choice experiment was as follows:

1. The design was generated in Ngene.[229]
2. Given the focus of this research on investigating the relative value between meta-health effects and health effects, the design was generated to maximise WTP efficiency (see below). This necessitates the use of some information on the expected coefficient values (priors) in order to generate the design.
3. Priors for the pilot study were obtained from the literature (see main text). Priors for the final survey were obtained from the pilot.
4. Ngene generated designs were manually checked for scenario dominance, and altered if required.
5. The efficiency of dominance corrected designs was re-tested by importing the altered design into Ngene for evaluation of the efficiency levels.
6. Comparisons of efficiency are provided using d-error, s-estimate (an estimate of the required sample size), and WTP efficiency.[229,346]

Generating the Design

The design was generated using the attributes and levels described in the main text, and as replicated in Table A 9. Priors for the design of the pilot survey were sourced from the literature.[90,223] For the purposes of the initial design, two aspects of the prior values were considered important: the sign and the rankings between coefficients (rather than their absolute values). The coefficient values used as priors in the syntax within Ngene to produce the DCE design were rescaled to produce smaller values that preserved the sign and rank order between coefficients.

Design coding within Ngene requires that categorical variables are treated as either dummy coded or effects coded. Within effects codes, the base variable is assigned a value of -1, thereby limiting the potential for confounding of attribute level effects with overall mean coefficient effects.[347] Typically, effects coding is recommended for use in DCEs conducted in labelled experiments, or which include a constant term in the analysis, so as to reduce the potential for confounding of the coefficient estimates with the overall mean effect, and to allow for the derivation of coefficient estimates for all attribute levels, including the base level.[189,348] Effects coding was therefore specified in the design within Ngene to avoid potential confounding of coefficient estimates and overall mean effects in the event that a constant term was specified in the final attribute analysis, and to clarify the interpretation of the coefficient estimates. Mode and frequency were entered as effects codes (to reflect their categorical nature), while efficacy, safety and cost were entered as continuous variables (coded as linear variables). Attribute coding is provided in Table A 9.

A WTP efficiency design was estimated in Ngene. The design included two restrictions on the levels of mode and frequency:

1. Where mode was a 15 minute IV infusion, frequency was restricted to exclude the possibility of including twice a day administration.
2. Where mode was a tablet, frequency was restricted to exclude the possibility of allowing monthly administration.

Restrictions were incorporated using the constraints function within Ngene.[229] The inclusion of restrictions in Ngene results in the loss of level balance within the design with respect to the attributes of mode and frequency. A total of 48 rows were used, in four blocks (incorporated into the design), specifying for Ngene to minimise the chance of dominant alternatives across choice sets.

The procedure for testing the efficiency parameters of the designs was staged as follows:

1. The initial design produced within Ngene was exported (as formatted comparative scenarios) into Excel for assessment for scenario dominance. This also considered whether or not scenarios were likely to be subject to extended dominance. That is, within the 'No-Efficacy Difference' frame it is assumed that *Efficacy* is the same between options. Thus scenarios in which alternatives differed only on the basis of *Efficacy*, or the difference in *Efficacy* was the only factor which prevented one alternative dominating the other, were considered to be subject to extended dominance since in the 'No-Efficacy Difference' frame one alternative would dominate the other.
2. For scenarios exhibiting extended dominance, one attribute level was modified to remove the potential for dominance. The resulting rebalanced design was re-evaluated in Ngene to assess the impact on the efficiency estimates of the design modification.[227]
3. The efficiency of the 'No-Efficacy Difference' frame was also assessed re-evaluated within Ngene. This is because the design for that frame was constructed by essentially removing *Efficacy* from the design produced by Ngene; by setting *Efficacy* to be the same across the two options, it removes it from the design.

Re-evaluating design efficiency within Ngene is possible by asking the program to evaluate the design using a pre-specified choice set design (in this case those modified from the initial choice set design), rather than choosing a choice set using the design generation algorithm within Ngene.

Table A 9: Design Parameters and Priors

Domain	Attributes	Levels	Coding for Attributes	Pilot Survey Prior	Final Survey Priors	
					Dummy Coded	Effects Coded
Mode	The treatment is: (3)	A 15 minute intravenous infusion.	0	-0.41	-0.92	-0.55
		An injection.	1	-0.09	-0.19	0.18
		A tablet.	2 (base)			
Frequency	You take the	Twice a day (*tablets*	0	-0.61	-0.65	-0.47

Domain	Attributes	Levels	Coding for Attributes	Pilot Survey Prior	Final Survey Priors	
					Dummy Coded	Effects Coded
	treatment: (3)	*only*)				
		Weekly	1	-0.12	-0.14	0.04
		Fortnightly	2	-0.45	0.07	0.25
		Monthly (*not tablets*)	3 (base)			
Efficacy	The chance that the treatment will control your condition, allowing you to do your usual activities is: (4)	Two in ten Four in ten Six in ten Seven in ten	2,4,6,7	0.33	0.08	0.08
Safety	The chance that you need to stop taking your medicine because you experience side effects is: (3)	One in one hundred Five in one hundred Ten in one hundred	1,5,10	-0.3	-0.07	-0.07
Cost (burden)	This treatment costs $2,000 per month. You pay: (4)	Nothing, the government pays all the costs. $1,000 and the government pays $1,000. $1,500 and the government pays $500. All of the costs.	0,1000,1500,2000	-0.003	n.a.	n.a.
	The cost to you per month of treatment is:	Nothing $40 per month $250 per month $500 per month	0,40,250,500	n.a.	-0.001	-0.001
	The cost to the government per person, per month of treatment is:	Nothing $500 per month $1,500 per month $3,000 per month	0,500,1500,3000	n.a.	-0.001	-0.001
Mode*Freq				0.05	n.a.	n.a.
Mode*Own OOP				n.a.	-0.0009	-0.0005
Own OOP* Govt.Cost				n.a.	0.00001	0.00001

Abbreviation: n.a., not applicable; Govt, government; OOP, out-of-pocket.

Design efficiency estimates: the Pilot Design

A resulting design was produced and selected based on when the system could no longer generate 1,000 iterations without improving efficiency. Overall, the design produced was balanced across attribute levels, with the exception of the attribute for frequency (due to the constraint imposed). The resulting estimates of efficiency were 203,340.39, 220.11 and

0.020 for the WTP efficiency, s-estimates and d-errors respectively. The WTP efficiency represents the estimated aggregate WTP (aggregating the mWTP across all attributes, given the pre-specified priors, using OOP as the numeraire and taking into account the Ngene generated distribution of choice probabilities), and essentially minimises the variance associated with taking the ratio of two coefficient estimates (in this case each attribute coefficient expressed as a ratio over OOP).[229]

The s-estimate is the minimum sample size required given the estimated t-statistics associated with the choice coefficients, pre-specified priors and distribution of choice probabilities.[346,349] An s-estimate is produced for each choice coefficient estimate, with the largest required estimate stated as the minimum required for that design.[229] In this particular design the largest s-estimate (220.11) was that associated with the interaction coefficient between mode and frequency of administration (as a single continuous variable), with all other s-estimates being less than 50. The d-error is estimated as the determinant of the asymptotic variance covariance matrix, scaled for the number of parameters being estimated.[349] For both the s-estimate and the d-error, smaller values are desirable as indicative of a more efficient design.[346,349]

The resulting design, formatted as choice scenarios, was imported into Excel for checking of scenario dominance. That is, while the option for alternative dominance was enacted in Ngene, the resulting choice scenarios were checked given that the use of constraints in the design might have resulted in some dominant scenario pairings. The design included one choice set which could be considered weakly dominated; one alternative was superior to the other in at least two attributes and there was overlap in the remaining three. Similarly, there were an additional three choice sets that would have been weakly dominated in the No-Efficacy Difference framing set.

While each of these four weakly dominated scenarios occurred in a different block, it was felt that to produce the most from the data (in terms of information gathered from a small sample size), the levels of the cost attribute (which overlapped in each scenario) would be modified so there was no longer a dominant alternative. The adjustments applied are summarised in Table A 10.

Table A 10: Adjustments to Address Dominance in Pilot Survey Design

Block	Scenario	Alternative	Cost Attribute		
			From	To	Compared with:
1	30	B	0	1,000	0
2	34	B	2,000	0	2,000
3	11	A	0	1,500	0
4	36	A	1,500	1,000	1,500

This would be expected to have some impact on the correlation structure of the data, and potentially its efficiency. However, it was felt that this was a reasonable adjustment to perform in the context of the pilot phase of the study. The revised choice design was imported into Ngene for re-evaluation of the efficiency estimates. This resulted in a WTP

estimate of 225,645.14, an s-estimate of 230.02 and d-error of 0.019, all marginally higher (indicating a loss of efficiency according to the s-estimate and d-error) than those of the original design. Similarly, the design for the 'No-Efficacy Difference' frame resulted in WTP efficiency of 274,934.00, an s-estimate of 280.21 and d-error of 0.020. This apparent loss in efficiency is due to the removal from the design of *Efficacy* for which there was a high degree of efficiency, as measured by the s-estimate (2.86) in the previous design evaluation, as compared with the majority of the remaining attributes (in the comparable design the s-estimate for *Injection*, the second level in the mode of administration, was 52.53). Thus removing an attribute from the design for which there appeared to be greater efficiency around the estimation of the estimated coefficient reduced the overall design efficiency.

Amendments to design: post pilot

Amendments to the survey design (other than to the cost and AE information – see Chapter 5) were considered following the feedback from the pilot survey and the CHERE internal seminar, namely:

- Whether to standardise the presentation of overlapping levels within attributes (as per the Efficacy frame version).
- Including an attribute for the duration of survival remaining.

The decision was made to only test one presentation of overlapping levels within the survey; that associated with *Efficacy* in the 'No-Efficacy Difference' frame. For all other attributes, overlapping levels would appear as two separate items for each choice option; as two separate levels for all attributes in the Base and 'Attribute' frames, and for all attributes other than efficacy in the 'No-Efficacy Difference' frame. Theoretically, it would be possible to test for an impact on choice outcomes due to differences in the presentation of overlapping levels. This could be achieved by comparing the choice coefficients resulting from the 'No-Efficacy Difference' frame with those from the 'Base' frame when the *Efficacy* levels overlap. However, this would require the same design coverage across the frames to achieve both sets of coefficients. This would not be the case for the 'Base' frame since it would be drawn from a subset of the design since only a subset of the *Efficacy* attributes had overlapping levels. The resulting coefficients would thus not be comparable with those from the 'No-Efficacy Difference' frame.

During the CHERE seminar it was discussed that inclusion of a survival attribute would allow the coefficient estimates to be expressed as utility values anchored against utility, and hence able to be used for the estimation of quality adjusted life years (QALYs). While such an outcome is of interest and would potentially expand on the role of meta-health effects as inputs to decisions of resource allocation, it would fundamentally change the design of the DCE. Moreover, it is unclear how respondents might choose between options, and trade between attributes, in the presence of both OOP costs and survival, and indeed whether separability of those 'payment vehicles' can be assumed for the purposes of estimating either mWTP or utility. Accordingly, this was not pursued further and the design was retained as described in Chapter 5.

Design efficiency estimates: the Final Survey Design

The priors used for the final survey design were taken from the coefficient estimates produced using the pilot survey results. The initial conditional logit regressions of the pilot survey were tested using dummy coding of the categorical variables (see Table A 19), as well as effects coding. The primary results were those using the effects coded categorical variables (for which the estimates are smaller than those when using dummy coding).

However, in entering the prior values into Ngene for use in the design of the final survey, the coefficient estimates from the dummy variable version of the analysis were used instead of those from the effects coded version. This means that for the categorical variables of mode and frequency, more relative importance was placed on the base attribute level (*Tablet* and *Monthly* respectively) compared with the other levels than would have occurred if the effects coded values had been used. For example, from the pooled analysis of the pilot survey data using effects code, the estimated choice coefficient for *IV: Tablet* was -0.55, compared with -0.92 from the dummy coded version. The same pattern was observed for the frequency of administration attribute.

Using the coefficient estimates from the dummy variable coded version (as described above), the estimated WTP, s-estimates and d-errors produced within Ngene were 2,527,297.19, 86.46 and 0.0002 respectively. Note, that given the change in the design between the pilot and the final survey design, the prior used for *Govt.Costs* was assumed to be the same as that for *Own OOP*. The design also allowed for two interactions, one between mode and *Own OOP*, and *Own OOP* and *Govt.Costs*, all modelled as continuous variables. If the design for the final survey had been generated using the coefficients from the effects coded conditional logit estimation, the corresponding efficiency estimates would have been 1,578,828.00 for WTP, 227.62 for the s-estimate (driven by the small coefficient on weekly administration compared with the other frequencies of administration), and 0.0002 for the d-error. Re-evaluating the design produced by Ngene with the dummy coded priors against the design code using the effects coded priors produces efficiency estimates of 2,373,597.09 for the WTP, 321.11 for the s-estimate, and 0.0003 for the d-error. This indicates that the use of the coefficients from the analysis using dummy coded categorical variables as priors in the design specifying effects coded treatment of categorical variables results in some loss of efficiency (indicated by the larger s-estimate and d-error).

Returning to the main design (using the prior values from the conditional logit specifying dummy coding of categorical values), the effects of re-balancing the design for potential dominance within scenarios was also examined. One scenario (Scenario 33) exhibited potential extended dominance in the 'No-Efficacy Difference' frame. To overcome this, the level on *Own OOP* for one option was changed from $500 to $0 (both were previously at $500; see Table A 11). The resulting design was imported into Ngene for re-evaluation

of the efficiency properties. The estimates were 2,543,649.34 for the WTP, 87.10 for the s-estimate, and 0.0002 for the d-error, indicating minimal change from the initial design.

Table A 11: Adjustments to Address Dominance in Final Survey Design

Block	Scenario	Alternative	Own OOP		
			From	To	Compared with:
1	33	A	500	0	500

Abbreviation: OOP, out-of-pocket.

Finally, the efficiency of the design was re-evaluated using the version in which there was no *Efficacy* difference between the options (the levels for the response attribute were the same between the two options). This produced efficiency estimates of 4,093,677.77 for WTP, 136.48 for the s-estimate, and 0.0003 for the d-error, indicating some loss of efficiency from the initial design.

Appendix 8: Chapter 5 - Results of the Pilot Survey

The key outcomes from the pilot survey are presented in this appendix. The survey went live on 27 October 2014 and was de-activated on 13 November 2014. During that time a total of 46 respondents logged on to the survey. However, only 33 completed all 12 choice sets.

Respondent Profile

Completed choice sets (n=12) were provided by 33 respondents, but only 32 completed the demographic information. A snapshot of their demographic information is provided in Table A 12.

Table A 12: Selected demographics of the survey respondents.

Variable			Respondents	
Name	**Category Level**	**Description**	**Number**	**%**
Age	1	16 to 34	2	6.25
	2	35 to 54	21	65.62
	3	55 & Over	9	28.12
Gender	1	Male	6	18.75
	2	Female	26	81.25
Income: Household income	1	Low ($0-$39,999)	1	3.12
	2	Medium ($40,000-$79,999)	4	12.5
	3	High ($80,000-$149,999)	4	12.5
	4	Very High (>$150,000)	12	37.5
	5	Did not know	8	25
	6	Declined	3	9.38
Education: Education attained	1	Primary only		
	2	Year 10 only	1	3.12
	3	Year 12 Only	2	6.25
	4	Vocational and Other	5	15.62
	5	University	24	75
Health: Self-reported health	1	Excellent	9	28.12
	2	Very good	13	40.62
	3	Good	7	21.88
	4	Fair	2	6.25
	5	Poor	1	3.12
Health: Chronic-health issues	0	No chronic health issues	20	62.50
	1	At least one chronic health issue	12	37.50
Medication Use	1	Yes	5	15.63
	2	No	20	62.50
Number of Medications	1	1-2	3	9.38
	2	2-4	0	0
	3	> 4	2	6.25

Of the 33 who completed the DCE survey, nine completed the frame testing the impact of changing the manner in which information on *Efficacy* is presented (No-Efficacy Difference), 14 completed the frame in which more information was provided about the mode of administration (Attribute), and 10 completed the standard frame (Base).

Survey Validation and Qualitative Feedback

Survey clarity and difficulty ratings

Clarity of the survey instructions, vignette, task and attribute descriptions were assessed using a rating on a scale from strongly disagree (1) to strongly agree (5). These ratings are provided below for 32 of the respondents who completed that aspect of the surveys (non-completers did not provide the rating information). These ratings show a high level of agreement regarding the clarity of the survey questions, attributes and that the task was not difficult to complete. Despite these strong average ratings, there were seven categories which scored an agreement rating of 2, but were only rated as such by one respondent at a time (ratings of 3 within the same category were given by as many as two respondents). In all, eight respondents gave at least one rating lower than four. Overall, these ratings would suggest minimal change is required to the wording or presentation of the survey.

Table A 13: Summary of agreement with survey validation questions.

	n = 32* Mean (s.d.); range
The instructions for the survey were clear.	4.44 (0.61); 3 - 5
The description of rheumatoid arthritis and its treatment was clear.	4.25 (0.61); 2 - 5
The description of rheumatoid arthritis was relevant to the task of answering the questions.	4.25 (0.71); 2 - 5
The language used in the questions was clear.	4.5 (0.5); 4 - 5
The questions were not difficult to understand.	4.44 (0.56); 3 - 5
The task was not difficult to complete.	4.44 (0.5); 4 - 5
The descriptions of the treatment were clear.	4.34 (0.69); 2 - 5
The descriptions about how often you took treatment were clear.	4.5 (0.5); 4 - 5
The text and pictures describing the chance of treatment affecting your condition were clear.	4.41 (0.49); 4 - 5
The text and pictures describing the chance of stopping treatment due to side effects were clear.	4.25 (0.56); 3 - 5
The description of the costs of treatment were clear.	4.32 (0.69); 2 - 5

Note: *One respondent did not answer the rating question on the description of the clarity of the costs of treatment.
Abbreviation: s.d., standard deviation.

Open-ended Feedback

Qualitative (open-ended) feedback was collected as part of the survey, either as a direct response to questions within the survey or in subsequent correspondence with respondents. The feedback is categorised in themes, with corresponding proposed actions, in Table A 14. Substantive actions to consider based on this feedback include:

- modifying the language used in the levels on the cost attribute (to note that it is per month);
- clarify in the introductory vignette that patients who stop treatment due to AEs are eligible to go on to another treatment;
- consider whether the manner in which attributes with level overlap are presented needs to be standardised; and
- implement changes to speed up the loading of the survey.

Table A 14: Survey feedback and possible actions.

Theme	Comments	Action
Survey access and loading	A number of respondents commented via email (one in the survey) that the vignettes were slow to load on their system.	Qualtrics have suggested reducing the size of the vignette pictures inserted into the online platform. Pictures will be resized.
	Vignettes (choice tasks) appeared sequentially on one large page rather than on separate pages on Internet Explorer.	This has been reported to Qualtrics and they are testing why it occurred. Follow-up with Qualtrics.
	"The multiple pair-wise considerations can become harder to follow further along. Splitting screens may help to break this up."	
Survey appearance	Use of blue for the efficacy attribute was very bright relative to the yellow for the safety attribute.	This might have been a computer specific issue. No other respondents raised this as an issue. No action at this stage.
	The progress bar did not move, or reflect the vignette number.	The progress bar reflects the total number of questions, not that to which respondents are randomised. It will be removed for the main survey.
Opt-out choices	Given that the survey is about a non-life threatening condition perhaps it should include an opt-out option as this might make it more realistic for individuals to answer.	Leave as a forced-choice since the interest is in relative effect sizes and not estimating absolute uptake rates. Moreover, the focus is on a methodological question rather than an applied question associated with RA treatment.
	How does the forced choice relate to the broader context with cost where government pays? The forced choice relates to the individual, but the cost brings in the concept of government (society).	No action.
Treatment descriptions	"My problem was that the treatment was described as controlling my condition (probability) yet at the same time the arthritis could be so bad I couldn't handle the tablets."	Check the internal consistency of the combinations.
	"More detail could have been provided regarding treatment could have been more expansive without leading compromising participants' responses."	One of the questions for the survey is to test the effect of varying how treatment is described (interestingly the first respondent here saw the 'Attribute' framed version where more information was provided on the mode of administration.)
	"I wonder if you may need to explain what an intravenous infusion is for some people. Maybe a hover or definition in the beginning (i.e. go to hospital, needle in the arm, drip)."	No action.

Theme	Comments	Action
	"The description is simplified so amyone [sic] who knows about RA might not be convinced."	Test if values are different among those with a chronic condition (unlikely to have sufficient numbers with RA). Comments regarding treatment description as above, and safety as below.
	"Description of RA was fine; but I think there should be some discussion of the treatment options; and the attributes before the choice tasks - I wanted to know what the harms actually were."	
Side effects	"A description of what the possible side effects might be may have changed my opinion on some answers."	The inclusion of more information on side effects would change the nature of the experiment to shift the focus to side effects.
		No action.
	"Might have helped if I knew what the side effects were."	
Introductory text	"Para 1, 2nd sentence maybe say "importance of different aspects" rather than each attribute as these are repeated in next para. Para 2: maybe mention hypothetical or imaginary? Para 3: interested in "your opinion" rather than "answers"?"	Review introductory text, although noting positive feedback received to date.
	"Might be good to start with "Imagine that you have..."	
Stopping treatment	"Interviewee knew whether they had the option or not to trial one treatment and then if that didn't work, you were able to choose the second option. It was presumed that this is the case, as it is a real life scenario, but if the treatment meant you had 'one shot' the responses would be different."	Including more information here about treatment sequencing might then make the exercise interpretable as a BWS, and would perhaps make it more relevant to include an opt-out. This could perhaps be included as part of the introductory vignette.
	"What does stopping treatment mean? Can another treatment be started?"	
	"yes, but I worry about whether this is conflating 2 separate issues - adherence and harms"	Whether or not a patient adheres to treatment might be quite independent from the occurrence of AEs. In this case, an adherent patient might need to stop treatment because of an AE (presumably non-adherent patients would have a reduced incidence of AEs due to reduced exposure). The wording is ok since no attribution of causality for stopping treatment other than an AE occurring is being made. No action.

Theme	Comments	Action
Costs	"The problem is that a payment of $1000 per month would be considered by some people as unaffordable. I wonder if this will affect the realism."	The levels reflect the monthly treatment costs for the bDMARDs. This was intentional to include the full range of possible cost if the government subsidy on the drug was removed. It is acknowledged that there are other costs to RA patients not captured here.
	"Out of pocket cost is high in many scenarios."	No action.
	"Each treatment involves additional costs, which were not specified."	
	"It is worth making it clear again in the level that the amount you pay is per month."	Potentially make this change.
Understanding/clarity	"The use of multiple methods to demonstrate risk is helpful. Understanding will be limited in some individuals. The repetition increases complexity."	No action.
	"Yes, but though 4 in 10 etc would be easier but understand the need to be consistent with the following attribute with rare numbers."	No action.
Risk (efficacy)	"I liked how the percentage was depicted as a graph"	No action.
	"Why include it if they are always the same across A and B?"	This referred to presenting the levels for the efficacy attribute in the No-Efficacy frame (which were the same). This was deliberate to test that method of presenting information.
		No action.
Framing	Is it internally consistent for the survey to present attribute levels that are the same across scenarios in different ways?	At the moment where there is level overlap, the levels appear in both scenarios for all attributes and versions except the efficacy attribute in the No-Efficacy frame.
		Consider whether this needs to be modified.

Abbreviation: AE, adverse event; bDMARD, biological disease modifying anti-rheumatic drugs; RA, rheumatoid arthritis.

Strategies for Answering the DCE

Respondent ratings of attributes in terms of what was most important, and (separately) what was least important, are provided in Table A 15.

Table A 15: Respondent rated attribute importance.

Attribute	Most important n (%)	Least important n (%)
Treatment type	2 (6.25)	6 (18.75)
Treatment frequency	1 (3.12)	12 (37.5)

Efficacy	18 (56.25)	1 (3.12)
Side effects	2 (6.25)	3 (9.38)
Cost	9 (28.12)	10 (31.25)
Total	**32**	**32**

In terms of how they answered the choice question, fifteen of the respondents answered that they focused only on those factors that were important (12) or had some other strategy (see Table A 16). Review of the responses provided showed that most focused on *Efficacy* (consistent with the majority of respondents considering this to be most important factor), then *Cost* in deciding between the alternatives. One respondent commented that they felt their responses might have changed if there was more information about the duration of the condition or how long they would be paying the cost of the medication. Another stated that they considered everything – except frequency.

Table A 16: Strategies employed when answering choice scenarios.

Strategy	**Frequency n (%)**
I did not have a strategy	0
I focused only on the factors I thought were important	12 (37.5)
I considered most of the factors all the time	9 (28.12)
I considered all the factors each time	8 (25)
Other	3 (9.38)
Total	**32**

Analysis of Results

Given the convenience nature of the sample involved, this sample and the results obtained are not considered to be representative. However, the data were analysed to:

- explore the underlying choice relationships with a view to testing the proposed approach to the analysis for the main study data, particularly with respect to the testing of the framing effects;
- investigate WTP for the various attributes;
- investigate potential hypotheses associated with preference heterogeneity; and
- to validate the prior information used in the DCE design process.

The analysis utilised conditional logit regression to explore the existing data.[235,236] The potential for interaction effects (between attributes and with demographic covariates) to be modelled will also be explored. Preference heterogeneity will be explored using latent class analysis and sub-group analyses (drawing on the underlying conditional logit regression analysis).[188,238]

Attribute balance and choice

Attribute level design frequency is presented in Table A 17. The resulting frequency of the attribute levels for *Frequency* in the preferred scenario in each choice set did not differ

markedly from those in the design, but there were slight variations noted in the frequency of levels for the remaining attributes. For example, there were higher proportions of scenarios chosen with higher levels of the *Efficacy* attribute, relative to lower levels, than were included in the initial design. Similarly, scenarios with the *Injection* or *Tablet* occurred in a higher proportion of preferred choice sets compared with *IV* relative to that in the design. Both of these results make intuitive sense; respondents prefer greater efficacy and less intrusive modes of treatment administration. While the resulting patterns of distribution for the levels for *Safety* and *OOP* differed from those in the base design there was no pattern discernible.

Table A 17: Comparison of attribute level balance and subsequent choice frequencies

	Level Distribution in Design		Frequency in Chosen Scenario
	Scenario A	Scenario B	
Mode			
IV	0.33	0.33	0.25
Injection	0.33	0.33	0.37
Tablet	0.33	0.33	0.38
Frequency			
Bi-daily	0.13	0.13	0.14
Weekly	0.31	0.31	0.32
Fortnightly	0.31	0.31	0.34
Monthly	0.25	0.25	0.20
Efficacy			
20%	0.25	0.25	0.11
40%	0.25	0.25	0.21
60%	0.25	0.25	0.31
70%	0.25	0.25	0.37
Safety			
1%	0.33	0.33	0.34
5%	0.33	0.33	0.27
10%	0.33	0.33	0.39
Out-of-pocket costs			
0	0.23	0.25	0.22
$1,000	0.27	0.27	0.32
$1,500	0.25	0.25	0.20
$2,000	0.25	0.23	0.26

Abbreviation: IV, intravenous.

Base results and framing effects

Results of the conditional logit analysis for the overall sample are presented in Table A 18. The results from the model "All 1" are for all three versions of the survey (Base, No-Efficacy Difference and Attribute) combined. From these results it can be observed that the likelihood of choice significantly increases with treatment effect, and significantly decreases with increasing out-of-pocket costs borne by the individual; both as expected. In terms of the meta-health effects, individuals were significantly less likely to choose an

option if it included an IV infusion compared with a tablet (as expected). While the direction and relative magnitude of the other level on mode, and all the levels on frequency were as expected they were not significant. Similarly, the direction of the sign on safety was as expected but it was not significant.

Comparison of these coefficients with those included as priors in the initial design in Ngene provides some assurance as to the face validity of the model; the direction and relative difference between the coefficients produced by the "All 1" specification using the pilot data and those from the priors used in the design is consistent. Performance of the base ("All 1") model was improved slightly taking into account the restrictions in the design imposed on mode and frequency. The results for "All 2" include interaction terms for mode and frequency and indicate that the significance of IV in "All 1" is mediated by its interaction with fortnightly frequency relative to any frequency of tablet. Inclusion of these interactions renders safety significant.

Table A 18: Conditional logit specifications – Pilot Survey.

	All 1	**All 2**	**No-Efficacy Difference**	**Base Only**	**Attribute**	**Priors**
Mode						
IV: *Tablet*	-0.550	-0.407	-0.719	-0.685	-0.221	-0.41
	(0.205)**	(0.250)	(0.341)*	(0.521)	(0.248)	
Injection: *Tablet*	0.179	0.312	0.148	0.245	0.246	-0.09
	(0.144)	(0.205)	(0.192)	(0.282)	(0.289)	
Frequency						
BD: *Monthly*	-0.470	-0.122	-0.727	-0.597	0.124	-0.61
	(0.328)	(0.436)	(0.599)	(0.589)	(0.321)	
Weekly: *Monthly*	0.039	-0.007	0.114	-0.262	0.193	-0.12
	(0.152)	(0.164)	(0.275)	(0.360)	(0.251)	
Fortnightly: *Monthly*	0.251	0.316	0.287	0.870	-0.038	-0.45
	(0.143)	(0.171)	(0.188)	(0.342)*	(0.201)	
Efficacy	0.078	0.086	-	0.109	0.063	0.33
	(0.011)**	(0.012)**		(0.025)**	(0.012)**	
Safety	-0.066	-0.075	-0.120	-0.070	0.021	-0.3
	(0.038)	(0.037)*	(0.060)*	(0.073)	(0.056)	
OOP	-0.001	-0.001	-0.001	-0.001	-0.000	-0.003
	(0.000)*	(0.000)**	(0.001)	(0.001)	(0.001)	
Mode by Frequency						
IV#Weekly		-0.346				
		(0.181)				
IV#Fortnightly		-0.383				
		(0.345)				
IV#Monthly		0.000				
Injection#Weekly		0.113				
		(0.466)				
Injection#Fortnightly		0.000				
		0.000				
Injection#Monthly		0.000				
Tablet#BD		0.000				
Tablet#Weekly		0.000				
Tablet#Fortnightly		0.000				
Observations	792	792	216	240	336	

	All 1	All 2	No-Efficacy Difference	Base Only	Attribute	Priors
N	33	33	9	10	14	
Wald Chi	83.17	94.03	114.37	4,692.20	135.44	
pseudo R-squared	0.34	0.36	0.09	0.59	0.38	
d.f.	8	11	7	8	8	
p-value	0.00	0.00	0.00	0.00	0.00	
Log-likelihood	-180.26	-176.75	-68.00	-34.20	-71.88	

Notes: * $p<0.05$; ** $p<0.01$
Italicised labels indicate the base level for that attribute (or interaction).
Numbers in parentheses below coefficients are standard errors.
Pseudo R-squared is produced by the clogit command in STATA.
Abbreviations: BD, twice daily; d.f., degrees of freedom; IV, intravenous; OOP out-of-pocket costs.

The use of dummy coding, as opposed to effects coding, for all categorical variables was tested for the set of analyses corresponding to those for which results are presented in Table A 18 and resulted in qualitatively the same results (see Table A 19).

Table A 19: Alternative results – Using Dummy Coded Categorical Meta-Health Attributes

	All 1	All 2	No-Efficacy Difference	Base Only	Attribute	Priors
Mode						
IV: *Tablet*	-0.920	-0.044	-1.291	-1.125	-0.195	-0.41
	(0.391)*	(0.494)	(0.686)	(0.873)	(0.380)	
Injection: *Tablet*	-0.192	-0.167	-0.424	-0.196	0.272	-0.09
	(0.299)	(0.334)	(0.483)	(0.431)	(0.458)	
Frequency						
BD: *Monthly*	-0.649	-0.205	-1.052	-0.587	0.403	-0.61
	(0.606)	(0.596)	(1.158)	(0.921)	(0.443)	
Weekly: *Monthly*	-0.141	0.142	-0.211	-0.252	0.472	-0.12
	(0.409)	(0.470)	(0.808)	(0.678)	(0.445)	
Fortnightly: *Monthly*	0.072	0.615	-0.039	0.880	0.241	-0.45
	(0.335)	(0.392)	(0.558)	(0.768)	(0.329)	
Efficacy	0.078	0.086	0.000	0.109	0.063	0.33
	(0.011)**	(0.012)**		(0.025)**	(0.012)**	
Safety	-0.066	-0.076	-0.120	-0.070	0.021	-0.3
	(0.038)	(0.037)*	(0.060)*	(0.073)	(0.056)	
OOP	-0.001	-0.001	-0.001	-0.001	-0.0002	-0.003
	(0.000)*	(0.000)**	(0.001)	(0.001)	(0.001)	
Mode by Frequency						
IV#Weekly		-1.037				
		(0.542)				
IV#Fortnightly		-1.225				
		(0.565)*				
IV#Monthly		0.000				
Injection#Weekly		0.263				
		(0.738)				
Injection#Fortnightly		0.000				
		0.000				
Injection#Monthly		0.000				
Tablet#BD		0.000				
Tablet#Weekly		0.000				

	All 1	All 2	No-Efficacy Difference	Base Only	Attribute	Priors
Tablet#Fortnightly		0.000				
N (observations)	792	792	216	240	336	
N (respondents)	33	33	9	10	14	
pseudo R-squared	0.34	0.36	0.09	0.59	0.38	
Wald Chi	83.17	94.03	114.37	4,692.21	135.44	
DF	8	11	7	8	8	
p-value	0.00	0.00	0.00	0.00	0.00	
Log-likelihood	-180.26	-176.75	-68.00	-34.20	-71.88	

Notes: * $p<0.05$; ** $p<0.01$
Italicised labels indicate the base level for that attribute (or interaction).
Number in parentheses below coefficients are standard errors.
Abbreviations: BD, twice daily; d.f., degrees of freedom; IV, intravenous; OOP, out-of-pocket costs.

The comparison across the models with respect to framing effects can be formed using the models labelled 'No-Efficacy Difference', "Base only" and 'Attribute'. In the "Base only" and 'Attribute' models, *Efficacy* had the only statistically significant coefficient (*OOP* reached a p-value=0.051 in the "Base only" model). Of the meta-health effects, *IV*: *Tablet* was significant in "No-Efficacy", while *Fortnightly: Monthly* was significant in "Base only". The former supports the presence of an efficacy framing effect in the role of mode of administration in choice. The results for the 'Attribute' framing set suggest that the addition of information reduces the disparity between the mode of administration (things might not be as bad as people imagine). Both results suggest that the hypotheses that efficacy framing affects the valuation of meta-health effects, and for the provision of more information to influence valuation, can be tested using the proposed approach.

Understanding differences: demographics, latent classes and prior experiences

The potential influence of demographic factors on choice was explored by interacting the relevant covariates with a dummy variable indicating whether or not the choice made appeared on the left or the right (the base). This is similar to interacting the covariates with an alternative specific constant (except in this case it is an unlabelled experiment). Five models were tested, including various combinations of *Age*, *Gender*, *Education* and *Income*. All five models were estimated using the data from all three frames within the same specification. One respondent had not provided any demographic information, reducing the sample available by one (n=32).

In general, inclusion of the demographic variables did not alter the magnitude or significance of the coefficient estimates ("All 1") previously reported in Table A 18 (see Table A 20). Of the demographic variables tested, *Gender* was not significant, and the significant effect found for *Age* in Model 1 dissipated in Model 5 that included both *Income* and *Education*. These results suggest an overall effect for *Income* and *Education* in choice, potentially describing two groups; those with income below $80,000 per year and non-university education, and the complementary group.

Table A 20: Conditional logit specifications incorporating demographic factors.

	Model 1	Model 2	Model 3	Model 4	Model 5
Mode					
IV: *Tablet*	-0.635	-0.634	-0.630	-0.639	-0.633
	(0.204)**	(0.203)**	(0.205)**	(0.204)**	(0.206)**
Injection: *Tablet*	0.135	0.132	0.151	0.127	0.152
	(0.149)	(0.149)	(0.152)	(0.150)	(0.152)
Frequency					
BD: *Monthly*	-0.623	-0.620	-0.585	-0.640	-0.594
	(0.312)*	(0.316)*	(0.313)	(0.317)*	(0.316)
Weekly: *Monthly*	0.004	0.002	0.040	0.004	0.039
	(0.165)	(0.166)	(0.176)	(0.167)	(0.178)
Fortnightly: *Monthly*	0.254	0.252	0.251	0.255	0.254
	(0.156)	(0.156)	(0.162)	(0.156)	(0.167)
Efficacy	0.081	0.081	0.082	0.082	0.082
	(0.011)**	(0.011)**	(0.011)**	(0.011)**	(0.011)**
Safety	-0.073	-0.074	-0.073	-0.075	-0.072
	(0.042)	(0.042)	(0.043)	(0.042)	(0.043)
OOP	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.000)**	(0.000)**	(0.000)*	(0.000)**	(0.000)*
0b.Age position interaction	0.000				0.000
16-24	-0.298				-0.516
	(0.084)**				(0.421)
25-44	-0.197				-0.227
	(0.141)				(0.419)
45-64	-0.110				-0.079
	(0.215)				(0.402)
0b.Gender position interaction		0.000			
Male		-0.181			
		(0.156)			
Female		-0.178			
		(0.135)			
0b.Income position interaction			0.000		0.000
Under $40K			0.561		0.783
			(0.122)**		(0.410)
$40-79K			-0.881		-0.611
			(0.455)		(0.680)
$80-$149K			-0.338		-0.155
			(0.279)		(0.481)
> $150K			-0.079		0.169
			(0.127)		(0.379)
Inc Declined			-0.017		0.284
			(0.233)		(0.441)
Inc Don't Know			-0.164		0.000
			(0.401)		
0b.Educ. position interaction				*0.000*	*0.000*
Year 10				-0.832	-1.041
				(0.123)**	(0.271)**
Year 12				-0.120	-0.006
				(0.084)	(0.260)
Tafe				-0.475	-0.174
				(0.340)	(0.374)
University				-0.121	0.000
				(0.134)	

	Model 1	Model 2	Model 3	Model 4	Model 5
Observations	768	768	768	768	768
N	32	32	32	32	32
pseudo R-squared	0.36	0.36	0.37	0.36	0.37
Wald Chi	172.15	86.74	.	.	.
d.f.	11	10	13	11	17
p-value	0.00	0.00	.	.	.
Log-likelihood	-170.18	-170.25	-167.92	-169.62	-167.21

Notes: * $p<0.05$; ** $p<0.01$
Italicised labels indicate the base level for that attribute (or interaction).
Numbers in parentheses below coefficients are standard errors.
Pseudo R-squared is produced by the clogit command in STATA.
Abbreviations: BD, twice daily; d.f., degrees of freedom; IV, intravenous; OOP out-of-pocket costs.

Following the exploratory analysis including demographic factors, a latent class analysis was conducted on the primary covariates only (excluding demographics) as a means of exploring preference heterogeneity within this relatively small sample. The number of classes was chosen by running the analysis with class numbers between 2-6 and choosing that with the best (lowest) AIC/BIC which could be estimated without returning program warnings errors.[188] This occurred with four classes (five and six classes produced lower AIC/BIC values but with error warnings in the construction of the variance matrix). The results are presented in Table A 21. The principal distinction between the classes is the pattern of significance of the meta-health effects, *Safety* and *OOP* in affecting choice; class 1 were influenced by mode (*IV: Tablet*) and *Safety*; class 2 by *Injection: Tablet*; class 3 by *IV: Tablet*, *Injection: Tablet*, *BD: Monthly, Safety* and *OOP*, and class 4 by *Injection: Tablet* and *OOP*. All four classes were influenced by *Efficacy*.

Table A 21: Latent class analysis: four classes.

	Class 1	Class 2	Class 3	Class 4
Share (%)	0.263	0.304	0.169	0.263
Mode				
IV: *Tablet*	-1.459	-0.167	-2.559	-0.810
	(0.605)*	(0.780)	(0.976)**	(0.605)
Injection: *Tablet*	0.355	1.106	1.024	-1.235
	(0.417)	(0.422)**	(0.441)*	(0.443)**
Frequency				
BD: *Monthly*	-0.370	1.146	-3.132	-0.579
	(0.908)	(1.233)	(1.263)*	(0.836)
Weekly: *Monthly*	-0.785	0.219	-0.025	-0.263
	(0.603)	(0.570)	(0.485)	(0.501)
Fortnightly: *Monthly*	0.325	0.204	0.840	-0.569
	(0.389)	(0.403)	(0.588)	(0.417)
Efficacy	0.149	0.191	0.099	0.121
	(0.042)**	(0.049)**	(0.032)**	(0.034)**
Safety	-0.348	0.009	-0.409	-0.254
	(0.150)*	(0.132)	(0.191)*	(0.186)
OOP	0.000	-0.001	-0.004	-0.006
	(0.001)	(0.001)	(0.002)*	(0.002)**
Share Constants	0.002	0.147	-0.439	
	(0.546)	(0.678)	(0.637)	

	Class 1	Class 2	Class 3	Class 4
Observations	768			
N	32			
Log-likelihood	-144.57			

Notes: * $p<0.05$; ** $p<0.01$
Italicised labels indicate the base level for that attribute.
Numbers in parentheses below coefficients are standard errors.
Abbreviations: BD, twice daily; IV, intravenous; OOP, out-of-pocket costs.

Estimating mWTP

mWTP was estimated based on the coefficient estimates from "All 1", the specification using the ratio of the coefficient on the attribute of interest to that of the cost attribute. Confidence intervals and p-values have been estimated by STATA using the product moment method.[248] Results for the mWTP estimates are presented in Table A 22. Only three estimates were significant; the shift to tablets from IV, additional disease control, and the additional chance of stopping treatment. None of the other meta-health effect estimates was significant, although this might be a function of the sample size. In terms of the significant estimates, the highest absolute value was for mode (*IV*) at $533. mWTP for the two health effects were substantially lower at 14.3% and 12.1% of the relative value of mode for *Efficacy* and *Safety* respectively. These relative effect sizes [251] were used to produce the rankings reported in Table A 22.

Table A 22: Summary of willingness-to-pay analyses.

	Mean (95% CI)	Relative Effect Size
Mode		
IV: *Tablet*	-$533 (-$896 to -$170) (p=0.004)	1
Injection: *Tablet*	$174 (-$128 to $475) (p=0.259)	0.326 (4)
Frequency		
BD: *Monthly*	-$455 (-$1,029 to $118) (p=0.12)	0.854 (2)
Weekly: *Monthly*	$38 (-$261 to $336) (p=0.806)	0.070 (7)
Fortnightly: *Monthly*	$244 (-$74 to $561) (p=0.132)	0.457 (3)
Efficacy: Additional disease control (per 1%)	$76 ($22 to $131) (p=0.006)	0.143 (5)
Safety: Chance of stopping treatment (per 1%)	-$64 (-$120 to -$9) (p=0.023)	0.121 (6)

Note: Relative Effect Size is estimated as the ratio of each mWTP to that of the largest individual mWTP (in this case the value for *IV*).
Abbreviation: BD, twice daily; CI, confidence interval; IV, intravenous; mWTP, marginal willingness to pay.

The impact of assuming *Efficacy* and *Safety* are linear in effect is explored by re-estimating the models using categorical coding for these two health effects. The results are presented in Table A 23 and Table A 24, and show that the use of categorical coding for health effects

elevates them to be the most influential attributes in explaining choice and relative value.

Table A 23: Impact of Effects Coding for Efficacy and Safety on Choice Estimates

	All 1	Effects 1
Mode		
IV: *Tablet*	-0.550	-0.590
	(0.205)**	(0.210)**
Injection: *Tablet*	0.179	0.183
	(0.144)	(0.148)
Frequency		
BD: *Monthly*	-0.470	-0.549
	(0.328)	(0.343)
Weekly: *Monthly*	0.039	0.027
	(0.152)	(0.165)
Fortnightly: *Monthly*	0.251	0.307
	(0.143)	(0.130)*
Efficacy		
Continuously coded	0.078	
	(0.011)**	
20%: *70%*		-3.844
		(0.587)**
40%: *70%*		-2.265
		(0.489)**
60%: *70%*		-0.496
		(0.421)
Safety		
Continuously coded	-0.066	
	(0.038)	
5%: *1%*		-0.464
		(0.377)
10%: *1%*		-0.705
		(0.371)
OOP	-0.001	-0.001
	(0.000)*	(0.000)*
Observations	792	792
N	33	33
pseudo R-squared	0.34	0.34
Wald Chi	83.17	84.88
d.f.	8	11
p-value	0.00	0.00
Log-likelihood	-180.26	-179.83

Notes: * $p<0.05$; ** $p<0.01$
Italicised labels indicate the base level for that attribute.
Number in parentheses below coefficients are standard errors.
Abbreviation: BD, twice daily; d.f., degrees of freedom; IV, intravenous; OOP, out-of-pocket costs.

Table A 24: Impact of Effects Coding on mWTP

	All 1		Effects 1	
	mWTP	**Weight (rank)**	**mWTP**	**Weight (rank)**
Mode				
IV: Tablet	-$533 (-$896 to -$170)	1	-$542 (-$879 to -$205)	0.153 (4)

	All 1		Effects 1	
	mWTP	**Weight (rank)**	**mWTP**	**Weight (rank)**
	(p=0.004)		(p=0.002)	
Injection: Tablet	$174 (-$128 to $475) (p=0.259)	0.326 (4)	$168 (-$139 to $475) (p=0.284)	0.045 (9)
Frequency				
BD: Monthly	-$455 (-$1,029 to $118) (p=0.12)	0.854 (2)	-$504 (-$1,031 to $23) (p=0.061)	0.143 (5)
Weekly: Monthly	$38 (-$261 to $336) (p=0.806)	0.070 (7)	$25 (-$277 to $327) (p=0.87)	0.007 (10)
Fortnightly: Monthly	$244 (-$74 to $561) (p=0.132)	0.457 (3)	$282 (-$7 to $572) (p=0.056)	0.080 (7)
Efficacy				
Additional disease control (per 1%)	$76 ($22 to $131) (p=0.006)	0.143 (5)		
20%: *70%*			-$3,532 (-$6,284 to -$780) (p=0.012)	1
40%: *70%*			-$2,081 (-$3,759 to -$403) (p=0.015)	0.589 (2)
60%: *70%*			-$456 (-$1,341 to $429) (p=0.313)	0.129 (6)
Safety				
Chance of stopping treatment (per 1%)	-$64 (-$120 to -$9) (p=0.023)	0.121 (6)		
5%: *1%*			-$427 (-$962 to $109) (p=0.118)	0.121 (7)
10%: *1%*			-$647 (-$1,122 to -$173) (p=0.008)	0.183 (3)

Note: All 1 shows the results for mWTP for the estimates of the conditional logit regression in which health effects are coded as continuous variables. Effects 1 shows the mWTP using estimates from the analysis in which health effects are coded as categorical variables.

Abbreviation: BD, twice daily; CI, confidence interval; IV, intravenous; mWTP, marginal willingness to pay.

Conclusions from Pilot Study

The results of the pilot survey appear to support the proposed study design in terms of investigating the value of meta-health effects, in this case convenience as evidenced through the mode and frequency of treatment administration. In addition, the results suggest that the DCE design is valid, and the proposed use of alternative versions of the survey (Base, Attribute, No-Efficacy Difference) to investigate the effects of framing on the resulting value derived for convenience is feasible.

Appendix 9: Chapter 5 - Supporting Evidence for the Main Survey

Table A 25: Mixed logit Regression – Log of Costs

	Pooled	Base	Attribute	No-Efficacy Difference
Coefficient Means				
IV: *Tablet*	-0.651	-0.577	-0.833	-0.554
	(0.090)**	(0.161)**	(0.159)**	(0.156)**
Injection: *Tablet*	0.065	0.146	0.207	-0.137
	(0.057)	(0.116)	(0.104)*	(0.096)
BD: *Monthly*	-0.108	0.039	-0.259	-0.079
	(0.081)	(0.135)	(0.156)	(0.140)
Weekly: *Monthly*	-0.068	0.080	-0.082	-0.177
	(0.046)	(0.084)	(0.083)	(0.083)*
Fortnighly: *Monthly*	-0.063	-0.216	0.002	0.034
	(0.049)	(0.098)*	(0.085)	(0.082)
Efficacy	0.068	0.074	0.062	
	(0.008)**	(0.011)**	(0.012)**	
Safety	-0.097	-0.123	-0.044	-0.137
	(0.013)**	(0.019)**	(0.019)*	(0.028)**
Log Own OOP	-0.541	-0.700	-0.480	-0.554
	(0.050)**	(0.111)**	(0.073)**	(0.078)**
Log Govt. Cost	-0.078	-0.052	-0.120	-0.060
	(0.021)**	(0.033)	(0.038)**	(0.037)
Standard Deviations				
IV: *Tablet*	0.940	1.033	0.914	0.951
	(0.101)**	(0.163)**	(0.168)**	(0.163)**
Injection: *Tablet*	0.382	-0.509	0.472	0.362
	(0.117)**	(0.163)**	(0.146)**	(0.129)**
BD: *Monthly*	-0.114	0.131	-0.328	0.067
	(0.100)	(0.120)	(0.393)	(0.095)
Weekly: *Monthly*	0.172	-0.205	-0.353	-0.242
	(0.222)	(0.114)	(0.150)*	(0.149)
Fortnighly: *Monthly*	-0.080	-0.122	-0.004	0.283
	(0.262)	(0.227)	(0.063)	(0.176)
Efficacy	0.064	0.061	0.064	
	(0.008)**	(0.010)**	(0.011)**	
Safety	0.132	-0.075	0.117	0.193
	(0.015)**	(0.037)*	(0.033)**	(0.028)**
Log Own OOP	0.461	0.629	0.437	0.380
	(0.049)**	(0.091)**	(0.075)**	(0.075)**
Log Govt. Cost	0.224	-0.196	0.271	0.285
	(0.033)**	(0.040)**	(0.066)**	(0.064)**
Observations	10,128	3,384	3,360	3,384
N	422	141	140	141
Wald Chi	184.75	110.36	72.47	63.32
d.f.	9	9	9	8
p-value	0.00	0.00	0.00	0.00
Log-likelihood	-2,575.85	-780.83	-863.80	-891.90

Notes: * p<0.05, ** p<0.01, *** p<0.0001

All categorical data are effects coded to minimise the potential for contamination of the base effect in the grand mean.

Base categories are shown in italics.
All analyses were conducted in STATA 12, with robust standard errors (in parentheses) to account for the survey nature of the data.
Analyses conducted with 1,000 replications.

Abbreviations: BD, twice daily; d.f., degrees of freedom; s.d. standard deviation; Ft, fortnight; Inj, injection; IV, intravenous; OOP out-of-pocket; Wk, weekly.

Table A 26: GMNL Specification – Comparing 500 and 1,000 Replications for 'Attribute' frame, with Effects Coded Health Effects

	1,000 Replications	500 Replications
Coefficient Means		
IV: *Tablet*	-1.268	-0.931
	(0.772)	(0.438)*
Injection: *Tablet*	-0.162	0.021
	(0.319)	(0.212)
BD: *Monthly*	-1.449	-0.864
	(0.676)*	(0.341)*
Weekly: *Monthly*	-0.108	-0.004
	(0.238)	(0.199)
Fortnightly: *Monthly*	-0.145	-0.022
	(0.231)	(0.19)
Efficacy 20%: *70%*	-5.453	-4.133
	(3.468)	(2.698)
Efficacy 40%: *70%*	-1.773	-1.175
	(1.502)	(1.106)
Efficacy 60%: *70%*	2.013	1.549
	(1.545)	(1.154)
Safety 5%: *1%*	0.172	0.164
	(0.310)	(0.206)
Safety 10%: *1%*	-1.306	-0.951
	(0.864)	(0.688)
Own OOP	-0.009	-0.006
	(0.003)**	(0.002)**
Govt. Cost	-0.001	-0.0005
	(0.001)	(-0.0003)
Standard Deviations		
IV: *Tablet*	0.902	0.939
	(0.208)**	(0.255)***
Injection: *Tablet*	0.523	0.504
	(0.128)**	(0.198)*
BD: *Monthly*	0.499	0.052
	(0.261)	(0.221)
Weekly: *Monthly*	-0.376	-0.446
	(0.140)**	(0.186)*
Fortnightly: *Monthly*	0.083	-0.041
	(0.092)	(0.137)
Efficacy 20%: *70%*	0.680	-0.506
	(0.264)**	(0.418)
Efficacy 40%: *70%*	0.535	0.788
	(0.189)**	(0.302)**
Efficacy 60%: *70%*	0.248	0.162

	1,000 Replications	500 Replications
	(0.281)	(0.33)
Safety 5%: *1%*	0.292	0.402
	(0.141)*	(0.128)**
Safety 10%: *1%*	-0.483	-0.588
	(0.192)*	(0.16)***
Own OOP	0.004	0.005
	(0.001)**	(0.001)***
Govt. Cost	0.001	0.001
	(0.000)**	(0.0002)***
Tau (constant)	-1.657	-1.661
	(0.388)**	(0.589)**
Observations	3,360	3360
N	140	140
Wald Chi	13.78	36.16
d.f.	12	12
p-value	0.31	0.0003
Log-likelihood	-848.71	-847.85

Notes: * p<0.05, ** p<0.01, *** p<0.0001
All categorical data are effects coded to minimise the potential for contamination of the base effect in the grand mean.
Base categories are shown in italics.
All analyses were conducted in STATA 12, with robust standard errors (in parentheses) to account for the survey nature of the data.

Abbreviations: BD, twice weekly; d.f., degrees of freedom; IV, intravenous; OOP out-of-pocket.

Appendix 10: Chapter 6 - Focus Group

Target recruitment for each group was 10 participants. To limit the extent of travel for participants, only women in the Sydney and Newcastle metropolitan areas were invited to participate in a two hour, focus group discussion. The groups were held separately to avoid the possibility of causing distress to participants arising from conflicting views amongst those who did or not undergo CPM.

The aim of the focus groups was to understand the factors that were considered by women as important when deciding how best to manage their ongoing breast cancer risk. The specific research question was: what are the factors that are important to women when making a decision about the ongoing management of breast cancer risk. This was explored via three sub-questions:

1. What factors were most important to women who underwent CPM?
2. What factors were most important to women who did not undergo CPM?
3. What would women tell someone making this decision that they had not been told?

The combined information from these three questions will provide an overview of the factors that the women in these groups consider to be relevant in making decisions about ongoing breast cancer risk. The last question was designed to elicit from women information beyond what might typically be listed as a characteristic of breast cancer risk management, but that was nonetheless germane to the decisions made by women. In addition a checklist of factors previously identified by Rosenberg et al. (2013) as considered by women deciding to undergo CPM was used as a checklist during the focus groups.[285]

Each focus group was facilitated by two investigators, and a breast cancer research nurse was also present in the event of clinical questions arising or incidence of participant unease. Discussions were audio taped for later transcription, along with physical minutes of the most salient points to aid in later transcription. Each woman provided prior consent to participate, and on arrival at the focus group provided de-identified information on her age, time since her diagnosis with early breast cancer, and time since completing treatment for breast cancer. The qualitative research was approved by the UTS Human Research Ethics Committee (No. 2014000423).

Analysis Plan

Data collected on the age, time since diagnosis (months), and time since primary treatment for breast cancer (months) were reported as aggregates (medians, min, max; mean). Transcribed information on women's decisions regarding CPM was analysed thematically. Using the discussion topics as a guide, transcripts of the sessions, along with notes taken by the investigators, were classified into core themes to describe the factors used by women in their decision-making. This not only considered whether a theme arises (e.g. fear of cancer recurrence) but also the

language used to describe that theme which might indicate its role in decision-making (e.g. it made women feel safer, less fearful).

The proposed list of emergent themes, how they were described and the extent of their importance, was reviewed by the fellow facilitator and other investigators on this project for consistency and agreement. Themes were grouped in terms of their apparent importance to women in the decision-making process.

Mastectomy Focus Group – Discussion Guides

Overview: to gain an understanding of the factors important to women when deciding how to best manage their ongoing risk of cancer recurrence.

Discussion

Sharing Experiences: Invite women to describe their experiences and how they managed their ongoing risk of cancer recurrence.

Why:

What factors were important to you when you made your decision about how to manage your ongoing risk of breast cancer?

- How do you describe those factors?
- What made those factors important to your decision?

Hindsight:

What are the important things you would say to another woman facing the same choices you faced?

- Were these things that you were told?
- Were these things that you wish you were told?

Checklist of Factors: Largely Developed from Rosenberg et al. (2013) unless otherwise stated.

Factors	**Raised in Discussion**	**Prompted**
Medically indicated		
• *Had an abnormal mammogram of other breast prior to surgery*	☐	☐
• *Had an abnormal ultrasound of other breast prior to surgery*	☐	☐
• *Had an abnormal MRI of other breast prior to surgery*	☐	☐
• *Had diagnosis of breast cancer in the other breast prior to mastectomy*	☐	☐
• *Have known genetic change (mutation) such as BRCA 1 or BRCA 2 mutation*	☐	☐
Prior risk factors		

Factors	Raised in Discussion	Prompted
• *Had prior radiation to the chest putting you at increased risk of new cancer in the other breast*	☐	☐
• *Have strong family history of breast cancer*	☐	☐
Reducing risk		
• *Desire to lower the chance of getting breast cancer in "other breast"*	☐	☐
• *Desire to prevent breast cancer from spreading to other places in body*	☐	☐
• *Feeling at increased risk of getting cancer in the "other breast"*	☐	☐
Safety (not in Rosenberg)		
• *Avoid risks of radiotherapy*	☐	☐
• *Avoid risks of surgery*	☐	☐
• *Impact of post-surgical recovery time*	☐	☐
Longevity		
• *Desire to improve my survival/extend my life*	☐	☐
• *Considerations about family and other commitments (not in Rosenberg).*	☐	☐
Peace of mind		
• *Desire for peace of mind*	☐	☐
• *Worry that screening wouldn't find cancer in the other breast*	☐	☐
• *Desire to do everything that could be done (taking back control)(not in Rosenberg)*	☐	☐
Cosmetic/aesthetic		
• *Desire to have both breasts look the same after surgery*	☐	☐
• *Desire to make breasts look better*	☐	☐
• *Appearance clothed or unclothed (not in Rosenberg).*	☐	☐
External advice		
• *Desire to follow my doctor's recommendation*	☐	☐
• *Advice from friends or family*	☐	☐
• *Self-sourced information from the internet, media or other sources (not in Rosenberg).*	☐	☐
Compliance		
• *Avoid ongoing mammography/breast imaging visits*	☐	☐
• *Avoid ongoing mammography/breast imaging costs*	☐	☐
• *Avoid ongoing monitoring discomfort/reminder of diagnosis*	☐	☐

Results

In all, eleven women attended focus groups held at the offices of the Centre for Health Economics Research and Evaluation (UTS) in Ultimo on 30th October 2014. Nine women participated in a focus group for women who had a CPM, and two in a group for women who had an unilateral mastectomy. The women in these groups had a prior diagnosis of breast cancer, had completed their primary therapy, and had a median age of 57 years (range 45-72).

There were five key factors that were important to the decisions women made:

1. Attitudes to fear of cancer recurrence: For women in the CPM group, a reduction in fear and risk were benefits of undergoing the bilateral mastectomy, but they understood that there was still some risk. For many, their decision was focused on actively working to avoid recurrence due to fear ("to get on with life") and there was anxiety associated with the cancer returning in the other breast. Some of this anxiety was associated with waiting for test results if they did not have the other breast removed. They felt it was important to avoid that anxiety. Women in the unilateral mastectomy group did not express the same fears or anxieties, and were willing to accept the risk of another cancer developing. They found it hard to understand decisions motivated by the fear of recurrence.

2. Learning from the past: This was a particularly strong theme among the women who underwent an unilateral mastectomy, for whom avoiding the pain and negative effects of surgery was very important. For a number of women in the CPM group, a prior undetected or missed diagnosis meant they no longer trusted that mammography or other forms of monitoring would be able to detect a new cancer in the contralateral breast. This, in part, motivated them to have the unaffected breast removed. They also wished to avoid ongoing monitoring and the intrusiveness of further testing. Having observed family members with breast cancer was also cited as a factor affecting women's decisions about whether or not to have the contralateral breast removed.

3. Costs: Avoiding the costs of ongoing monitoring was a factor for some women in deciding to undergo a CPM. Cost was not explicitly an issue for women in the unilateral mastectomy group, but they agreed that ongoing monitoring was an imposition on women. Women also noted that besides the costs of monitoring, there is potentially a high burden placed on women in terms of direct costs (surgery, reconstruction) and indirect costs (lost income) when they have breast cancer. They felt that this is not often discussed with women.

4. Unwanted effects. A factor arising in both groups was the unwanted effects of breast cancer treatment. For women in the unilateral mastectomy group this was dominated by the pain associated with surgery, and was the primary motive for avoiding CPM. For those in the CPM group there were a variety of unwanted effects, but numbness and loss of sensation in the breast area were perhaps the most common.

5. Body image and identity. Aspects of body image and identity arose as a factor influencing the decisions women made with respect to whether or not they underwent a bilateral mastectomy, had subsequent reconstructions or used prostheses, and their perception of no longer having breasts. These factors arose for participants who underwent bilateral mastectomy, but were not raised by those in the unilateral mastectomy group as factors in their decision-making.

While these were the five main factors raised during the groups, there were a number of others focusing on communication (with doctors or family/friends), the stage of life at which women developed breast cancer and the compressed timelines for decision-making.

Allocation of Focus Group Themes

Themes *(may not be mutually exclusive)*	**CPM group**	**Mastectomy**
Trust (or confidence in monitoring)	X	X
Learning from the past (to avoid events)	X	X
Communication - family, friends	X	X
Communication – doctors (involvement in decision-making)	X	X
Body image and identity	X	
Sensuality/sexuality	X	
Fear and risk/anxiety	X	X
Waiting	X	
Costs	X	X
Convenience	X	X
Pain & unwanted effects	X	X
Post validation	X	X
Life stage (includes effects on family & friends)	X	X
Unencumbered/Peace of Mind	X	
Timing – Instant/Delay	X	
Genetics	X	X
Cancer Discrimination	X	

X indicates whether this theme was noted in the transcript for that group.

Trust and confidence

Trust and confidence arose as a major theme in the focus groups, both in terms of the confidence women had in their medical teams and importantly the extent to which they trusted current surveillance methods to manage their ongoing risk of cancer. The principle theme was the lack of trust expressed by women who underwent CPM in the ability of regular monitoring to detect new cancers. This was perhaps the primary motivator for these women as a group, although not the universal motivator, for the decision to remove the contralateral breast: they no longer trusted that mammography or other forms of monitoring would be able to detect a new cancer in the contralateral breast.

Post:	An acceptance that the normal checks and regular monitoring are sufficient, and are all that can be done to minimise the ongoing risk of cancer. There was some sense of acceptance of responsibility for managing ongoing risk related to these comments; the sense that nothing more could reasonably be done other than the regular monitoring, and that it is the responsibility of the woman.

P10: ***You just, you pay attention.*** ***You check*** *your breasts.* ***You have your mammograms.*** ***You have your ultrasound****, you do your normal checks. You have your checks at the hospital. Having said that, the second time I had cancer it came back 5 months after I'd had my mammogram and checks.*

P11: *Same, mammogram, ultrasound* ***that's all you can do****.*

P10: *With regard to the risk of it developing in the other breast* ***the normal checks are pretty good.***

P4: *…my way of dealing with preventing it* ***is just going for my routine screens****, following-up with my medical team*

Pre:	Women expressed a confidence in their medical teams to provide them with relevant information and to guide them in their decision-making. These women had not expressed a loss of trust in ongoing monitoring, despite one being recently diagnosed with abnormal ovarian masses.

P7: *I have* ***total confidence*** *in my medical team. I was totally guided by them. I don't think I Dr Googled.* ***They supplied me with information****. Again* ***I just had total confidence*** *in them.*

P4: *What have you done, what's the outcome like, but then I had a quite, not only was she involved in that research but* ***she was quite experienced*** *so* ***I felt quite confident accepting her advice*** *on whether that was an acceptable treatment or not for you*

P4: *We don't know what they are yet, so we'll see how that pans out. So I guess my worry is,* ***I work closely with my medical team****.*

P1: ***You have confidence*** *in them?*

P4: ***Yeah I do****, and trying to listen to your body.*

P6: *Yes, sometimes you get too many, and the general public can create doubt in your mind but* ***you do get a lot faith in your surgeons*** *and your team and I often say to people that are newly diagnosed that* ***you have to believe that they are there trying to save your life*** *and* ***you work with them,*** *you* ***listen to what they offer, you can get your information, ask your questions, but they are your team****.*

Pre:	The key theme under the banner of trust; this reflected a lack of confidence in ongoing monitoring due to women having the experience of primary cancers or recurrences being missed. It is intertwined with the theme of learning from the past and experience based preference formation.

It was discussed by the group that mammography is uncertain and will not pick up all cancers (particularly non-ductal carcinoma in-situ (DCIS) cancers), eroding trust in it as a risk management option for these women. After those initial experiences mammography is no longer an option for these women, so bilateral mastectomy was how they chose to manage ongoing risk. Lack of trust and confidence in mammography means they did not trust that it would pick up any further recurrences.

A counter point was offered by one participant who noted that perhaps some of the lack of trust is due to unreal expectations about what the medical system can deliver in terms of its ability to detect cancer, and that it is not perfect in this regard. Another offered a distinction between a cancer being missed (and the negative effect this would have on trust) and not being detected due to the lack of a suitably sensitive test.

P6: *In 10 months I went for an out of cycle mammogram and* ***it was found by accident*** *and it was already 4 cm 10 months* ***after a routine one*** *and I'd had lumpy breasts* ***too they were watching. I'd had annual checks but nothing ever found. Biospies, that sort of thing***.

P9: *…a mammogram guided ultrasound and they thought there was a 2 cm lesion there on ultrasound* ***so i'd been missed up to that point***.

P1: *So basically I don't, and* ***you hear so many anecdotal stories that it was missed on mammogram. I see mammography as being very uncertain*** *and really, very..*

P9: ***I think the general public need to know that it doesn't pick up everything, that it's not clear****...*

P1: ***That it picked up two of mine but not the third. So I didn't trust mammography anymore****..*

P2: ***So ongoing mammograms aren't an option***.

P8: *Yet* ***DCIS only shows up on mammogram***

P9: *And that's why because* ***its 80% of all breast cancers****, they've got that out there as the screening process that will pick the majority of cancers up****, the rest will just get missed***.

P9: *I* ***didn't trust the scanning the process***.

P2: *Well I'll just say briefly that it wasn't, and I don't think the system is wrong, it's just a realisation that* ***medicine is an imperfect science and that monitoring isn't fool proof*** *and we all watch too much CSI and medical shows and* ***think they can detect anything, find anything, and that's not so****, and*

as I say it's possibly not even operator error or anything ***it's just simply, it's not perfect****, and so*

P9: *My dad is the doctor, I'm the physio and I just went for my mammogram because* ***my sister was diagnosed and unfortunately it was missed****, and just* ***realised that it's not infallible****, you can't just, and I had ridging on that side and they couldn't see anything. And I just made the decision, look I just don't even want to go there and think that it might possibly not be.*

P4: *There is a* ***big difference between being missed and undetected****. Missed, I would very much lead into your distrust of anything going forward. Whereas undetected, it wasn't showing, or the test wasn't specific enough or fine enough, to me that would give me a reassurance..*

P9: *They even tried to do more fine needle biopsies in my armpit to try to check instead of having to do the node and yeah, shocking, but* ***they all missed it as well.***

Learning from the past

Learning from the past, and seeking to avoid prior negative experiences, was another key theme that emerged from the group discussions. This was a particularly strong theme among the women who underwent an unilateral mastectomy, for whom avoiding the pain and negative effects of surgery was paramount. For those in the CPM group, avoiding ongoing monitoring and the intrusiveness of further tests was a theme that arose. Having observed family members with breast cancer was also cited as a factor affecting women's decisions about whether or not to have the contralateral breast removed.

Pre:	The decision to not undergo a CPM was primarily motivated by a desire to avoid further surgery. Prior surgeries were associated with pain and ongoing discomfort. Women did not see a need to subject themselves to further pain for what they considered potentially unnecessary surgery.

P11: *I had a* ***nerve cut*** *which now I have* ***chronic pain*** *so its* ***put me off*** *quite honestly and I had a* ***haematoma*** *while I was in hospital. I did not get out until* ***day 11*** *from hospital so that* ***put me off*** *too.*

P11: *That's right,* ***the idea of any more pain****, my pain tolerance, even if I get a headache.*

P10: *And so,* ***and the pain. And you think, no I don't want to go through this again, I don't want anymore operations***

P10: ***Why have surgery if you don't need it?***

P11: *…even the following day I said* ***I'm never having surgery ever again for the rest of my life***

P10: ***Constant trouble and pain*** *is difficult to deal with.*
Even if it is not very strong pain it is ***very hard to deal with*** *if it is there all the time.*

Pre:	Two participants referred to the influence of prior family experience, either direct or indirect, as affecting decisions around the removal of the unaffected breast. In one case, this was an explanation for undergoing prophylactic mastectomy (in the absence of a cancer diagnosis) observed from a third party as wishing to avoid what her mother had been through.

P10: *"I'm going to* ***get rid of everything*** *and not going to have it because I* ***saw my mother die****"*

P1: *I had* ***lost my mother*** *to breast cancer which* ***played a part*** *in my* ***decision*** *to have* ***prophylactic mastectomy****.*

Pre:	Once women have had one recurrence of cancer (or second new cancer), the desire to avoid further recurrences, or another new occurrence in the contralateral breast, also motivated decisions to remove the second breast.

P4: *…when that* ***second occurrence*** *of cancer came, it was still a fairly aggressive and nasty one I think I'd* ***already decided*** *way back when that if it ever came back that would be it,* ***two strikes and we're not going to muck around****.*
P8: *For me,* ***I'd had cancer before*** *so it was a no brainer.*

Pre:	Associated with convenience, is avoiding the ongoing testing and invasiveness of monitoring. Women expressed the feeling of being poked and prodded, and the desire to avoid that by having the bilateral mastectomy. They had been through it with the primary cancer and did not want to continue going through the invasiveness of monitoring and testing, so removing the contralateral breast seemed like a good option to avoid that ongoing invasiveness.

P7: *Because it was quite a lot to absorb, the idea of being quite young, then being tested, having them removed, oh that was like, no. Then two years later when it was settled and* ***I'd been poked and prodded I thought why not****?*
P9: *I want a* ***double mastectomy*** *because I* ***don't want to have to worry*** *about having to* ***go through all this again****, and put the family through it.*
P9: *I just didn't want to* ***have to keep going back to have ultrasounds, MRIs, and I just don't have time you know.***

Post:	Avoiding the side effects of radiotherapy was another reason cited for having the contralateral breast removed.

P2: *I was taken aback by the **discomfort and ongoing discomfort from radiation**. I wasn't prepared for it, I just thought radiation, it's going to be nothing after chemo.*

P9: *They **might contribute to your decision to have the other breast removed** because if you do have DCIS or LCIS you are still going have to be irradiated aren't you, so wouldn't you be better off?*

P9: *If you've experienced something terrible on that side wouldn't you be better off **getting the mastectomy to prevent getting the radiotherapy**?*

Communication (family, friends)

Communication with family and friends emerged as a theme important in how women processed the information to enable them to reach their decisions, and importantly how they coped with their diagnosis and treatment (including the implications for their families).

Pre:	Decisions sometimes involved others in the family unit, either in terms of advice or in actually sorting through all the information, as there were many options for women to consider, often in a short space of time and under difficult circumstances. However, even among those women there was a reluctance to provide advice to other women; but a view that such advice might better come from doctors.

P10: *I did **discuss** it with **my husband**.*
*"What do you think I should do?" He said "**I don't want to influence you**." I said you **won't influence me**, I just **want to know what you think and I'll make my own decision**.*

P2: *My husband is an engineer, he literally **had a spreadsheet** from my diagnosis with all the variables **to handle that part of the process**. Do I do chemo, do I do radiation? **Thankfully he just got back to me**. I just don't know how your average Joe blow would even begin in the two weeks they given you to decide that.*

P10: *I **definitely wouldn't tell** anyone what they should be doing, or even say...it's up to them totally. **Doctors** are probably the better people **to tell them**.*

Post:	The communication between mothers and daughters was highlighted, particularly as a way of daughters expressing their fears (or risks), and how this impacted on women with breast cancer. One participant expressed hurt that despite trying to assure her daughter otherwise, she was convinced she would get breast cancer.

P9: *I've also got four daughters so I, I was **quite conscious** on **how I speak** to them, to be as **strong** as I could be present like to my children, **so they felt secure**.*

P2: *And of course they are **four daughters** so that I'm guessing they have breasts, so this is..*

P9: ***They've all asked**, the older two have asked*

P8: *My older **daughter is convinced she'll get cancer** when she grows up and we are going to have the gene test. My surgeon is convinced that I don't need it, that I don't have the gene. **She says, just matter of factly, "I'm pretty sure I'll get it", and that hurts me. I don't want to give her that**, and there's no evidence of it, but **I can't convince her**. We'll go and have the gene test and hopefully that will come back negative*

Post:	There was a sense of individuality and isolation in the experience for one participant, both in terms of the enormity of the experience itself but also in not being able to have others understand the experience.

P6: *Most people around you will **pretend that it's all over**, you've had your surgery and **it's all hunky dory**.*

P6: *And I think the difference is that the cancer, even though that on paper it might appear the same as someone, **each individual is an individual case** and the risks of recurrence or something happening need to be nailed down. **Explaining that to someone that hasn't had a piece of paper put in front of them is very difficult, they tend to see it black and white**.*

P6: *And so **it's not really understood** how big, and **all the drains, its huge surgery**. **If you go through, and add in the reconstruction and you've got more drains, more stuff, it's very significant surgery that is underrated**.*

P6: ***I think personally it is underrated. People don't understand**. They are very **supportive** and they're **doing the best for you**, which is often **visiting and being with you which is exhausting in itself**.*

Communication with doctors and decision-making

The themes which emerged around communication with doctors centred largely on the decision-making process, whether women felt they were involved in that process, and the manner in which they were informed of their cancer diagnosis. Involvement in the decision-making process was the key aspect of this theme in terms of women being participants in that process, but feeling they could seek the advice of their doctors.

Pre:	Some participants seemed to express displeasure at the manner in which doctors communicated with them as part of the treatment decision-making process. In some cases this seemed to indicate that women were

not being involved in the decision-making process, but rather told what would happen (rather than be presented a course of possible options) and were passive recipients of those decisions. This was particularly apparent among the two participants in the unilateral group who were both BRCA positive, but chose to not have the unaffected breast removed (even though it was presumed they would).

The corollary was that other participants very much saw themselves as the decision-makers, the options are presented by the clinicians, but the patient makes the decision.

P11: *It wasn't suggested, it was **automatically presumed** that that's what I would do.*
*I find that's **very much how they speak to** you.*
*And I was always told, "**Don't worry** when you have your reconstruction we'll do this we'll do that". And I'm like, **I haven't thought** about that yet.*

P11: *You know, and yeah **they talk about your mastectomy**, and **tell you all you're going to have done** which I find..*

P10: *I didn't find they tell you, they strongly suggest..*

P11: *I wasn't told, **it's the way they speak to you**. It's, the way they speak to you, it's **you'll be doing that**.*

P6: ***They go through the options, but ultimately it's your decision**. When they ring to give you that result I was jotting down lumpectomy, radiation, mammogram and as the conversation unfolded I just went: 1 mastectomy, 2 mammogram, so I was **trying to digest the information** and jot it down. It's your gut feeling that comes clearer, but **you need the back-up information to justify that decision** I suppose to yourself if you are going to put yourself through **that major surgery the total, the distress to your family**.*

P8: *Do you think I should get them both off or you know the right breast is healthy. "**My personal opinion get them both off, get a reconstruction at the same time and be done with it**".*

P9: *So I had the ovaries out and she was saying: "**If I was you I'd have the other breast off, get it all done, you'll be feeling ok**".*

P6: *Yes, sometimes you get too many, and the general public can create doubt in your mind but **you do get a lot faith in your surgeons** and your team and I often say to people that are newly diagnosed that **you have to believe that they are they're trying to save your life** and **you work with them,** you **listen to what they offer, you can get your information, ask your questions, but they are your team**.*

Post:	However, doctors were not always willing to accept the patients' decisions about whether to have the unaffected breast removed or not. A number of participants noted that their doctor attempted to change their initial decision.

P10: ***Every time*** *I see my surgeon she is* ***on at me about having a prophylactic one****.*

P2: *Yes, she did try to* ***talk me out of it****. Did a couple of sessions to say,* ***we could have an inflator****.*

P9: ***I want*** *a double mastectomy because I don't want to have to* ***worry*** *about having to go through all this again, and put the family through it, but* ***he was sort of reluctant*** *to because he felt the* ***tissue on the other side was healthy****, let's just get through this first.*

Post:	Experiences with decision-making and the manner of communication from doctors was mixed. For some participants, there was a feeling that the communication from doctors had been poor and this had impacted negatively on the treatment experience. This was not universal, some participants were happy for their doctor to take control of the situation and to make decisions on their behalf and inform them over the phone. This seemed to convey the sense of urgency of the situation. There was a suggestion that this might have been due to the medical team being all female.

P3: *He* ***didn't say anything*** *about why he was going to take the glands out on the right side, and* ***I still don't know why*** *he did. I asked him.*

P4: *The only time I've come up against any resistance to treatment in my journey of all of this was when I was initially being investigated right at the* ***beginning I was sent for ultrasound and mammogram and they refused to do the mammogram****, and I was a* ***bit angry*** *about that at first. …. Having said that when I went in for surgery* ***they did do a mammogram****, but at the time I was a bit cranky about it,* ***thinking why not****?*

Moderator: *Perhaps they should have explained that a little better.*

P3: *My* ***doctor rang me*** *up and* ***told me it was cancer*** *which* ***I wasn't impressed*** *about.*

Moderator: *You would have preferred to have been called in to be told in person?*

P3: *I* ***would have liked, I'm not sure that's how you tell anyone****. And even the surgeon, he told me if they find a second lump you'll have to think about it, and* ***he even rang me up while I was in hospital and told me he found the second lump. And I went you could have come in and told me****.*

P4: *What happened to me X with my second one, they sent me a copy of the pathology report, a copy of the report so I knew before I even got to my doctor. An envelope came and I thought it was the bill I opened and all I could see was...*

P7: *My doctor told me slightly differently and **I didn't mind** at all…. she rang me about 7 o'clock at night and said "I've booked you in to see a surgeon, I haven't got the results yet but that's what I'm thinking", and I said **"so that's not good"** and she said "yeah", I've but **I didn't mind that because I found that there was a sense of urgency**, …, I didn't mind being rung up.*

P3: ***I didn't think it was the correct thing to do because he didn't do anything he just rung up and said yes it is cancer**.*

P6: *So were you feeling blindsided?*

P3: *And **I told him you he should never do that again**.*

P4: *It's like the feeling you get when you are having your scanning done and they leave the room to get someone*

P5: *And **they sort of forget there is a person attached to the bit of body they are looking at**, that there is someone else in the room.*

P7: *I had all women again which made, the **multidiscipline approach I had was all women which I think made a bit of a difference**.*

Body image and identity

Aspects of body image and identity arose as a theme influencing the decisions women made with respect to whether or not they underwent a bilateral mastectomy, had subsequent reconstructions or used prostheses, and their perception of no longer have breasts. These themes arose for participants who underwent bilateral mastectomy, but were not raised by those in the unilateral mastectomy group as factors in their decision-making.

Pre:	One of the considerations for some participants was the impact of having their breasts removed on balance, both in terms of the visual aesthetic and the physicality of having two breasts. There was a feeling that this could not necessarily be addressed with prostheses, that these either were not appropriate or would not always result in the desired symmetry.

P5: *Having had **two boobs**, and **one boob** I'd rather have **no boobs** at all.*

P4: *I was actually very small busted and to actually get a bra to fit I ended up having, **using prostheses** on both sides, bilaterally **to balance things out**. And so I'd had the opportunity to have a couple of months, unbeknownst to me I was going to have both breasts removed, **of wearing prostheses**. It really helped me deciding what I was going to do, **I'd been down that street**. **It's not for me.** I'm not someone who gets up in the morning and puts lots of makeup on and looks in the mirror and spends lots of time grooming, I just*

chuck clothes on and fly out the door. ***I found they weren't the thing for me. I think that swayed me as well.***

P9: ***It's a bit weird*** *when you are* ***uneven*** *and you're trying to put it altogether. So I'd pretty much decided still that I* ***wanted to have the other one off****, and I could have skin sparing on that side and have an implant.*

P2: …but people who aren't involved think it's so simple, you just ***buy some new breasts*** *and they'll be ok, they'll* ***be all perfect*** *you know,* ***symmetrical****, it will be* ***problem free****, but of course you know, thankfully* ***it is most of the time but it's not always, it's not always symmetrical, it's not always perfect****.*

P2: And ***then you have the symmetry. It's like carrying a shoulder bag.***

Post:	For some participants, there was a benefit from bilateral mastectomy from no longer having to wear bras. Another expressed disappointment because it meant an unexpected change in her behaviour not necessarily related to her breasts, but associated with wearing a bra (the use of bra straps to store hankies or eye-glasses).

P2: If there is one ***upside of bilateral mastectomy*** *it's that* ***you don't have to wear a bra. Carry this extra weight.***

P1: I just think you can talk about your own personal experience, for me that's been the relief, and ***it's great not to have to wear the bras anymore****.*

P5: I ***miss wearing my glasses there****. That's the most annoying thing.* ***Where do you put your hankies without a bra strap****?*

Pre:	How participants related to their breasts also differed, impacting differently on the decisions they made and how they felt about those decisions. While there was a general theme among the participants who raised this topic of not being defined by their breasts, one expressed that they are a key part of her femininity and she mourned their loss after a bilateral mastectomy. This was a key motivator for her to have a reconstruction. In contrast, another participant saw her breasts as separate from her identity, a piece of equipment, and a threat due to the potential for them to become cancerous; and therefore something that "had to go". This raised the question of which generated more fear – losing the breasts or keeping them. For another, keeping her breasts was not worth risking her life.

P5: *But now, it's* ***so long since I had any breasts****, I don't really care because* ***I don't think who I am is defined by the shape of my breast****.*

P7: The thing is it's a ***big decision*** *to have your* ***breasts removed as a woman****.*

P7: ***Not*** *that we are* ***defined by our breasts****, but I think it's something that probably the* ***odd man doesn't understand****. ...it* ***doesn't define you*** *but it is*

very much a part of your femininity *hence* ***why I had a reconstruction um, having breasts****.*

P7: *I loved my breasts, but* ***I mourn them****. I have absolutely* ***no regrets*** *but.*

P8: *It's not worth* ***risking your life for them****.*

P6: *I think that's again a stage of life, I was 54 and thought,* ***I don't need these anymore, they're equipment.***

P6: *I didn't feel that. I felt* ***they were threatening me and had to go****. So* ***I had no qualms****, I'd made my decision very slowly.*

P6: *..what's* ***scaring you the most, keeping it or getting rid of it****?*

Post:	Aspects of body identity and aesthetic were also important to the decisions women made regarding the immediate post surgical period. One strategy used by a number of participants to enable them to cope with the removal of their breasts was to have reconstruction, or skin sparing, at the same time as their mastectomy. Participants expressed a desire to not wake up to an empty bra or "slits" immediately post surgery. However, one participant noted that despite skin sparing, it was still confronting and emotional the first time she was topless for radiotherapy after her surgery. For another, her decision about reconstruction was influenced by her mother's desire for her to 'look normal again'.

P9: *….having the* ***expander*** *when I had my* ***first mastectomy helped me come to terms*** *with what happened with me because it happened so quickly at the beginning, being able to* ***wake up*** *after having the skin sparing mastectomy with the expander and* ***have something in the bra helped me come to terms with it at the beginning.***

P4: *Yeah I did say to my surgeon, and* ***I want bilateral****, and* ***I want reconstruction*** *at the time. She did question me about that and I said, "****When I wake up and look down I don't want to see two big scars, I want to see something****."*

P6: *Yeah, I had that feeling* ***I didn't want to look in the mirror and see some slits****.*

P9: *I found it* ***quite confronting*** *because it was really the* ***first time that my chest was exposed to quite a range of people standing around*** *me and I was like this on the table and I found it* ***quite confronting to start with****. And I don't think they meant to be but the way they draped me and did the,* ***I found it really exposing****. I had a* ***really strange emotional reaction after the first one, after that I was ok****.* ***I must admit, I shed a few tears****.*

P6: *I have very **elderly parents**, 94 and 93 and I think that my **mother** just wanted me to **look normal again**.*

Sensuality/sexuality

Post:	There were differences in views regarding the impact of bilateral mastectomy on sensuality and sexuality. For one participant, this did not pose an issue in terms of the loss of sensation, but for others the numbness and loss of sensuality was very important.

*P2: And I guess when we **talk about making a decision**, and I don't know how you want to compress this, but for me I didn't have super sensitive breasts I mean sort of so big, from the **sensuality perspective I wasn't you know so anxious about that numbness about or loss of sensation**, so..*

*P7: I must say **I can't join you in that**...*

*P2: **For some women**, that's you know, **clinging to her other side because they mean so much to her its great**.*

*P6: **I really miss the skin sensation**. **I think that's very under-rated, that numbness**. I think that more research into nerve conduction and getting over that scar tissue.*

Fear and risk

The themes of fear and risk were very important in the decision-making process, and to how participants perceived the decisions made by other women. Participants in the CPM group expressed that a reduction in fear and risk were benefits of undergoing the bilateral mastectomy, but that they understood there was still some risk. Participants in the unilateral mastectomy group felt they were less fear focused and were willing to accept the risk of another cancer developing; they would deal with it if it happened.

Pre:	Participants in the unilateral mastectomy group were not motivated by fear, and questioned the view that cancer risk should be avoided at all costs. The perception was that some women who undergo prophylactic mastectomy are fear driven. One felt that perhaps it was linked to life-stage and could understand it applying in young women. Cancer was seen as something to deal with if it happens again. For these women, they were willing to accept the possible risk of a new cancer in preference for avoiding the possible pain associated with another operation.

*P11: But all I heard from her was how **she had everything removed**. Then she went on to talk about how they had found something in the intestine or something. And I sat there thinking, "**Oh for heaven's sake don't tell me you're looking into something like that now**".*
*She was **so fear driven by cancer** and when we all went for lunch that was the*

discussion at lunch. I sat there thinking, everyone was talking about how ***terrified this woman was of getting cancer****. And I thought* ***you could get it anywhere. Or you could have a car accident and die****.*
I just found the whole notion, well ***it was fear of death really, that's what she had.***
*I just found that really...****I've never struck that, even when I've had treatment and been speaking to other people. I found that odd****.*
I wonder if in a way you are desensitised to it too?

P10: *And* ***I'm not terrified of developing it in the side that's left anyway****.*
Some people are. ***If it happens it happens****.*

P10: *Well by that time I'd had time to think because from diagnosis to operation was only a couple of days so I had time to think, and as you said the pain. I'm thinking I don't want to go through it,* ***it's perfectly healthy, I might never ever develop cancer in that side.***

P10: *I can* ***understand young women choosing to have them off to reduce the risk****, umm it's easier when you are older, but even so* ***some people want to reduce their risk down as low as they can****.*

Pre:	Participants also revealed general attitudes to risk that may have influenced their decisions and behaviour. One CPM participant noted that she is generally particular about all types of screening, due to family history, suggesting risk aversion. For another, prior family experience of cancer led her to have a mammogram, suggesting risk aversion. In contrast, a participant in the unilateral mastectomy group seemed to be less risk averse and downplayed fear of breast cancer as just another risk of death.

P1: *So* ***it's not*** *for me* ***just about breast cancer*** *I'm very* ***particular about having follow-ups and screenings*** *and I've had some precancerous polyps removed.*
P9: *I had* ***mammograms*** *from the* ***age of 40*** *because my* ***oldest sister was diagnosed*** *with ductal carcinoma so I* ***just decided to start having mammograms.***
P11: *And I thought* ***you could get it anywhere. Or you could have a car accident and die****.*

Pre:	An important theme among participants in the CPM group was the idea that if you remove your breasts, the potential cause for fear is also gone; there is no longer a source for worry, suggesting reassurance or reduced anxiety, because the breasts are not there.

	However, there was a feeling that this varies with cancer type and that any treatment is not a 100% cure, there will always be some residual risk.

P5: *So I **don't have to worry** about **getting cancer** in my breasts **because I don't have any**. But you still worry that something could recur in the little bits that they do have.*

P8: *But it's the health thing, **worrying** every year that you know **she might get cancer in the other one**.*

P7: *I don't think the **removal of the healthy breast** though is unnecessary if that's what the person chooses to, because the mind is just as important, so **worrying** about it can be just as **detrimental.***

P6: *I agree with that because **they assume that by having surgery you are cured**. Ok, it **gives a 98% chance of not getting another cancer, but it is cancer**, and they are cancer cells. Everyone knows **there can be mistakes** if something has got away when you're having the hook wire biopsy. If **something escapes** you know **you are still at risk**.*

P7: *…she highlighted the **risk of it coming back.***

P2: *So **it doesn't wipe that away completely**, it's not some magical you know, **"I've got new breasts I'm never going to get disease."***

P7: ***I don't worry about getting breast cancer at all**.*
*When I was diagnosed it was quite you know, aggressive and all of that and I knew it was bad but I knew I'd get through it and I always had the feeling it would come back. I had an optimist pessimistic feeling. **I actually don't worry about getting breast cancer.***

P8: ***I don't feel like I have cancer anymore, I don't feel like I'm at risk**. I don't have that "ooo" hanging over me at all. I feel better now than I have for a long time. I don't know why.*

P8: ***I skip around the room, thinking I don't have breast tissue**.*

Anticipatory anxiety

Apart from the expression of fear associated with cancer recurrence, or a new cancer forming, participants also expressed a desire to avoid anxiety associated with waiting for test results and the uncertainty of a cancer detection. While this influenced decisions to undergo CPM it was not raised by participants in the unilateral mastectomy group.

Pre:	For some participants, one of the things motivating their decision to have a bilateral mastectomy was the desire to avoid the waiting and associated anxiety of further testing and procedures if they kept the contralateral breast.

P7: *And then that **waiting** each time there's a lump*
P7: *…being tested all the time, that **waiting**, and **anxiety**. I thought **I don't want to live like that…***
P5: *It's the **uncertainty** every time.*
P7: *It's the **anxiety** I think, the **uncertainty** every time.*
P5: *It's also the **trauma of wondering** what they're going to find, the **worry**.*
P7: ***Waiting** for the results. Yeah.*
P6: *Being the last one in the waiting room. Everyone else has gone home.*
P5: *Everyone else has had their results.*
P7: *That **waiting was doing my head in all the time**, needed to have a break for a couple of days.*

Post:	One participant expressed anxiety about what type of reconstruction to have and its impact on subsequent monitoring and the ability to detect a recurrence.

P2: *What sort of monitoring, because that's some of **my anxiety about reconstruction**, is that it could **hide the detection** in the chest wall when it recovers.*

Costs

Costs arose as an important theme among participants in the CPM group, in two ways; in terms of the impost to women associated with ongoing monitoring and testing once they have been diagnosed with cancer, and the costs (including their work time lost) of undergoing treatment for breast cancer and reconstructive surgery. Avoiding the ongoing monitoring costs was a factor for some in deciding to undergo a CPM. Cost was not explicitly an issue for women in the unilateral mastectomy group, but they agreed that ongoing monitoring was an impost on women.

Pre:	For some women, avoiding the ongoing costs of monitoring for cancer (screens, biopsies, MRIs, etc) was an important factor in their decision to undergo a CPM. It was suggested that undergoing CPM was perhaps less expensive than the accumulated cost of ongoing monitoring. The key issue here was that once diagnosed with breast cancer, women are no longer eligible for services provided by BreastScreen and are often required to pay for services privately (no bulk-billing). This was not a consideration for women in the unilateral mastectomy group.

P10: *I **don't** find costs a problem.*
P11: *No, the **MRI is** though.*
P10: *But the last time I had my mammogram it was bulk-billed. I've never had that before.*

Mod: *No I suppose it depends on where you are in you are in your life stage, if you've got children and expenses, but once they grown up it makes a difference.*

P10: *Yes, you don't worry about finding the money. Yes*

P5: *Because it's going to be a matter of* ***coming back*** *for* ***more core biopsies***, ***MRIs***. *And apart from anything else, the* ***cost***, *the* ***financial cost*** *from all* ***these things*** *because* ***once you're diagnosed*** *with cancer* ***BreastScreen wipes you off*** *their books and* ***you have to pay*** *for things which for me should be the reverse.*

P8: *..* ***it is expensive***, *all the things, then the* ***biopsy***, ***mammograms***, *I don't know* ***what people who are skint do***. ***Do you just die****?*

P5: *That's right they write you off. This surgery and that surgery. All the things for this I had to pay for, with the idea that my surgeon was saying that* ***even if you just have the lumpectomy*** *on this side* ***it would be the same as you having to have the MRI and core biopsies***, *and she actually said that I've already decided and I've told my family, if you tell me I need a lumpectomy I'm telling you I need a mastectomy, and she said "X, that will make your life and my life much easier, good go for it".*

Post:	Participants noted that reconstruction was costly. Those who disclosed their status in the CPM group had all undergone private surgery. They justified the costs of the CPM and reconstruction because it meant they were no longer subject to ongoing tests, inconvenience and had a reduced risk of breast cancer.

P7: *the* ***expense*** *you speak of…*

P7: *the reconstruction was done by a plastic surgeon.* ***Lots of expense***, *but I'm* ***not being prodded and poked****...*

P8: *I went into hospital as a private patient in a public ward because it was plastic surgery I was in a burns ward. They sent me estimates, the plastic surgeon and the surgeon sent me estimates of what it was going to cost me and so far, touch wood, I don't know what's happening..*

P7: *I was* ***out of pocket around the $8K*** *for the reconstruction surgery and that's 7 years…*

P8: *I haven't paid for mine yet, but* ***I'd say with all those other things it's thousands of dollars.***

P8: ***What does someone without money do?***

P4: ***They wait?***

P8: ***They wait then they die. That's not fair.***

P9: *I um, yeah, well* ***I went private***, *I had top hospital cover. I didn't have extras but* ***I'm out of pocket about $4K.***

Post:	Apart from the direct financial costs, participants in the CPM group also highlighted the impact of having breast cancer, and its treatment, on their productivity and loss of earnings. They saw this as adding to the overall costs.

P2: *Not to mention the **impact on your productivity**, you're **not at work**..*

P4: *….it's only now 10 years later in retrospect that I can look and see **what impact financially** it's had on our family actually. It's actually had **a significant impact**.*

P6: *And I think too you are **in shock that first 12 months** until you get over that benchmark of surviving the first 12 months, I've made it this far you carry on. **But the work thing.** You don't want to lose all these parts of your life because **you have to go to the doctors** and **have all these treatments and spend time in hospital.***

Post:	Communication around costs seemed to be a concern in that some women felt they had received little information about potential costs, while others were unsure of how such information might be communicated without seeming to be unethical. There was also discussion that perhaps the assessment of costs, and the potential trade-offs between having a CPM and the ongoing cost of monitoring, might start to be used as a justification to push women into having a CPM where they might not otherwise have one.

P4: *One thing that was **never mentioned** for me when making the decision about what to do and what not to do, just to go back to the **financial side** of it, so no-one ever said to me "**if you keep that second breast over the next 10 years it is going to cost you X,** you are only 30 and **you've got another 50 plus years of this expense, of ongoing expense**". So yeah it's something that's never mentioned by the medical profession what the **ongoing costs** will be.*

P7: *It's a hard one, for **ethically speaking of costs** when someone is making a decision based on that. There are so many different ways you can explore to support payments and everything like that I suppose.*

P2: *It's a dilemma too, because there are **real costs** to ongoing monitoring as well to the **Medicare system** so I'm not surprised that the Medicare system will pay for non-diseased breast viewed from a purely you know medical viewpoint, but then you wonder will it reach a point where somebody does do the math and you know, **holy cow it's heaps cheaper to take it off** you know. Then is there a push to have people forced down that path.*

*Feeling some pressure, because there are kids that need immunisation, and **if you keep that breast it will cost our country**. You know I mean, that's silly I know but…*

P4: *We have moved to some of, partly to the American health system, and I know their system is different to ours but there are situations **where if someone was***

to stipulate to you prophylactic because they didn't want the ongoing expense of this or that, we are not in that situation yet

Convenience/imposition

There were two main aspects to the theme of convenience. One which arose in both groups centred around physical accessibility to services. The other which arose in only the CPM group was related to the impact of having continuing monitoring and testing, following a diagnosis of breast cancer. Some women considered themselves lucky at being able to access the services they needed. Avoiding the need for ongoing monitoring and testing was a factor for some participants in their decision to remove the unaffected breast.

Post:	One participant noted some inconvenience in parking to see her specialist due to lack of parking at her hospital. While there were no women present from rural areas, two participants presented anecdotal accounts of the difficulty for rural women of accessing services for treatment. They commented that the lack of services in those areas meant those women often did not have a choice as to how to manage their cancer and ongoing risk – it was mastectomy/CPM since other services were not an option.

P10: ***Physical access****, because my surgeon is right near the hospital as well it doesn't help because it is* ***always impossible to park*** *near hospitals, but I'm well enough that if I have to park blocks away I can just walk down anyway.* ***Parking is a big deal*** *for anyone trying* ***to get into hospitals or doctors****.*

P4: *I think X has pointed out having a family and particularly being a sole mum where you have to deal and still have to look after all those things for your children stuff like that* ***your situation really impacts on it a lot****, and I think you'd get* ***different answers if we were a remote or rural community*** *sitting down to talk. I think* ***we've all****, while we might make different assumptions, having the* ***luxury of making different decisions*** *on the basis that we are* ***able to access services****. Yes I know I can do this because I can go and have my screening done da da da because I can, I just pick up the phone and make the appointment,* ***I don't have to travel 8 hours to get to the nearest whatever, or juggle too many other commitments.***

P6: *Rural women, according to my brother,* ***rural women will opt for a mastectomy because it reduces the amount of visits to town, or Sydney****.*

Post:	Some participants in the CPM group expressed a sense of being lucky, derived from being able to access services.

P2: *I'm being facetious.* ***Lucky****. But,* ***we had*** *the* ***services available****.*

P9: *I **feel really lucky**.*
P7: *Me too. **I think we all are**.*

Pre:	Apart from the impact on costs, participants also cited the imposition on their time and inconvenience of the ongoing testing (mammograms, biopsies, MRIs), as a factor in deciding to have a CPM.

P5: *Because it's going to be a matter of **coming back for more core biopsies, MRIs**….*
P7: *…during a 2 year period **I was being tested, poked and prodded and tested***
P5: *…the idea that my surgeon was saying that **even if you just have the lumpectomy** on this side **it would be the same as you having to have the MRI and core biopsies**, and she actually said that. I've already decided and I've told my family, if you tell me I need a lumpectomy I'm telling you I need a mastectomy, and she said "X, **that will make your life and my life much easier**, good go for it".*
P7: *…they **wanted to keep doing MRIs** and they thought that would be the thing that would be picking things up, and yeah I just thought **yeah I don't need to**.*
P6: ***Time management** does come into play doesn't it? **Do I really want to do this every three/six months? I've got better things to do with my time**.*

Adverse effects of Treatment

A strong theme for both groups were the adverse effects or sequelae associated with the primary treatment received for breast cancer. For women in the unilateral mastectomy group this was dominated by the pain associated with surgery and its sequelae, and was the primary motive for avoiding CPM. For those in the CPM group, there were a variety of adverse effects, but numbness and loss of sensation in the breast area were perhaps the most common. One participant noted that fear of an adverse effect – infection secondary to prophylactic mastectomy delaying chemotherapy – resulted in her delaying her decision to have a CPM.

Post:	Pain associated with the primary mastectomy and subsequent sequelae was a primary motivator to avoid CPM in women who had an unilateral mastectomy and opted for ongoing regular monitoring as their means of risk management.

P11: *I had a nerve cut which now **I have chronic pain** so **it's put me off** quite honestly and **I had a haematoma** while I was in hospital. I did **not get out until day 11** from hospital **so that put me off too**.*
*That's right, the **idea of any more pain**, my pain tolerance, even if I get a headache.*
P10: ***Pain** has a lot to do with it. You **feel pretty bad** after having **mastectomy**.*
P10: *Constant **trouble and pain** is difficult to deal with.*

*Even if it is not very strong **pain** it is very **hard to deal with** if it is there all the time.*

Post:	One participant provided anecdotal evidence of a side effect of reconstruction; a feeling of being full due to stomach tightness.

P10: *I know I spoke to a lady once who had actually had the tram flap done. And I said what's it like? And she said "good" but she said "my **stomach is very tight** if I have a big meal I'm very conscious of it". And that is something that is **never mentioned**.*

Post:	Aligned with body image, two participants in the CPM group noted the loss of balance associated with having a breast removed, and potential problems with using prostheses that do not stay in place.

P2: *I was very big busted, so the **idea of asymmetry** was very **unappealing** and I can't imagine, certainly once I had **radiation I had burns**, trying to **wear bras to support the other breast**. You had just one was it difficult to..*

P5: *The **balance was really bad**. Your **centre of gravity is put out**. It was just really weird. I felt as if I was **falling forward all the time,** as if I had a weight on one side of me but not on the other, but that sort of dissipated after I had the other, the second mastectomy.*

P5: *- I forget what it's called - with the prosthesis in it, and you wear that bra that's like a harness, but during the day you look down and **you can see the prosthesis, you've got it poking out. They are not practical.***

Post:	One participant in the CPM group expressed a fear of radiation exposure associated with ongoing monitoring.

P2: *Also the MRI. This is **lots of radiation exposure**, how many years do I want to have biannual MRIs or whatever else I mean.*

P5: *MRIs don't have radiation.*

P2: *Ongoing everything else though as well.*

Post:	The issue of lymphoedema was raised in the CPM group, with one participant suggesting that perhaps there was no point in having a reconstruction if it resulted in her experiencing lymphoedema.

P7: *I think I'm actually getting past that, I think they say 10 years or that's what I've always been told.*

P2: *I would still **be very vigilant** on not having anything done on that side.*

P7: *Oh I don't have blood tests or blood pressure, but I've got to say I carry stuff all the time, I..*

P2: *Oh they say now do everything, its not like the old days when they said don't carry anything. It's like when I think about reconstruction,* ***what's the point of having fabulous cleavage*** *if I end up with a* ***giant lymphoedemic arm*** *because of the extra burden of the surgery...*

Post:	For participants in the CPM group, the loss of sensation and numbness due to the removal of both breasts and reconstruction was an issue. This not only leads to loss of feeling, but to potential secondary injuries (such as sunburn).

P8: *I'm* ***numb*** *down here,* ***numb*** *under here,* ***numb*** *here*
P6: ***Chasing things*** *down your chest.*
P7: *I* ***burnt myself quite badly*** *once and didn't know and wound up with a very bad blister, so.*
P6: *Because* ***you just can't feel****.*

Pre:	One participant in the CPM group expressed anxiety about developing an infection post CPM delaying chemotherapy.

P2: *And I think that the bilateral, and I was* ***anxious about developing an infection*** *in the healthy breast. I had a friend that happened to and that* ***delayed her then moving to do her chemotherapy*** *for months. Well I mean...[sigh]...you had this, what is* ***perceived as an unnecessary surgery and that's delayed your cancer therapy****. Well* ***nothing could be more terrifying****. And I think, that's that, well thankfully that doesn't happen now.*

Post:	The overall impact of surgery was an important point for one of the participants in the CPM group.

P6: *Because I* ***looked all right****, you* ***look healthy****,* ***feeling good****,* ***feeling positive*** *but you're* ***exhausted****,* ***mentally****,* ***emotionally,*** *you* ***need to lie down****.*
The ***body hasn't healed****. It's* ***huge surgery****.*

Post:	Linked to body image, some participants in the CPM group referred to their unaffected breast as separate entities, seeking revenge for being removed. Another referred to breasts as "misbehaving" and something the woman can not control.

P2: *Then I* ***did have cerronies*** *and* ***lots more problems on the non-cancer side****.*
I thought ***it's seeking revenge****, but that was* ***my choice****, I* ***can't blame anybody*** *else.*
P5: *….****it has given me more grief*** *in the last year, no problems with the right side at all, and* ***I get lymphoedema all around the scar****. And I have this* ***constant dragging feeling****, um and as you said* ***I feel as if it feels as if it was left out****.*

"Attention! Why didn't I have cancer? Now I'll give you something to moan about".

P6: *This is where a lot of the **misbehaving happens**, it seems to me when a breast starts **misbehaving** you can't control that, **it's just going to spring back whether it's lobular or invasive**.*

Post:	Participants in the CPM group also noted that there were side effects associated with radiotherapy.

P2: *They say a **bit of sunburn**. I had **skin peeling**, it was just so….*
P8: *I mean for the last week where they intensify, they zoom in on a little spot, they got a **little bit sunburned**, not bad. Maybe two hours in the sun that sort of sunburn.*

Post validation

A number of women in the CPM group reported that cancer, or abnormal cells, had been detected in the unaffected breast after its removal. This is a form of post validation and adds to their lack of confidence in the detection process. However, not all women reported this to be the case with an instance of cancer recurrence in the treated but not the untreated breast used to justify the decision to not undergo CPM (by a participant in the unilateral mastectomy group). These post-validations were sometimes reflexive, with anecdotal stories (of breast cancers missed, or regret at having undergone CPM) offered by two participants.

P10: *But I've **not developed breast cancer** again **in the untreated breast** it came back in the treated breast, **so I could have had extra surgery unnecessarily**.*

P2: *Of course **I'm two years now post my diagnosis** and a year finishing treatment and of course I guess what I know now is that is that **it's not that simple** and I'd really love to be able to talk to the earlier me. Although, having said that my **left breast which was negative** on the MRI, ultrasound and **mammogram did have cancer in it and was taken prophylactically**, so I can't, can't beat on myself about that decision.*
P6: *And **taking the other breast prophylactically was a good decision because there was hyperplasia and that was misbehaving, so it was prevention.***
P8: *So I'm really happy with my decision. Oh and also, **they found cancer in the healthy breast**.*
P9: *So I've now got implants and **they found lobular in situ on this side**.*
P2: *I'm **not regretful because we did find cancer on this side**,*

P8: *But the **number of people in this room who'd said that they had the healthy breast taken off and it was found to be cancerous, that's telling isn't it**?*

P8: *..a friend who had a single mastectomy and when I told her what I was going to do she said "It was the best decision. **I have great regrets I kept this breast and got that one cut off**" and now she has to go every year and have the mammogram, ultrasound on one breast. And she said **what's the point**, you know why, and she **almost wants to go in and have the healthy breast, you know, cut off years later**.*

P11: *I **talked to a lot of women who have gone through it** [CPM] and they're **not happy**.*
*There just seems to be a lot of women, not that I've spoken to hundreds, but the majority that I've spoken to are **unhappy with the reconstruction and the whole process** of it and they've said that they **didn't really feel like they needed it, shouldn't…. no needed've**.*
*I think some of them **felt there was a little pressure from family members and so forth**.*

Life stage

Where women were at in their lives, in terms of age and having children was discussed in both groups as a factor affecting the decision to remove the contralateral breast. One participant in the CPM group raised the topic that life stage is also important in dealing with effects of breast cancer, and in her case this was complicated by elderly parents.

Pre:	The view that breast conserving surgery is perhaps linked to age and life stage, for younger women and those with young children, or the potential to have young children.

P10: *All **my children were older** the second time which **made it easier***

P4: *At the time **I was still breast feeding** my 1 year old, so I made the **decision to go for a lumpectomy and the other treatments**, radiotherapy and chemotherapy, and **opted to have the follow-up screening** through my medical team which I did.*
I thought I might actually use that other breast.

P4: *I often wonder **if age has a lot to do with it**, where you are in your life span, because I know **for me it really did** and I've been very comfortable with all the decisions I've made along the way, even looking back now I wouldn't have changed a thing*

P6: *Even at **my age, where you're finished with children** and all things are out of the way **people still ask me did I think I made the right decision**.*

P3: *I decided the **best time to get breast cancer is when you're 60** because you **didn't have any responsibilities, you're still all right**, so always tell everyone if get your breast cancer get it at 60.*

P4: *I think I was **lucky to get it young**, because I was **younger fitter and healthier**.*

P1: ***I wasn't interested at my age and stage** [in breast conserving surgery], I think it could have been done better at the time..*

Post:	One participant expressed the view that attempting to deal with the stress of breast cancer with the immediate family is difficult, and it is exacerbated by the presence of elderly parents who are also dependent on you for help.

P6: *I **don't want to have anything else** done **while they're alive** because it's **just too stressful**. Umm, it's very hard.*

P2: *That's very interesting.*

P6: *Juggling **your own children** they **hover**, they **hover**, they're **very concerned** and **scared** for you and them, but **juggling elderly parents** who are **not able to help** in any way or form is **even more stressful** because **you have put on this front of managing.***

Relief and Peace of Mind

For some participants in the CPM group, having the procedure enabled them to obtain relief and peace of mind. It was the motivating factor for the decision to have a CPM for one participant. Making a choice based on what gives you peace of mind was the advice one participant would give to other women with breast cancer.

P9: *I just didn't want to have to keep going back to have ultrasounds, MRIs, and I just don't have time you know. **Just get it sorted**. And I do, I **feel like this weight has lifted** from my shoulders, I **can just get on with life**, **get back to work**, 2-3 days a week, you know the **kids all feel more secure**, it's much better. That's why I made my decision.*

P1: *I just think you can talk about your own personal experience, for me **that's been the relief**, and it's **great not to have to wear the bras anymore**.*

P1: *And we know tumours can be present and not detected in mammograms, **it's just the relief it gives you.***

P3: *"Sorry nanna that hurt, it hurt mummy". I said "no it didn't". She said, he said "why". I said "I've got pretending breasts" [Audible laughter]. And he said "what are pretending breasts". And I said "nanna had something called breast cancer and they had to take them both off and nanna's here to play with you". And he says "can I see them"? And I said "yes you may", and I showed him. And a couple of weeks later he asked again, that's it. Until a friend of mine*

*died and he wanted to know what sort of cancer that person died of, and it was stomach cancer. "**Oh good, my nannas still alive**".*
*That was interesting, that he, wanted to see a prosthesis, and **he's happy as anything as long as nanna is here to play with him**.*

P4: *I think go with **whatever makes you feel comfortable**. Whatever **gives you peace of mind**.*

P4: *It depends. Likewise to have the mastectomy, or just a lumpectomy, or just the single mastectomy, if **you feel that's what you'd like** to have done go with that **if that's what's going to give you peace of mind** because that can, that allows you to **feel comfortable in your body**, whatever it is, yeah go with what, it's very individual as X said.*

P2: *I had had, once again no family history of breast cancer but fibrous cystic breast changes so **over the years** have had **multiple breast cyst aspirations** and **biopsies** and so **during** all those previous tests **formulated**, oh my gosh **if its cancer** I'm just going to **lop both of them off** and **get on with my life** you know.*

Timing – Instant/Delay

Time was also noted as a potentially important factor in decision-making, either in decisions actually made, or in terms of the advice that participants would give to other women. For one participant in the unilateral mastectomy group, having extra time after surgery allowed her to reflect on her decision and to think about the pain and decide against the CPM. Among participants in the CPM group, it was discussed that the timeframes for decision-making are very compressed, and it is a big decision. Participants felt that perhaps women should give themselves more time to make these decisions.

P10: *Well **by that time I'd had time to think** because from diagnosis to operation was only a couple of days **so I had time to think, and as you said the pain**. I'm thinking I don't want to go through it, it's perfectly healthy, I might never ever develop cancer in that side.*

P2: *My **husband is an engineer**, he literally **had a spreadsheet** from my diagnosis **with all the variables to handle that part of the process**. Do I do chemo, do I do radiation? Thankfully he just got back to me. I just **don't know how** your **average Joe** blow would **even begin in the two weeks they give you to decide that**.*

P2: *I think certainly you're, for me my initial diagnosis, you **just want to be rid of that, you just want to be rid of the cancer** so I think that it **isn't a great time to be making**, probably these really **big decisions** about one or the other*

because um you know ***you're not thinking about all the other complex things****. You're just thinking* ***I just want this out of me of me****, whatever, so I think the* ***compressed time period*** *to make a decision is retrospectively was the* ***most difficult thing****. Ok, they say best practice is they want this out in two weeks so the surgeons say and that's* ***not a lot of time****.*

P2: *..just said to X I'd now say to someone, you know maybe don't,* ***wait and maybe have a tissue sparing, nipple sparing reconstruction*** *because it is so much more complex to have a reconstruction. Yeah, even in two years time my viewpoint has changed.* ***Let's attend to this, the side you know it is on****, once again it is very personal. I'm* ***not regretful because we did find cancer on this side****, but there are lots* ***more options now****, you* ***don't have to make your decision in two weeks****.*

P2: *I agree with you, and my message would be* ***don't rush with your decision****. Take your time. You* ***don't need to make your big decision in two weeks****.*

P4: ***Do women consider*** *that when they consider, as part of their decision-making process,* ***what's available now*** *to me what are my options now* ***as opposed to if I wait a bit and have a single mastectomy or lesser treatment now, what would that mean for the future time*****?**

Genetic

Of the women who disclosed they underwent genetic testing and were positive, both chose to have unilateral mastectomy. Those in the CPM group who referred to genetic testing did not disclose being BRCA positive (or disclosed being BRCA negative), or stated that they were likely to have a genetic pre-disposition but one that could not yet be tested. There was a sense that genetics was perhaps not yet informative in these women.

P11: *Mind you* ***I don't think I was tested then****, no-one knew that I was…,* ***although it was quite obvious*** *as* ***all the women in the family all get breast cancer*** *and it comes back in the bone and that said, um yeah.*

P10: *I just look at the fact that when I had it the first time round* ***they discovered I had the gene*** *after I had my operation, then the* ***surgeons were recommending I had the other side off****.*

P1: *I had* ***lost my mother to breast cancer*** *which* ***played a part in my decision to have prophylactic*** *mastectomy.*
Although we definitely have a cancer gene *in the family, we just* ***don't know which one*** *yet.*

P4: *I had highly aggressive, highly invasive cancer, and I'm triple negative, I was also* ***screened genetically for BRCA*** *which was* ***negative*** *with a view with*

*how, or what to **do with my ovaries** because apparently women with my type of cancer are more likely to develop ovarian cancer.*

*P9: I **didn't need** to have a **genetic test**, that there could be a genetic link if I did have family history but **there isn't any test that would pick it up yet**.*

Cancer Differentiation

For participants in the CPM group, there was a thread in the discussion around breast cancer not all being the same, and that there are different forms that require different means of follow-up and potentially treatment. Moreover, some women were concerned that the language used around some types of cancers – namely DCIS – underplayed its seriousness and might leave some women with the mistaken impression that it is either not as serious a condition, or that they had been cured. Not all women had disclosed their type of cancer so this is not reported in the analysis.

*P5: With DCIS, I think **doctors should be careful not to** say DCIS. They **say "Oh it's just DCIS"***

*P3: I was told **lobular cancer can come back** and it **doesn't show in mammogram or ultrasounds**, or I know mine didn't.*

P6: But I think that's a problem with DCIS too that they don't screen for genetics.

*P5: Did anyone ever think or **realise there are so many different sorts of breast cancer**? I was astounded because the specialist says we have to wait to get the results back to **find out if you are HER2 positive or HER2 negative** whether you are this or that, **and if you're this you have to have this done, and if you're this you have that done**.*

*P2: And **how that affects your treatment and decision options**.*

*P5: Then you get your family and friends saying why are you having that done because so and so's mother had it and they had blah, but **you're saying they're all different**.*

*P9: But **maybe with certain types of cancers they need different diagnostic tools to pick it up**. Yeah I don't know. Core biopsy is maybe better for lobular and MRI.*

Appendix 11: Chapter 6 - Development of Attributes – Mapping from Qualitative Results

The second phase of the research investigating women's preferences for how to manage the ongoing risk of breast cancer recurrence is a quantitative discrete choice experiment (DCE). This builds on, and is informed by, the qualitative work with women with early stage breast cancer to value the factors that are important to women when deciding how to manage their ongoing risk of breast cancer recurrence. The central question of interest being explored in this DCE is the value placed by women on different methods of managing the ongoing risk of cancer recurrence, with an emphasis on trying to identify the role (and value) of the fear of recurrence in that decision-making process.

Broadly, the process that has been adopted to develop the DCE is as follows:

1. Domains were drawn from the focus group discussions held with 11 women who have undergone a therapeutic mastectomy (TM), nine of whom also underwent a contralateral prophylactic mastectomy (CPM) for breast cancer.
2. The process used to develop the draft attributes for the DCE was as follows:
    a. Each of the domains arising out of the qualitative interviews was deconstructed into those aspects that can be described as pertaining to a process aspect of health care delivery (and potentially amenable to change via policy) and those which reflect a response to that health care (e.g. feelings of anxiety/femininity/sexuality).
    b. For each of the domain elements pertaining to a health care process, a possible attribute describing an aspect of managing the risks of breast cancer recurrence was suggested.
    c. The suggested attributes were rated as being of high, medium or low importance in terms of what is likely to most influence the choice of risk management option. The ratings were based on the feedback from the qualitative interviews.
    d. Levels for the derived attributes were determined based on the published literature, clinical input and the results from the qualitative interviews.
    e. Two choice questions were formulated for potential inclusion in the DCE: one to test the idea associated with directing participants to focus on the concept of interest (in this case a reduction in fear of recurrence, or peace of mind), another using more standard wording.

Mapping of Domains

The development of draft attributes from the domains identified in the qualitative interviews is provided in Table A 27.

Table A 27: Mapping of qualitative domains to suggested DCE attributes

Domain from qualitative work	Categorisation (Process or Response).	Aspect potentially captured as an Attribute.	Relevance of possible attribute to primary DCE question.
Fear & Risk			
Cancer recurrence	Process & Response	Recurrence rate associated with the proposed method of managing risk.	High. This is important as it relates to the "efficacy" and recurrence aspects being tested.
Anxiety of testing (anticipatory anxiety)	Response	Frequency of ongoing testing. Type of ongoing testing.	High. These are critical aspects of describing the ongoing monitoring.
Peace of mind	Response	Not suitable as an attribute; could be used to form underlying choice question.	
Relief	Response	Not suitable as an attribute; could be used to form underlying choice question.	
Learning from the Past			
Missed & undetected cancers	Process	False negatives/false positives of testing. Potentially as part of background scenario only.	Medium. Will it overly complicate the DCE to have recurrence and accuracy?
Surgical & radiotherapy pain/complications	Process Response	Side effects associated with TM and radiotherapy interventions. Potentially as part of background scenario only since these complications relate to primary therapy.	Low. Describe as part of the background scenario only. Response to pain etc cannot be affected by policy.
Costs			
Monitoring	Process	Costs of monitoring.	High.
Surgery/ reconstruction	Process	Costs of surgery/ reconstruction; as a separate attribute.	High.
Access to services	Process	Need to travel for care. Need to see another team for care. Time to wait for care.	Low. Raised in limited number of cases by women in the qualitative work.
Loss of income	Process	Time to return to usual duties.	Low. There is more than just the physical at play here, so it is difficult to state when women would be able to return to usual duties.
Unwanted Effects			
Pain	Process	Side effects of intervention.	Medium. Potentially incorporate with other side effects.
Numbness	Process	Side effects of intervention.	Medium. Potentially incorporate with other side effects.
Loss of balance	Process	Side effects of intervention.	Medium. Potentially incorporate with other side effects.
Body Image			
Freedom from bras	Response	Not suitable as an attribute;	

Domain from qualitative work	Categorisation (Process or Response).	Aspect potentially captured as an Attribute.	Relevance of possible attribute to primary DCE question.
		potentially include as an attitude question within the survey?	
Relationship to breasts	Response	Not suitable as an attribute; potentially include as an attitude question within the survey?	
Aesthetics	Process & Response	Potentially captured via prosthetics (see below).	Medium.
Breast symmetry	Process	Could be captured using an attribute for prosthetics.	Medium.
Coping	Response	Not suitable as an attribute; could form part of the choice question.	
Sensuality	Response	Not suitable as an attribute.	
Timing	Response & Process	Not suitable as an attribute.	
Life stage (age, children, family)	Response	Collect demographic information. Not suitable as an attribute.	
Medical involvement (communication)	Process	Vary degree of involvement in ongoing risk management.	Medium.
Family communication (anxiety, guilt)	Response	Not suitable for an attribute. Collect demographic information.	

Abbreviation: DCE, discrete choice experiment.

Timing, and its impact on decision-making, was an important factor raised by women in the focus groups informing the development of the DCE. However, an attribute for timing has not been included due to the difficulty of capturing all the aspects of timing in the one attribute in a meaningful manner. That is, there is a difference between a woman choosing to delay her decision about undergoing a CPM, and deciding to undergo the CPM but needing to delay the surgery due to medical reasons (comorbidities and the enormity of the surgery, particularly if it includes reconstruction). The former is not amenable to inclusion as an attribute in a DCE ('you decide to delay your decision' is prescriptive and describes an action the woman would take rather than an attribute of the approach to risk management). Similarly, the latter relies on a medical assessment about the woman's ability to tolerate surgery and therefore is not specifically an attribute of the approach to risk management. Moreover, it is difficult to conceptualise how long a delay might be and what happens to women (in terms of monitoring) during such a delay. For these reasons, the impact of timing on the choice undergo a CPM is not tested via the DCE.

Appendix 12: Chapter 6 - DCE Experimental Design and Survey Generation

Overview

The process for the generation of the design for the discrete choice experiment was as follows:

1. The design was generated in Ngene.[229]
2. Given the focus of this research on investigating the relative value between meta-health effects and health effects, the final design was generated to maximise WTP efficiency (see below).
3. Priors for the final survey design were obtained from the pilot survey. The pilot survey design did not use prior information. Accordingly, it was based on maximising mnl efficiency (since WTP efficiency requires the use of priors).
4. Effects coding was used for all categorical variables.
5. Ngene generated designs were manually checked for scenario dominance and the frequency of appearance of attribute level pairings, and altered if required.
6. The efficiency of subsequently designs was re-tested by importing the altered design into Ngene for evaluation of the efficiency levels.
7. Comparisons of efficiency are provided using d-error, s-estimate (an estimate of the required sample size needed to test the WTP at the 5% significance level), and WTP efficiency.[229,346]

General information on generating designs in Ngene is provided in Chapter 4 and Appendix 7. Both designs were generated using a modified Federov algorithm to apply restrictions to the attribute levels across the choice options. The first was that *CBC Risk* on the CPM only option could not exceed that of routine monitoring. The second was that *Pain Risk, Sensitivity Risk* and *Surgical OOP* were not to appear in routine monitoring. The levels for *Pain Risk* and *Sensitivity Risk* do not include zero, so for design purposes two alternative methods of generating the design were tested. For the pilot design, the attribute level combinations for *Pain Risk* and *Sensitivity Risk* were generated by specifying five levels for those attributes, including zero. For the CPM option, all levels could apply except zero, while for the monitoring option only zero could apply. For the final survey version, only the four actual levels were included in the design, restricting that appearing in the monitoring only option to the lowest level only so as to preclude the inclusion of a fifth level in the attribute design. The resulting fixed level for both attributes in the monitoring only option was then manually changed to zero and the impact on the efficiency measures evaluated.

A total of 48 rows were used, in four blocks (incorporated into the design), specifying for Ngene to minimise the chance of dominant alternatives across choice sets.

Generating the Design: The Pilot Survey

Specific to the research in this chapter, the pilot survey was generated using zero priors, and specifying an mnl efficient design. This produced a d-efficiency score of 0.002. Within this design, there was one scenario (Scenario 19, block 2) in which the CPM only option was likely to be perceived as dominated by the Routine Monitoring option in the No-Efficacy Difference framing set. This was not corrected in the pilot survey.

An attempt to evaluate the efficiency of the No-Efficacy Difference framing set (in which the attribute for *Other BC Risk* is essentially removed from the design) did not produce a d-efficiency score within Ngene. This is potentially due to the presence of that dominated scenario.

Generating the Design: The Final Survey

A WTP efficient design was generated using the attributes and levels described in the main text, and as replicated in Table A 9. Priors for the design of the final survey were sourced from the results of the conditional logit regression analysis of the pilot survey. These are provided in Table A 9, alongside the attribute level coding.

The procedure for testing the efficiency parameters of the designs was as described in Appendix 2 of Chapter 4. Due to the inclusion of restrictions on a number of attribute levels (and combinations – list from the Ngene design), this initial design was not balanced in levels, and failed to produce comparisons of all relevant levels. This resulted in some changes to attribute level combinations. Subsequently modified designs were re-evaluated in Ngene to assess the impact of any modifications to attribute levels on the design efficiency.

Table A 28: Design Parameters and Priors

Attribute		Levels	Coding	Priors
CBC Risk	The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:	3 in a 1,000 (0.3%) 10 in a 1,000 (1%) 20 in a 1,000 (2%) 50 in a 1,000 (5%)	Continuous	-0.089
Other BC Risk	The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:	100 in a 1,000 (10%) 150 in a 1,000 (15%) 200 in a 1,000 (20%)	Continuous	-0.041
Monitoring	In addition to your regular self-checks, you will need to have the following tests:	Mammogram MRIs of your breast area Ultrasounds of your breast area	0(base) 1 2	 0.006 -0.006

Attribute		Levels	Coding	Priors
Frequency	You are scheduled to have your follow-up tests every:	Six months Year Second year	0(base) 1 2	 0.029 -0.047
Monitoring OOP	The cost for monitoring each year is $900, and you pay:	0, 300, 600, 900	Continuous	-0.025
Surgery OOP	The cost for surgery associated with managing your ongoing risk of cancer recurrence is $15,000, and you pay:	0, 5000, 10000, 15000	Continuous	-0.003
Pain Risk	The chance you will experience ongoing pain is:	400 in 1,000 (40%) 300 in 1,000 (30%) 200 in 1,000 (20%) 100 in 1,000 (10%)	Continuous	-0.001
Sensitivity Risk	The chance you will experience an ongoing loss of sensitivity in your breast area is:	600 in 1,000 (60%) 500 in 1,000 (50%) 400 in 1,000 (40%) 300 in 1,000 (30%)	Continuous	0.003
Involved	Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence.	Not very often Always	0 (base) 1	0.066
ASC		Monitoring: CPM		0.512

Note: Within Ngene, effects coding is activated by specifying "effects" prior to attribute coefficient. Level codes within the attribute name are entered as would be any categorical variable e.g. 0, 1, 2, with the order specifying the base level (last level taken as the base).

Abbreviation: n.a., not applicable; Govt, government; OOP, out-of-pocket.

The resulting efficiency measures were WTP=879.16; s=9390.27, and d-efficiency scores of d=0.009. Note that the large s estimate in this design is primarily driven by the small coefficient estimates for mode (0.006) and pain (-0.001), which when coupled with their attribute levels would imply an overly large sample is required to detect a significant difference of that size. In addition, the implied preference inversion for *Monitoring* and *Frequency* observed in the pilot survey would require a large sample to confirm.

The restrictions specified in the design resulted in situations where there would be a dominated scenario in the No-Efficacy Difference framing set (Scenarios 17 and 26). To account for this, for Scenario 17, the levels on *Frequency* were shifted from biennial (2) to biannual (O) for Routine Monitoring, and the level on type of monitoring from MRI (1) to ultrasound (2) on CPM. For Scenario 26, the levels on *Monitooring OOP* of $900 and $300 for CPM and Routine Monitoring initially produced by Ngene were swapped. In addition, the levels for *Pain Risk* and *Sensitivity Risk* for the monitoring only option were set to zero for all scenarios as described above.

The efficiency values for the dominance altered and zero corrected design were: WTP= 889.61; s= 9213.06; and d=0.009. Similarly, the no-efficacy difference version of the design (in which *Other BC Risk* does not enter the design) was imported into Ngene for evaluation of efficiency. This resulted in WTP efficiency and s-estimates that were largely in-line with those that apply to the base design: WTP= 871.38; s= 9504.53; d=0.009.

Appendix 13: Chapter 6 - Results of the Pilot Survey

Preferences for managing ongoing breast cancer recurrence risks: Pilot DCE Findings and Future Questions.

Meeting Notes and Action Items

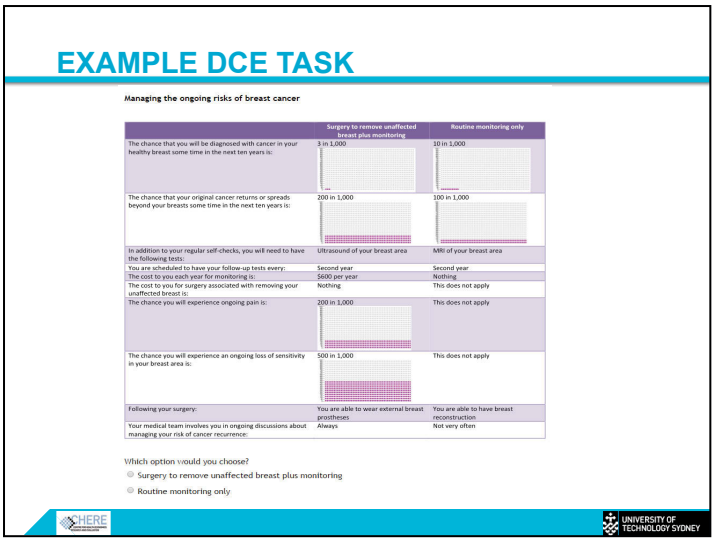
Attendees: Marion Haas (MH), Jane Hall (JH), Patsy Kenny (PK), Richard De Abreu Lourenco (RL), Kim Parish (KP), Rosalie Viney (RV)
Apologies: Domini Stuart (DS)
Date: 22nd September 2015
Venue: CHERE, UTS

These notes are not direct minutes of the meeting. Rather, they highlight relevant discussion points and action items arising during course of the meeting.

For ease of presentation, and where necessary, the meeting notes are presented alongside the presentation slide/information that engendered that discussion.

RL provided brief introductions for KP to the group, apologies from DS, and provided a brief overview of the purpose of the meeting.

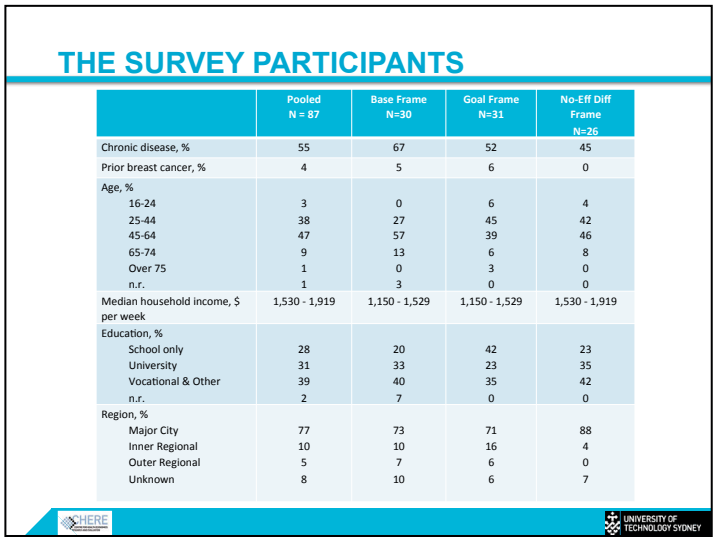
DCE Overview



- It was noted that the DCE scenarios are more legible when they appear online, rather than as "screenshots" in a presentation.
- RL commented that during the testing phase of the survey, the technical support person at Pureprofile (the online survey panel) thought the survey was not advancing because successive scenarios looked similar in terms of colour and layout. RV felt this was worth noting for the future. It was also noted that this called into question what was actually being testing by Pureprofile if they were not reading each scenario.

- RL provided background on the various versions of the survey being used to test the framing effects (the Base version, Context and No-Efficacy difference) of providing different types and amounts of information.
- KP asked for background on how women are selected to participant in online surveys. RL advised that Pureprofile has around 64,000 active members per month. Each member has their own log-on page (akin to a Facebook page) that allows them to see active surveys. When they log-in, they can view which surveys are active and decide in which they'd like to participate.
- It is possible to target surveys to specific groups. In this case, the survey only appeared on the log-in page of adult women. There were no other restrictions.

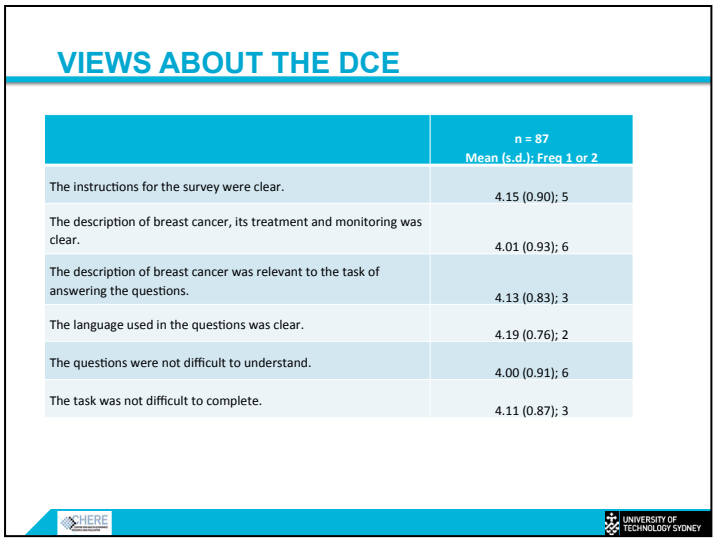
Participant Demographics



	Pooled N = 87	Base Frame N=30	Goal Frame N=31	No-Eff Diff Frame N=26
Chronic disease, %	55	67	52	45
Prior breast cancer, %	4	5	6	0
Age, %				
16-24	3	0	6	4
25-44	38	27	45	42
45-64	47	57	39	46
65-74	9	13	6	8
Over 75	1	0	3	0
n.r.	1	3	0	0
Median household income, $ per week	1,530 - 1,919	1,150 - 1,529	1,150 - 1,529	1,530 - 1,919
Education, %				
School only	28	20	42	23
University	31	33	23	35
Vocational & Other	39	40	35	42
n.r.	2	7	0	0
Region, %				
Major City	77	73	71	88
Inner Regional	10	10	16	4
Outer Regional	5	7	6	0
Unknown	8	10	6	7

- A total of 87 women completed all 12 choice sets in their survey. Of those, only 85 provided demographic information.
- It was noted that there were some minor differences across the women included in three frames with regard to their demographics. These differences were unlikely to affect the analysis of the choices women made. Moreover, due to the sample size, analyses by demographically defined subgroups (e.g. age group) were not possible due to the likelihood that the choices from such sub-groups would not cover all 48 choice sets from the design.
- There were only two women of the 87 who reported experience of breast cancer. However, the question on prior breast cancer experience was only asked of those women who reported another chronic health condition. KP confirmed (after the meeting) that this is likely to have missed some women with a prior experience of breast cancer who do not have other chronic health issues. This will be revised for the final survey.

Survey Language Clarity and Relevance



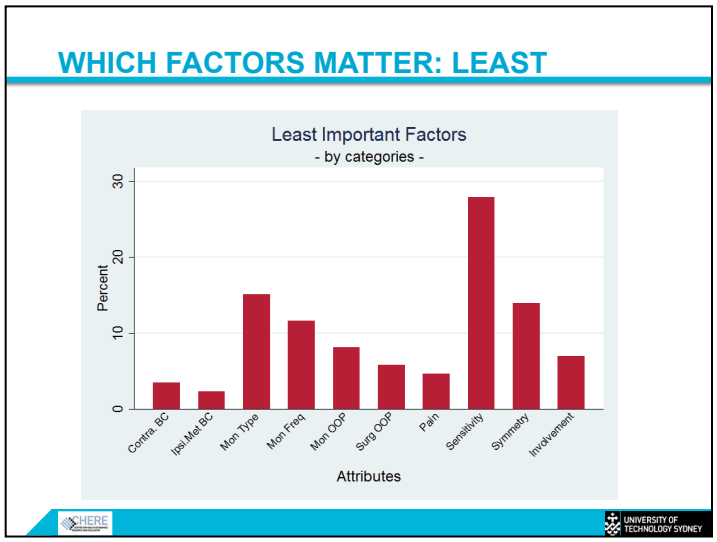
- The results with respect to the survey language and clarity were reviewed and it was agreed that the results were very positive. Given that all the mean values were over four (indicating a high degree of agreement with the statements on survey clarity/understanding and relevance), these results show that in general women did not have difficulty with the understanding or completing the survey.
- RL noted that one woman had commented that the description of breast cancer had been so clear she could really imagine having breast cancer. Post-script: in the meeting, RL incorrectly described the woman's feedback as describing this as "scary" – the actual comment from the woman was as follows: *"Almost too clear, so much so I could almost imagine what it would be like to go through this"*.
- This generated discussion around how participants might react to such descriptions, and the potential that it might result in some women changing their screening behaviour.

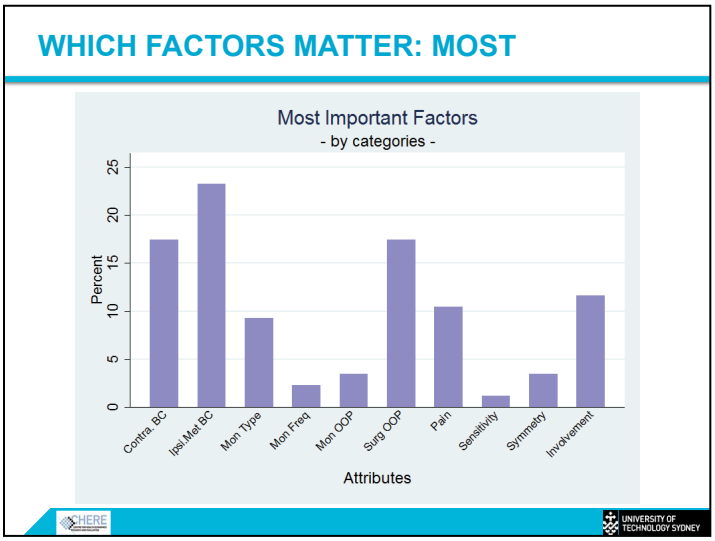
- RL apologised for having excluded the results to the question regarding whether there were too many attributes in the survey, and whether it affected the manner in which women answered the survey. The responses to the clarity questions suggest this was not of concern.
    - *Post-script: The results for how women rated the number of attributes that appeared are below. Overall, 41% of women felt there were too many attributes, but 28% reported that they dealt with this by only focusing on those which they felt were important. This is lower than the proportion of women who reported that they had a strategy when answering the DCE of only focusing on those factors that mattered to them (37%). In all, 58% reported that the number of factors did not affect how they answered, and the only open text comment supported the number of factors provided:*

Options	***%***
The number of factors did not affect how I answered	*58*
There were too many factors and it made it difficult to	*13*

answer the questions	
There were too many factors so I only focused on the ones I thought were important	*28*
Other (open text): *"I found it important to have as many factors available"*	*1*

Factor Importance



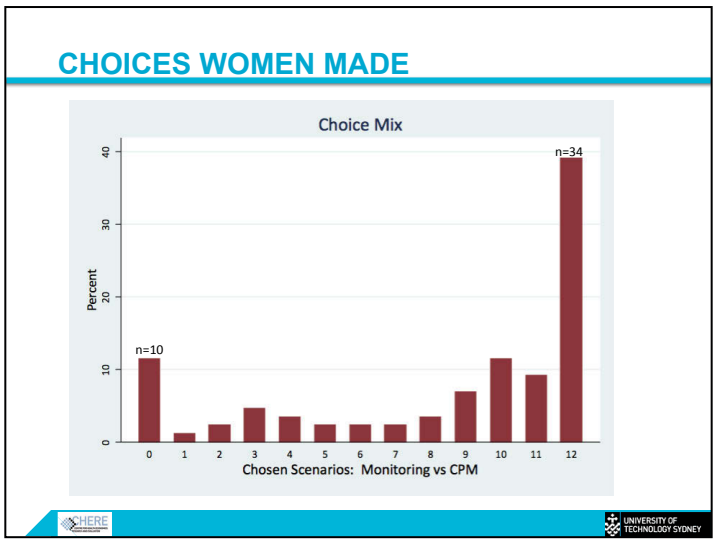


- In general everyone was comfortable with the response to the strategies used by women to answer the survey; 37% of women reported that they focused only on the factors they thought were important (*Post-script: this accords with the proportion who felt there were too many* factors – 41%). Among those women, the order in which factors were considered was also reported. It was agreed that this was not particularly surprising, but it was noted that the factors addressing prostheses/reconstruction and breast sensitivity were last.

- The factors that all women (regardless of whether or not they applied a strategy when they were making their choices) considered to be most important (risk of cancer recurrence) and least important (breast sensitivity) were subsequently discussed.

RV commented that these questions were perhaps more informative than those regarding the ordering of factors women used when answering the choice questions. She suggested leaving the latter out of the main survey.

Choosing Between Monitoring and CPM



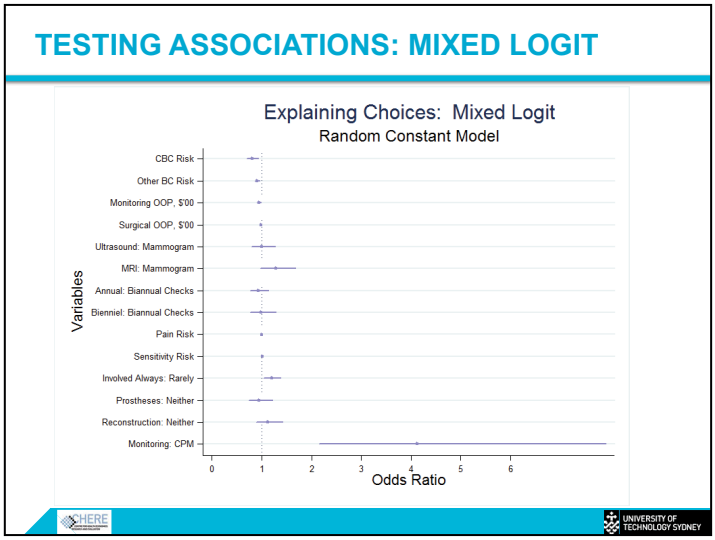
- Each woman in the survey answered 12 choice questions (across 87 women this provided 1,044 choice observations). Among the 87 women 44 always chose one option over the other; 34 always chose the monitoring only option, and 10 always chose the surgery only option.
- These results show that women hold strong preferences over how they might deal with the decision of whether or not to have a CPM. There was agreement that while these women represent "non-traders" (they have not altered their choice based on the variability of the factors) they should be retained in the analysis.

- There were a number of suggestions about understanding what might be driving the preferences of these women:
    - Conduct a multi-nomial logit (MNL) analysis in which women are classified as 'CPM choosers' (always choose CPM), 'monitoring choosers' (always choose monitoring), and 'traders' (switch between them) using key demographics (and potentially frame) as regressors.

    - These strong non-trading preferences would appear to reflect attitudes to risk. It would be informative therefore to collect more information in the final survey on women's attitudes to risk and health seeking behaviours e.g.:
        - Whether or not women attend for regular cervical screening?
        - Have you ever had a mammogram/cervical screen? How often?
        - Tobacco and alcohol consumption behaviour?

    - It was also noted that as well as own breast cancer experience, familial and friends' experiences with breast cancer could influence women's preferences with regard to subsequent treatment. Additional questions in the final survey might also include (to test for halo effects):
        - Have you had a family member with breast cancer?
        - Have you had any close friends with breast cancer?

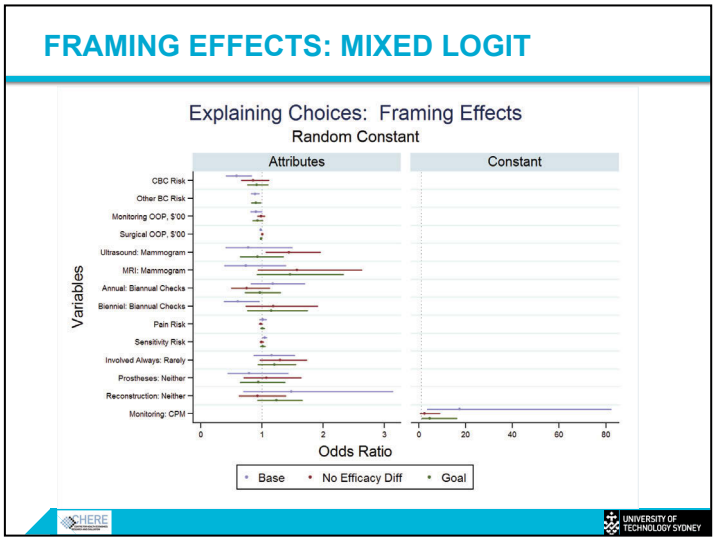
Understanding Choices



- RL took the meeting through the various regression methods for exploring what was driving women's choices; starting with the conditional logit analysis that recognised repeated choices were being made by the same individual but not that they varied in their responses, to the mixed logit that took account of the between women differences in how they responded.

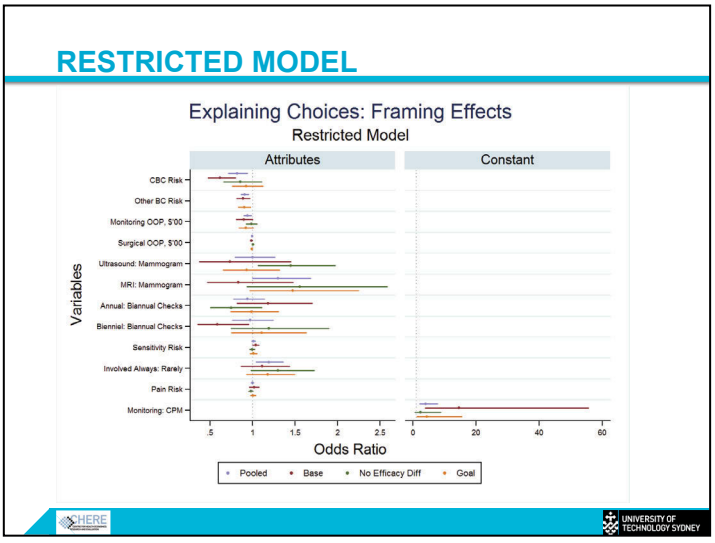
- In short, it was clear from the results that women held strong preferences between CPM and monitoring, as witnessed from the value for the constant. The high value for the *Monitoring: CPM* odds ratio indicates that women have a higher odds (nearly double) of choosing monitoring compared with CPM, regardless of the other factors. This is to some extent expected given the pattern of choices reviewed previously.
    - RL noted that of the remaining factors, cancer recurrence was significant, as was surgical cost, for both analyses. In the mixed logit, monitoring costs and the extent of involvement were also significant. There was some discussion around whether the form of the surgical cost variable should be modified since the coefficient was still so small; despite it being expressed in hundreds. RV suggested making it a squared term. Alternative forms (including log-cost) will be tested in the analysis of the final survey.

Framing Effects



- There was discussion of whether or not framing effects across the Base, Context and No-Efficacy difference versions, were evident. From the results of the mixed logit this appeared to be the case, with the Base version showing the largest difference in mean coefficients from the other two versions. It was agreed that framing in terms of excluding a difference in efficacy, and including more information about the history leading up to the initial diagnosis of BC in the background vignette was resulting in a change in coefficient values, and in what was significantly affecting choices.

Dropping Prostheses/Reconstruction



- There were two factors that did not appear as significant in any of the frames: pain, and the ability to wear prostheses or have reconstruction. However, pain was ranked the fourth most important factor by women, and there was a significant difference between women in the variance around how they considered this factor. It was agreed that this should therefore be retained.
- There were no significant effects for the prostheses/reconstruction factor (included to capture the effects on breast symmetry). KP noted that there had been some difficulty in the wording used to describe this attribute, and she felt that perhaps it was still not conveying the right concepts. She noted that the question of having a reconstruction or not is

one that is often not addressed until some time after women have had their mastectomies. PK commented that some women in the focus groups certainly talked about the issue of symmetry for them as a factor in their decision-making. However, it was noted that even for some of those women there was a gap between when women had their therapeutic mastectomy, subsequent contralateral mastectomy and reconstruction.

- The other potential complicating factor with this attribute is that it refers only to the decision about the CPM. Yet the background vignette tells women they would have had at least a therapeutic mastectomy. It is possible that women were making inferences about how the reconstruction (or otherwise) following their therapeutic mastectomy was being addressed, and that this was affecting their views on this factor. This made it conditional on what was happening with the reconstruction (or otherwise) following the therapeutic mastectomy. However, it was agreed that it could potentially bias women's preferences if we included information in the vignette about what occurred with the reconstruction (or otherwise) following the therapeutic mastectomy.

- RL also presented results in which the prostheses/reconstruction attribute was excluded from the mixed logit analysis. These showed that excluding that factor could be done without loss of explanatory power. JH asked whether this would lead to concerns of a loss of face validity; would the experiment be criticised for excluding a factor that should be there? KP felt that given the issues with the timing of the decision for reconstruction, this factor could be excluded without a loss of face validity.

- It was discussed that perhaps the most appropriate manner in which the issue of prostheses/reconstruction/breast symmetry could be addressed would be to include in the vignette the fact that the patient would at some point consider the question of reconstruction.

.

Appendix 14: Chapter 6 - Supporting Evidence for the Main Survey

Table A 29: Conditional Logit Analysis

	Pooled	Base	No-Eff Diff	Goal
CBC Risk	-0.081	-0.062	-0.139	-0.028
	(0.018)**	(0.029)*	(0.032)**	(0.032)
Other BC Risk	-0.020	-0.023		-0.013
	(0.008)*	(0.013)		(0.013)
Monitoring OOP, $'00	-0.022	-0.031	-0.020	-0.015
	(0.006)**	(0.010)**	(0.010)*	(0.012)
Surgery OOP, $'00	-0.003	-0.004	-0.003	-0.003
	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Pain Risk	-0.002	-0.000	0.002	-0.007
	(0.003)	(0.005)	(0.004)	(0.005)
Sensitivity Risk	-0.000	-0.002	0.001	0.001
	(0.003)	(0.006)	(0.005)	(0.006)
Ultrasound: *Mammogram*	0.068	0.051	0.078	0.075
	(0.025)**	(0.038)	(0.041)	(0.050)
MRI: *Mammogram*	-0.023	0.038	-0.040	-0.066
	(0.025)	(0.047)	(0.042)	(0.044)
Annual: *Biannual Checks*	-0.025	-0.024	0.035	-0.098
	(0.027)	(0.050)	(0.044)	(0.049)*
Biennial: *Biannual Checks*	0.033	0.014	-0.028	0.127
	(0.030)	(0.046)	(0.053)	(0.055)*
Involved Always: *Rarely*	0.092	0.057	0.123	0.093
	(0.031)**	(0.057)	(0.051)*	(0.056)
Monitoring: *CPM* Constant	0.516	0.505	0.591	0.444
	(0.099)**	(0.171)**	(0.174)**	(0.169)**
N	11,136	3,792	3,864	3,480
pseudo R-squared	0.21	0.23	0.19	0.23
Wald Chi	222.17	80.00	93.34	73.52
d.f.	12	12	11	12
p-value	0.00	0.00	0.00	0.00
Log-likelihood	-3,042.75	-1,015.87	-1,085.51	-931.59

Notes: * $p<0.05$; ** $p<0.01$
All categorical data are effects coded to minimise the potential for contamination of the base effect in the grand mean.
Base categories are shown in italics.
All analyses were conducted in STATA 12, using robust standard errors (in parentheses) to account for the survey nature of the data.
Analyses conducted with 1,000 replications.

Abbreviations: BC denotes breast cancer; CBC contralateral breast cancer; CPM contralateral prophylactic mastectomy; d.f. degrees of freedom; OOP out-of-pocket.

Table A 30: Alternative Specific Constant Logit Analysis – Incorporating Demographics: Pooled Dataset

	Model 1	Model 2	Model 3	Model 4
CBC Risk	-0.081	-0.083	-0.079	-0.081
	(0.018)**	(0.018)**	(0.019)**	(0.019)**
Other BC Risk	-0.020	-0.019	-0.022	-0.021
	(0.008)*	(0.009)*	(0.009)*	(0.009)*

	Model 1	Model 2	Model 3	Model 4
Monitoring OOP, $'00	-0.022	-0.023	-0.023	-0.024
	(0.006)**	(0.006)**	(0.006)**	(0.007)**
Surgery OOP, $'00	-0.003	-0.003	-0.004	-0.004
	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Pain Risk	-0.002	-0.002	-0.002	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)
Sensitivity Risk	-0.000	-0.000	-0.001	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)
Ultrasound: *Mammogram*	0.068	0.072	0.066	0.069
	(0.025)**	(0.025)**	(0.026)*	(0.026)**
MRI: *Mammogram*	-0.023	-0.025	-0.032	-0.034
	(0.025)	(0.026)	(0.026)	(0.027)
Annual: *Biannual Checks*	-0.025	-0.028	-0.026	-0.028
	(0.027)	(0.028)	(0.029)	(0.029)
Biennial: *Biannual Checks*	0.033	0.035	0.025	0.027
	(0.030)	(0.030)	(0.031)	(0.031)
Involved Always: *Rarely*	0.092	0.094	0.089	0.093
	(0.031)**	(0.032)**	(0.033)**	(0.033)**
Monitoring: *CPM* Constant	1.033	1.295	1.062	1.313
	(0.197)**	(0.213)**	(0.293)**	(0.298)**
Demographics and Influencers				
Cancer Concern: *Not Concerned*		-0.517		-0.508
		(0.116)**		(0.122)**
Income $79K: *$39K*			-0.201	-0.215
			(0.164)	(0.166)
Income $149K: *$39K*			0.215	0.235
			(0.174)	(0.175)
Income Over $150K: *$39K*			-0.052	-0.026
			(0.240)	(0.248)
Age 16: *25*			-0.658	-0.690
			(0.370)	(0.365)
Age 45: *25*			-0.171	-0.161
			(0.217)	(0.217)
Age 65: *25*			0.054	0.092
			(0.252)	(0.250)
Age 75: *25*			1.098	1.055
			(0.627)	(0.619)
Outer Regional: *Major City*			-0.066	-0.060
			(0.240)	(0.236)
Inner Regional: *Major City*			-0.026	-0.061
			(0.185)	(0.183)
Ed Uni: *School*			-0.061	-0.037
			(0.127)	(0.127)
Vocational: *School*			0.158	0.108
			(0.128)	(0.128)
Observations	11,136	11,136	10,344	10,344
Individuals	464	464	431	431
Wald Chi	94.73	113.33	99.50	118.09
d.f.	11	12	22	23
p-value	<0.001	<0.001	<0.001	<0.001
Log-likelihood	-3,042.75	-2,954.72	-2,795.07	-2,717.84
Pseudo R^2	0.013	0.042	0.094	0.119

Notes: * $p<0.05$; ** $p<0.01$

Model 1 contains DCE attributes only; Model 2 adds the categorical variable for whether or not women expressed concern over cancer recurrence when answering the choice scenarios; Model 3 is Model 1 with the addition of key demographics using a complete case analysis; and Model 4 combines Model 2 with Model 3.
All categorical data are effects coded to minimise the potential for contamination of the base effect in the grand mean.
Base categories are shown in italics.
All analyses were conducted in STATA 12, with clustering on ID to account for within respondent variability (standard errors in parentheses) and to account for the survey nature of the data.
Analyses conducted with 1,000 replications.

Abbreviations: BC denotes breast cancer; CBC contralateral breast cancer; CPM contralateral prophylactic mastectomy; d.f. degrees of freedom; Ed education; OOP out-of-pocket; Uni university.

Here the alternative specific constant logit regression modifies the standard conditional logit regression (equation 10) as follows:

$$U_{ijt} = \beta x_{ijt} + (\zeta A_i)' + \varepsilon_{ijt}$$

where the A defines a vector of individual (i) specific demographic variables that enter as a matrix, and do not vary over the alternatives (j) or choice (t).[213,235]

Table A 31: Alternative Specific Constant Logit Analysis – Incorporating Block Specific Dummies

	Pooled	Goal
CBC Risk	-0.081	-0.024
	(0.019)**	(0.033)
Other BC Risk	-0.021	-0.015
	(0.009)*	(0.014)
Monitoring OOP, $'00	-0.024	-0.019
	(0.007)**	(0.013)
Surgery OOP, $'00	-0.004	-0.004
	(0.001)**	(0.001)**
Pain Risk	-0.002	-0.007
	(0.003)	(0.004)
Sensitivity Risk	-0.001	-0.000
	(0.003)	(0.006)
Ultrasound: *Mammogram*	0.069	0.084
	(0.026)**	(0.053)
MRI: *Mammogram*	-0.034	-0.070
	(0.027)	(0.045)
Annual: *Biannual Checks*	-0.028	-0.105
	(0.029)	(0.048)*
Biennial: *Biannual Checks*	0.026	0.108
	(0.031)	(0.056)
Involved Always: *Rarely*	0.093	0.087
	(0.033)**	(0.057)
Monitoring: *CPM* Constant	1.285	0.909
	(0.333)**	(2.793)
Block	0.003	0.022
	(0.014)	(0.149)

	Pooled	Goal
Demographics and Influencers		
Cancer Concern: *Not Concerned*	-0.509	-0.348
	(0.122)**	(0.218)
Income $79K: *$39K*	-0.214	0.141
	(0.165)	(0.315)
Income $149K: *$39K*	0.233	-0.089
	(0.174)	(0.321)
Income Over $150K: *$39K*	-0.024	-0.027
	(0.248)	(0.489)
Age 16: *25*	-0.686	-0.804
	(0.366)	(0.899)
Age 45: *25*	-0.159	-0.466
	(0.218)	(0.365)
Age 65: *25*	0.093	0.109
	(0.250)	(0.442)
Age 75: *25*	1.046	1.879
	(0.622)	(0.773)*
Outer Regional: *Major City*	-0.062	0.030
	(0.236)	(0.431)
Inner Regional: *Major City*	-0.058	-0.284
	(0.184)	(0.319)
Ed Uni: *School*	-0.039	-0.078
	(0.127)	(0.229)
Vocational: *School*	0.108	0.324
	(0.128)	(0.244)
Observations	10,344	3,288
Individuals	431	137
Wald Chi	118.63	54.47
d.f.	24	24
p-value	<0.001	<0.001
Log-likelihood	-2,717.69	-834.05
Pseudo R^2	0.12	0.73

Notes: * $p<0.05$; ** $p<0.01$

All categorical data are effects coded to minimise the potential for contamination of the base effect in the grand mean.

Base categories are shown in italics.

All analyses were conducted in STATA 12, with clustering on respondent ID to account for within respondent variation.

Abbreviations: BC, breast cancer; CBC, contralateral breast cancer; CPM, contralateral prophylactic mastectomy; d.f., degrees of freedom; Ed, education; OOP, out-of-pocket; Uni, university.

Table A 32: Pooled Regression with Sample Specific Interactions (Chow Test)

	Main Effects		Framing Interactions: Compared with Base	
	Mean Coefficient	**s.d.**	**No Efficacy Difference**	**Goal**
CBC Risk	-0.147	0.412	-0.193	0.144
	(0.050)**	(0.058)**	(0.069)**	(0.075)
Other BC Risk	-0.054	-0.075	n.a.	0.025
	(0.019)**	(0.043)	n.a.	(0.019)
Monitoring OOP, $'00	-0.081	0.082	0.030	0.013
	(0.018)**	(0.042)	(0.023)	(0.026)
Surgery OOP, $'00	-0.013	0.014	0.001	-0.001
	(0.002)**	(0.003)**	(0.002)	(0.002)
Pain Risk	-0.012	-0.016	0.011	-0.014
	(0.006)*	(0.008)*	(0.008)	(0.008)
Sensitivity Risk	-0.018	0.047	0.000	0.003
	(0.007)**	(0.008)**	(0.009)	(0.009)
Ultrasound: *Mammogram*	0.206	0.167	-0.006	0.012
	(0.063)**	(0.109)	(0.080)	(0.090)
MRI: *Mammogram*	-0.010	-0.345	-0.047	-0.078
	(0.060)	(0.178)	(0.084)	(0.088)
Annual: *Biannual Checks*	-0.038	-0.016	0.237	-0.220
	(0.063)	(0.099)	(0.082)**	(0.093)*
Biennial: *Biannual Checks*	0.037	-0.011	-0.205	0.275
	(0.056)	(0.075)	(0.077)**	(0.090)**
Involved Always: *Rarely*	0.170	-0.333	-0.001	0.079
	(0.052)**	(0.110)**	(0.070)	(0.076)
Monitoring: *CPM* Constant	1.196	2.225	0.167	-0.234
	(0.232)**	(0.198)**	(0.260)	(0.266)
Observations	11,136			
N	464			
Wald Chi	162.41			
d.f.	35			
p-value	<0.001			
Log-likelihood	-1,729.41			
Pseudo R^2	0.066			

Notes: * p<0.05, ** p<0.01, *** p<0.001
All categorical data are effects coded to minimise the potential for contamination of the base effect in the grand mean.
Base categories are shown in italics.
All analyses were conducted in STATA 12, using robust standard errors (in parentheses) to account for the survey nature of the data.
Analyses conducted with 1,000 replications.

Abbreviations: AIC, Akaike Information Criterion; BC denotes breast cancer; BIC, Bayesian Information Criterion; CBC contralateral breast cancer; CPM contralateral prophylactic mastectomy; d.f., degrees of freedom; OOP out-of-pocket; s.d., standard deviation.

Table A 33: Mixed Logit with Categorical Coding of Risk Attributes

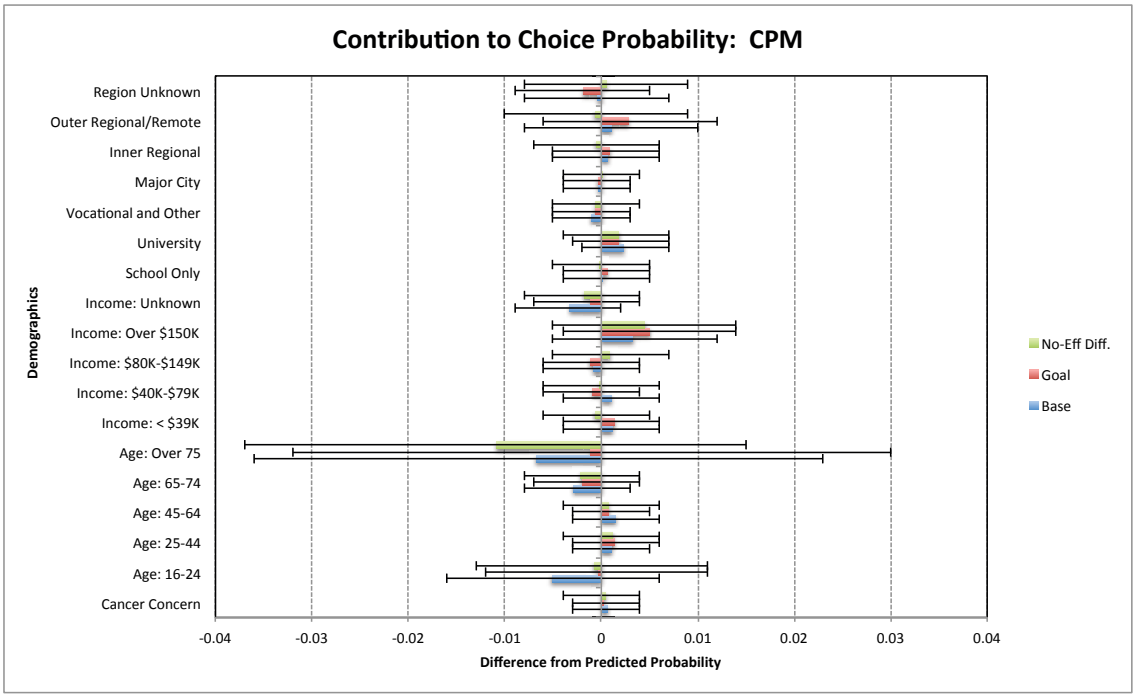
	Base	Goal	No-Efficacy Difference
Attribute Means			
CBC Risk 2%: *5%*	-0.902	0.931	0.054
	(0.381)*	(0.476)	(0.236)
CBC Risk 1%: *5%*	1.332	-0.376	0.099
	(0.472)**	(0.416)	(0.232)
CBC Risk 0.3%: *5%*	0.828	-3.188	1.066
	(0.438)	(1.261)*	(0.331)**
Other BC Risk 15%: *20%*	0.258	-0.198	n.a.
	(0.334)	(0.309)	
Other BC Risk 10%: *20%*	1.759	0.384	n.a.
	(0.578)**	(0.391)	
Monitoring OOP, $'00	-0.435	-0.266	-0.035
	(0.120)**	(0.109)*	(0.034)
Surgery OOP, $'00	-0.086	-0.035	-0.013
	(0.019)**	(0.007)**	(0.004)**
Pain Risk 30%: *40%*	1.671	-0.973	-0.049
	(0.637)**	(0.602)	(0.258)
Pain Risk 20%: *40%*	-2.088	2.442	-0.136
	(0.694)**	(0.873)**	(0.322)
Pain Risk 10%: *40%*	0.946	0.962	-0.008
	(0.619)	(0.397)*	(0.283)
Sensitivity Risk 30%: *60%*	-0.390	1.056	0.002
	(0.837)	(0.819)	(0.293)
Sensitivity Risk 40%: *60%*	-0.980	0.777	-0.068
	(0.641)	(0.511)	(0.251)
Sensitivity Risk 50%: *60%*	3.182	-0.443	-0.068
	(0.921)**	(0.734)	(0.280)
Ultrasound: *Mammogram*	0.505	-0.409	0.428
	(0.278)	(0.296)	(0.147)**
MRI: *Mammogram*	0.211	-0.223	-0.052
	(0.257)	(0.292)	(0.127)
Annual: *Biannual Checks*	0.397	-0.620	0.281
	(0.282)	(0.396)	(0.139)*
Biennial: *Biannual Checks*	-0.054	0.190	-0.225
	(0.193)	(0.279)	(0.106)*
Involved Always: *Rarely*	-0.251	0.641	0.305
	(0.291)	(0.257)*	(0.135)*
Monitoring: *CPM* Constant	4.759	4.638	2.156
Attribute Standard Deviations			
	(0.979)**	(0.843)**	(0.448)**
CBC Risk 2%: *5%*	0.039	2.221	-0.033
	(0.302)	(0.571)**	(0.448)
CBC Risk 1%: *5%*	3.443	3.250	0.402
	(0.748)**	(0.664)**	(0.660)
CBC Risk 0.3%: *5%*	2.712	3.908	2.083
	(0.590)**	(0.860)**	(0.524)**
Other BC Risk 15%: *20%*	1.023	1.540	n.a.
	(0.339)**	(0.358)**	
Other BC Risk 10%: *20%*	-3.201	2.236	n.a.

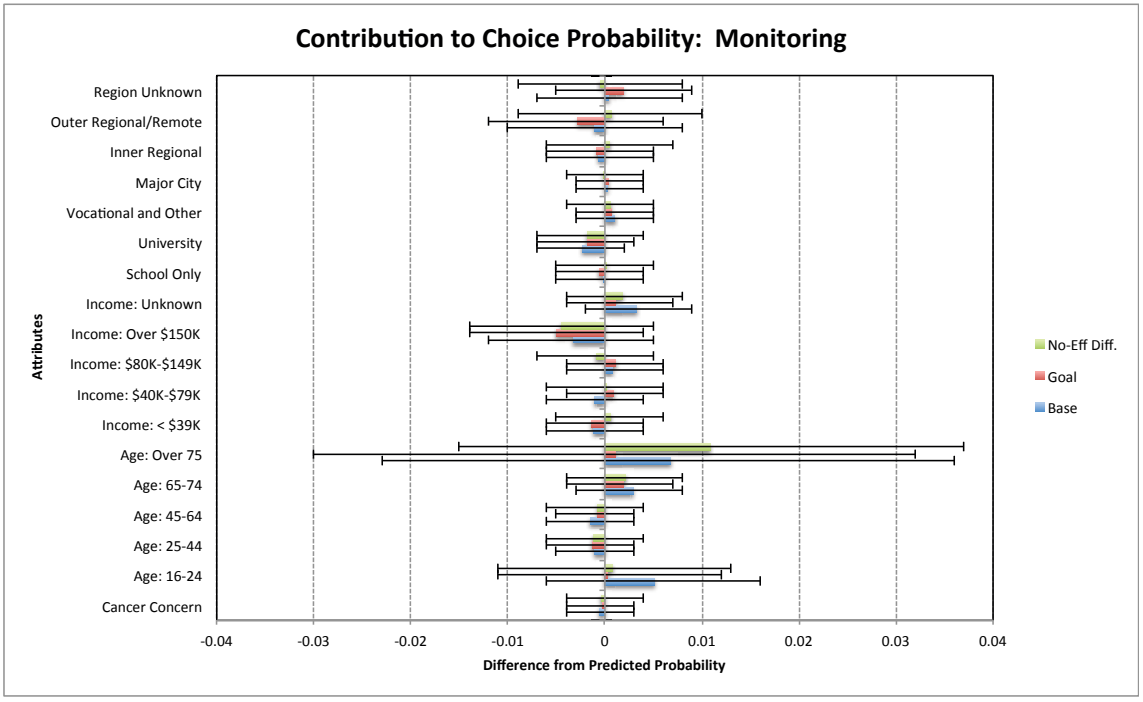
	Base	Goal	No-Efficacy Difference
	(0.756)**	(0.559)**	
Monitoring OOP, $'00	-0.488	0.346	-0.052
	(0.113)**	(0.118)**	(0.058)
Surgery OOP, $'00	-0.094	0.026	-0.004
	(0.020)**	(0.005)**	(0.004)
Pain Risk 30%: *40%*	1.192	-2.883	-0.926
	(0.727)	(0.625)**	(0.444)*
Pain Risk 20%: *40%*	-1.211	-1.697	-0.393
	(0.452)**	(0.510)**	(0.281)
Pain Risk 10%: *40%*	3.678	-2.508	-0.947
	(0.817)**	(0.503)**	(0.342)**
Sensitivity Risk 30%: *60%*	1.470	0.239	1.210
	(0.349)**	(0.380)	(0.410)**
Sensitivity Risk 40%: *60%*	0.498	0.070	-0.365
	(0.249)*	(0.414)	(0.408)
Sensitivity Risk 50%: *60%*	0.074	-2.171	-0.696
	(0.282)	(0.540)**	(0.307)*
Ultrasound: *Mammogram*	-0.178	1.751	-0.244
	(0.231)	(0.394)**	(0.282)
MRI: *Mammogram*	-2.502	2.682	0.372
	(0.576)**	(0.511)**	(0.388)
Annual: *Biannual Checks*	1.400	0.836	-0.064
	(0.310)**	(0.200)**	(0.118)
Biennial: *Biannual Checks*	0.114	-1.367	-0.049
	(0.179)	(0.324)**	(0.117)
Involved Always: *Rarely*	0.649	-0.979	0.729
	(0.240)**	(0.239)**	(0.210)**
Monitoring: *CPM* Constant	7.977	7.919	3.052
	(1.678)**	(1.505)**	(0.632)**
Observations	3,792	3,480	3,864
Individuals	158	145	161
Wald Chi	58.76	104.86	46.14
d.f.	19	19	17
p-value	<0.001	<0.001	<0.001
Log-likelihood	-532.51	-511.70	-631.68

Notes: * p<0.05; ** p<0.01
All categorical data are effects coded to minimise the potential for contamination of the base effect in the grand mean.
Base categories are shown in italics.
All analyses were conducted in STATA 12, with robust standard errors (in parentheses) to account for the survey nature of the data.
Analyses conducted with 1,000 replications.

Abbreviations: BC, breast cancer; CBC, contralateral breast cancer; CPM, contralateral prophylactic mastectomy; d.f., degrees of freedom; OOP, out-of-pocket.

Table A 34 Predicted Mixed Logit Probabilities and Demographics

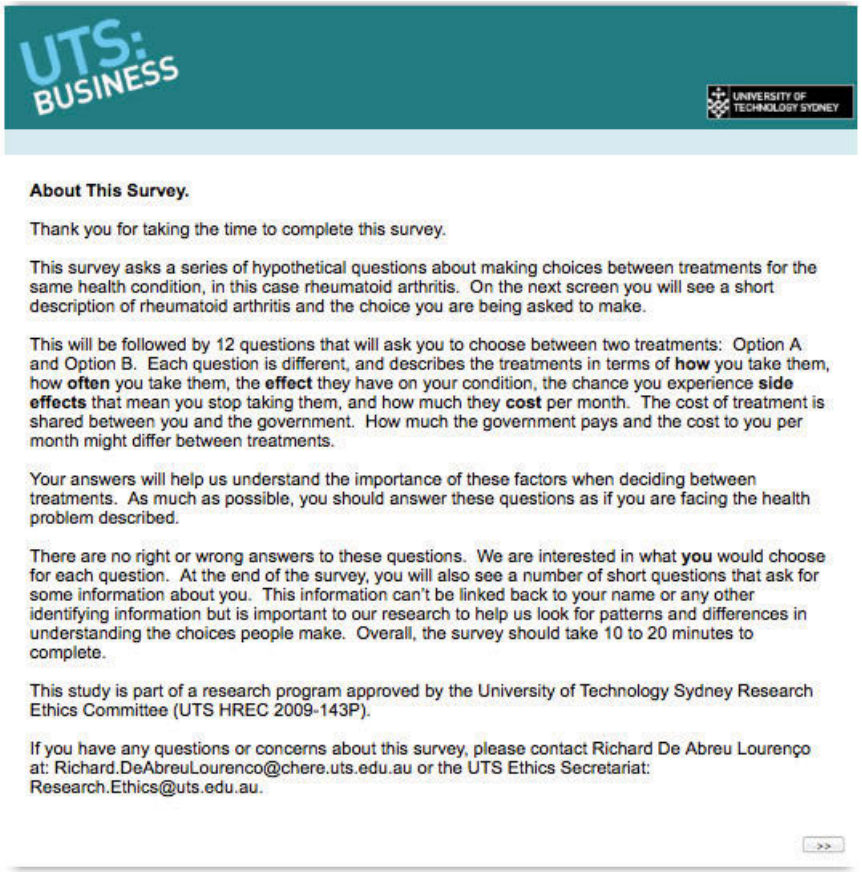


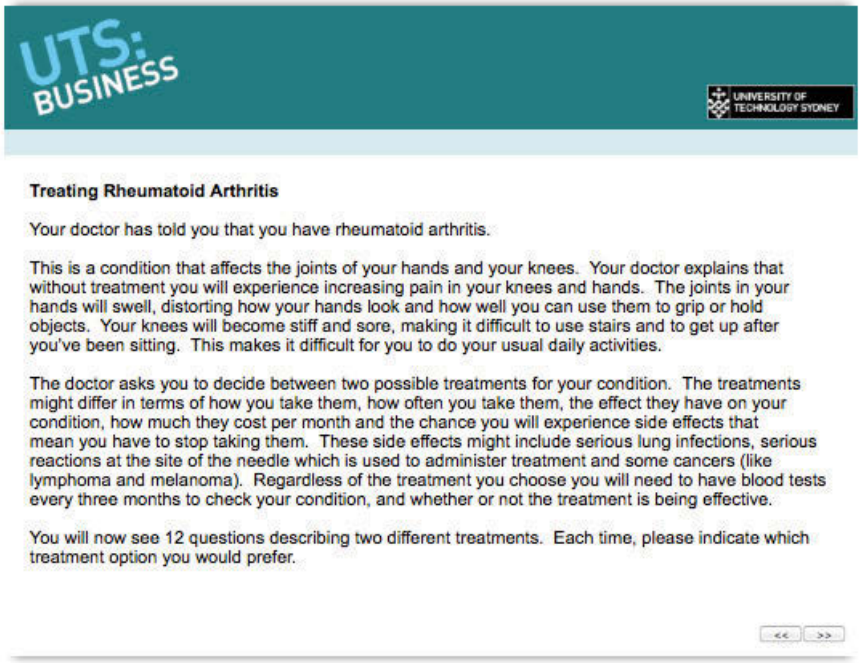


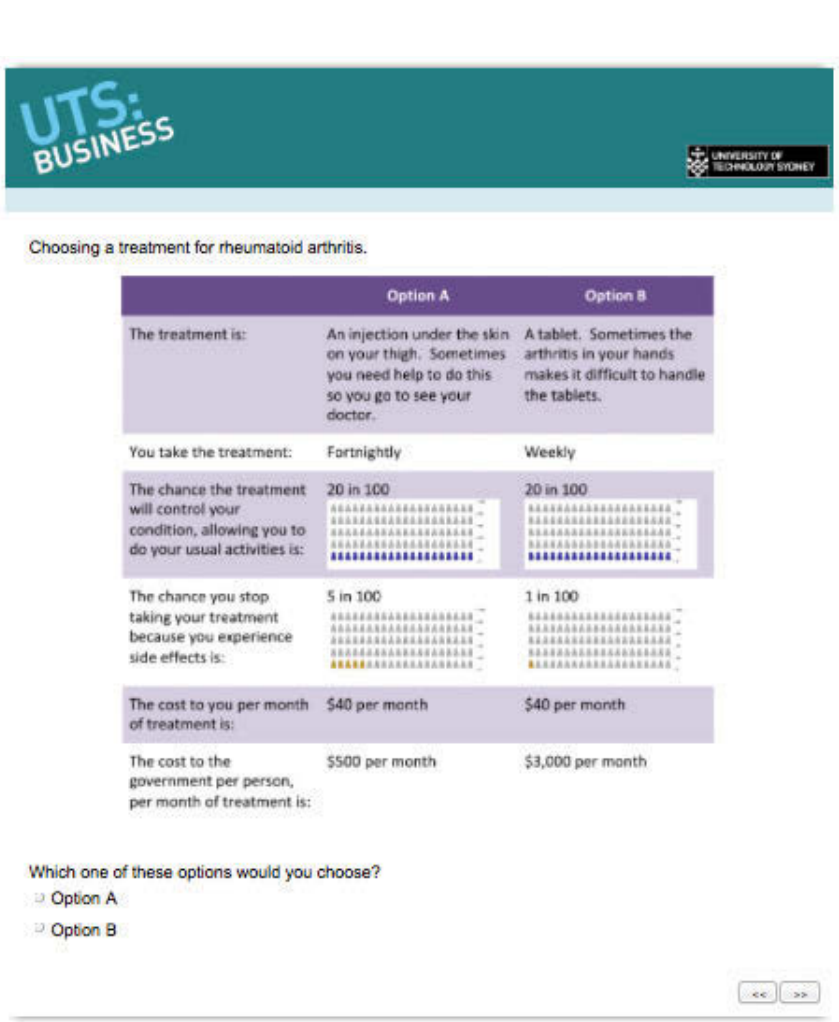
Appendix 15: Screenshots of Online Surveys

Survey Screenshots for Choices for the Treatment of RA

'Attribute' frame: Design block 2 (survey allocation randomised as block 6).







UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

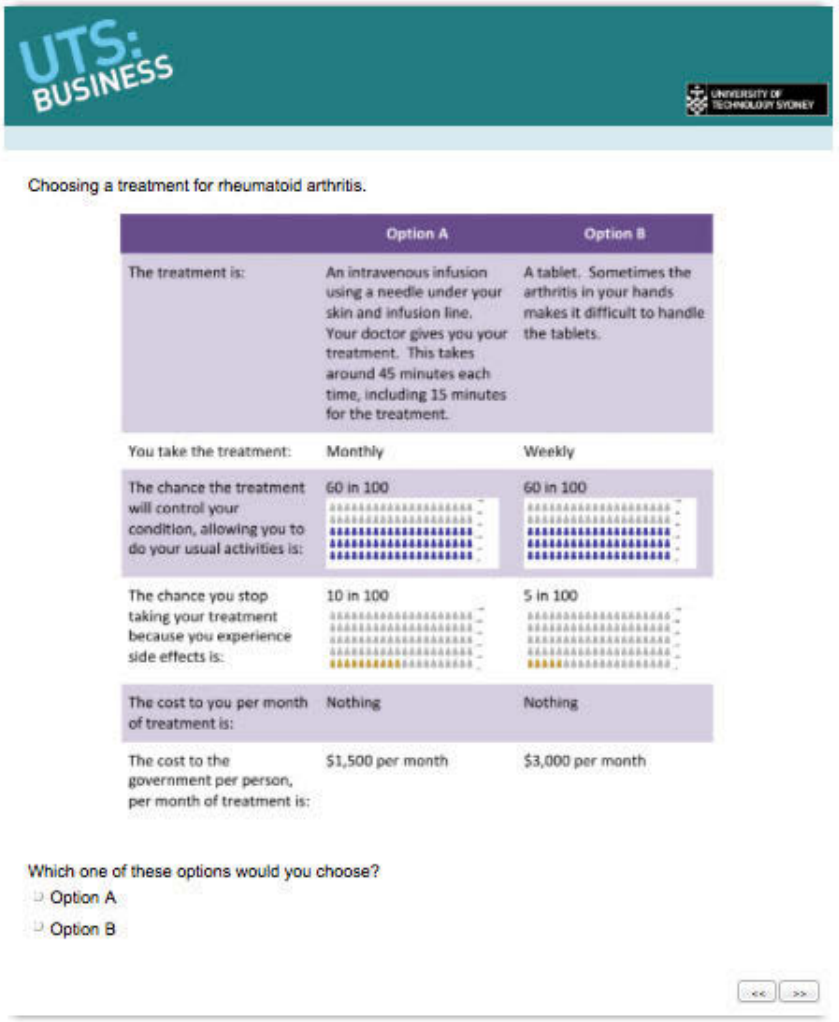
Choosing a treatment for rheumatoid arthritis.

	Option A	Option B
The treatment is:	An injection under the skin on your thigh. Sometimes you need help to do this so you go to see your doctor.	A tablet. Sometimes the arthritis in your hands makes it difficult to handle the tablets.
You take the treatment:	Fortnightly	Weekly
The chance the treatment will control your condition, allowing you to do your usual activities is:	20 in 100	20 in 100
The chance you stop taking your treatment because you experience side effects is:	5 in 100	1 in 100
The cost to you per month of treatment is:	$40 per month	$40 per month
The cost to the government per person, per month of treatment is:	$500 per month	$3,000 per month

Which one of these options would you choose?

- Option A
- Option B

<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

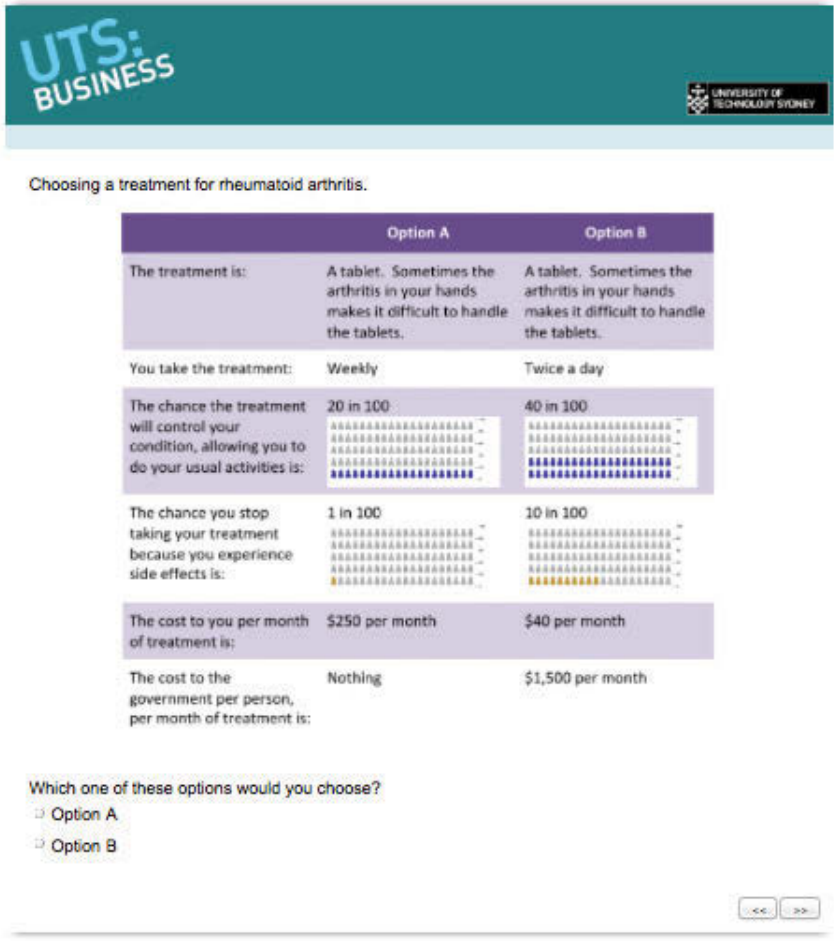
Choosing a treatment for rheumatoid arthritis.

	Option A	Option B
The treatment is:	An intravenous infusion using a needle under your skin and infusion line. Your doctor gives you your treatment. This takes around 45 minutes each time, including 15 minutes for the treatment.	A tablet. Sometimes the arthritis in your hands makes it difficult to handle the tablets.
You take the treatment:	Monthly	Weekly
The chance the treatment will control your condition, allowing you to do your usual activities is:	60 in 100	60 in 100
The chance you stop taking your treatment because you experience side effects is:	10 in 100	5 in 100
The cost to you per month of treatment is:	Nothing	Nothing
The cost to the government per person, per month of treatment is:	$1,500 per month	$3,000 per month

Which one of these options would you choose?

- Option A
- Option B

<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

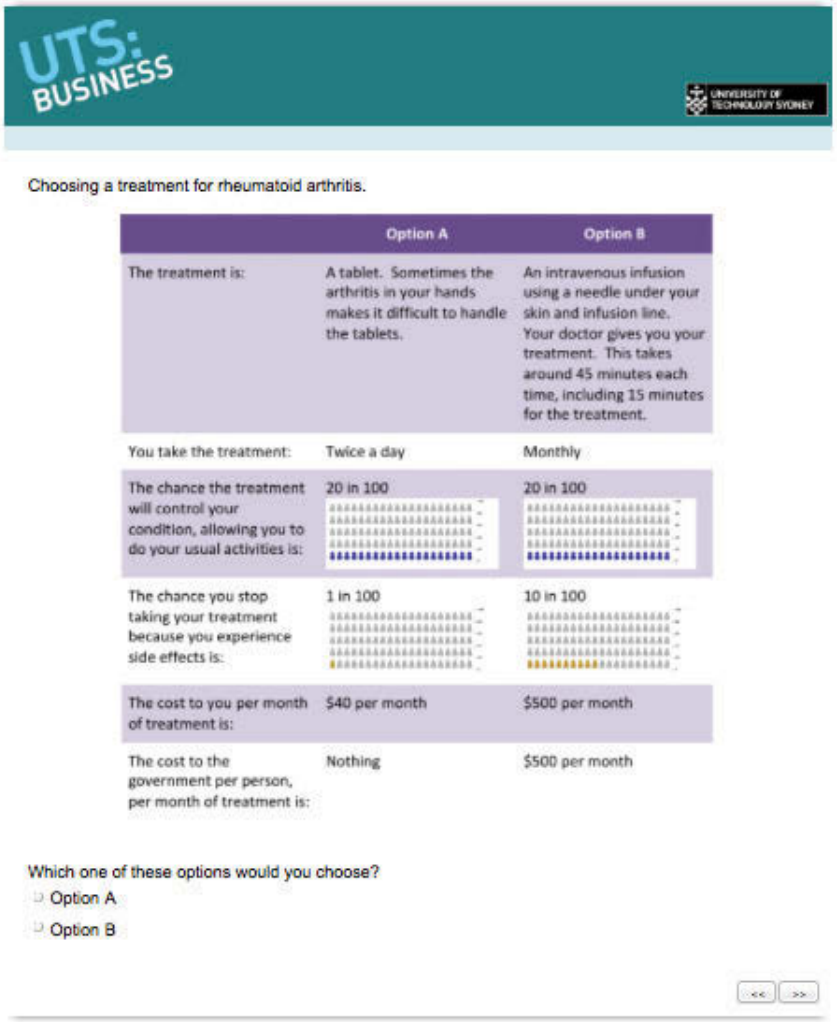
Choosing a treatment for rheumatoid arthritis.

	Option A	Option B
The treatment is:	A tablet. Sometimes the arthritis in your hands makes it difficult to handle the tablets.	A tablet. Sometimes the arthritis in your hands makes it difficult to handle the tablets.
You take the treatment:	Weekly	Twice a day
The chance the treatment will control your condition, allowing you to do your usual activities is:	20 in 100	40 in 100
The chance you stop taking your treatment because you experience side effects is:	1 in 100	10 in 100
The cost to you per month of treatment is:	$250 per month	$40 per month
The cost to the government per person, per month of treatment is:	Nothing	$1,500 per month

Which one of these options would you choose?

- Option A
- Option B

<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

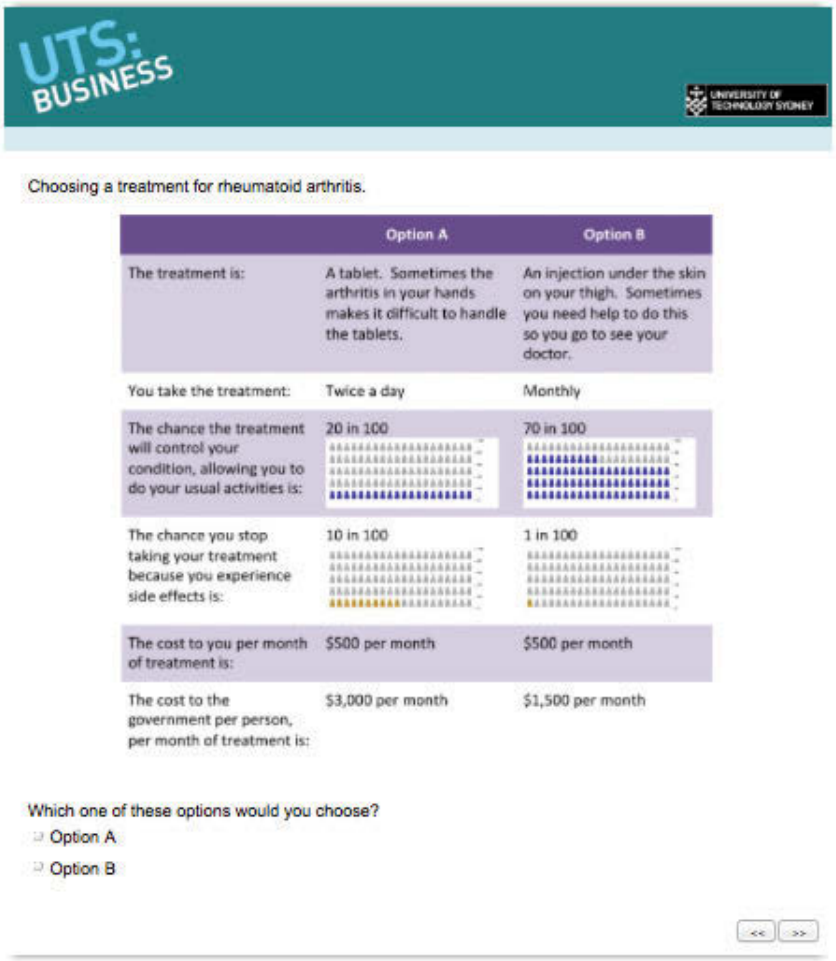
Choosing a treatment for rheumatoid arthritis.

	Option A	Option B
The treatment is:	A tablet. Sometimes the arthritis in your hands makes it difficult to handle the tablets.	An intravenous infusion using a needle under your skin and infusion line. Your doctor gives you your treatment. This takes around 45 minutes each time, including 15 minutes for the treatment.
You take the treatment:	Twice a day	Monthly
The chance the treatment will control your condition, allowing you to do your usual activities is:	20 in 100	20 in 100
The chance you stop taking your treatment because you experience side effects is:	1 in 100	10 in 100
The cost to you per month of treatment is:	$40 per month	$500 per month
The cost to the government per person, per month of treatment is:	Nothing	$500 per month

Which one of these options would you choose?

- Option A
- Option B

<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

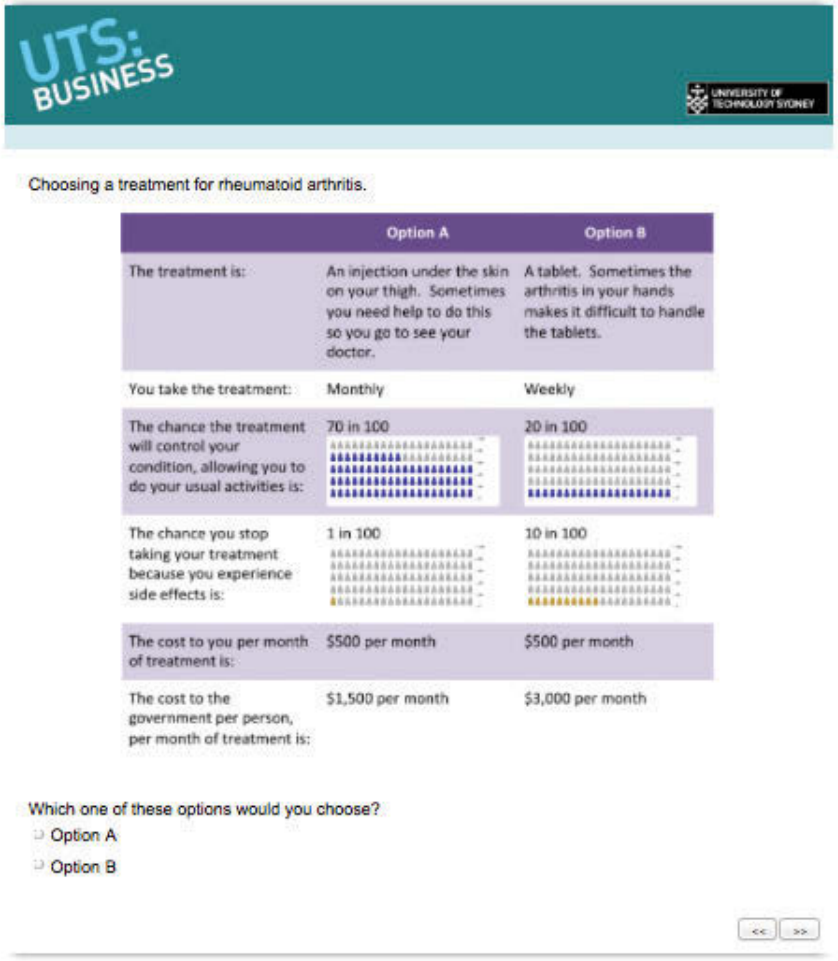
Choosing a treatment for rheumatoid arthritis.

	Option A	Option B
The treatment is:	A tablet. Sometimes the arthritis in your hands makes it difficult to handle the tablets.	An injection under the skin on your thigh. Sometimes you need help to do this so you go to see your doctor.
You take the treatment:	Twice a day	Monthly
The chance the treatment will control your condition, allowing you to do your usual activities is:	20 in 100	70 in 100
The chance you stop taking your treatment because you experience side effects is:	10 in 100	1 in 100
The cost to you per month of treatment is:	$500 per month	$500 per month
The cost to the government per person, per month of treatment is:	$3,000 per month	$1,500 per month

Which one of these options would you choose?

- Option A
- Option B

<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

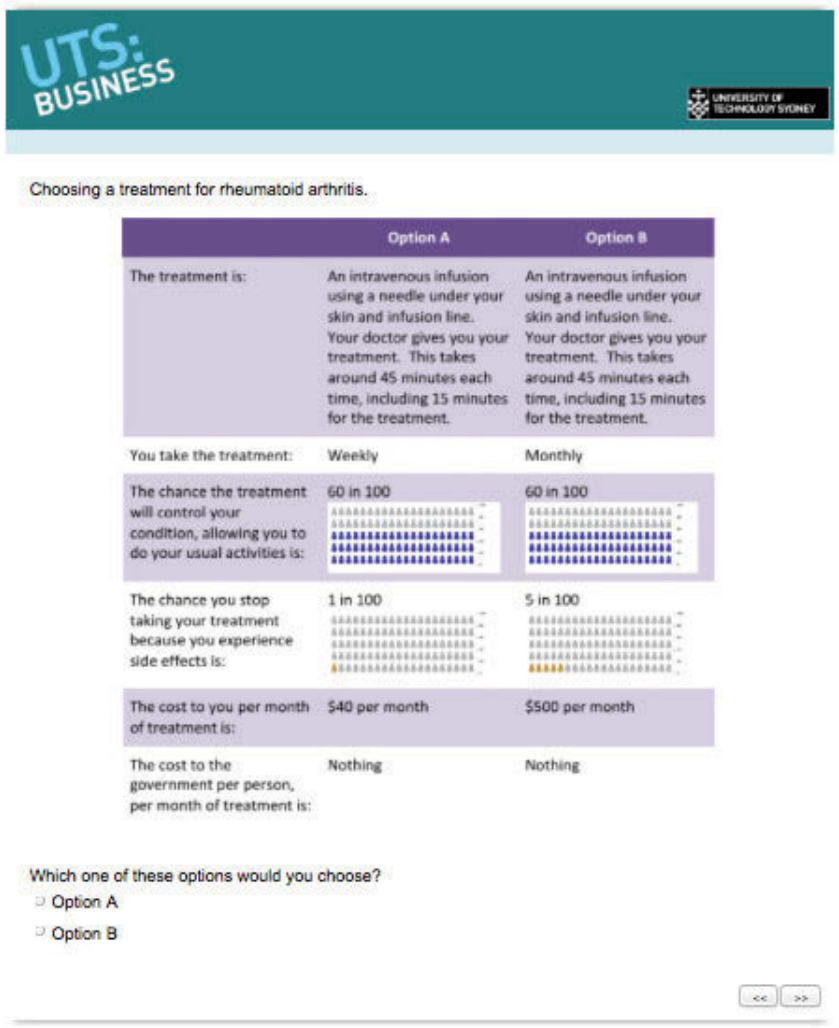
Choosing a treatment for rheumatoid arthritis.

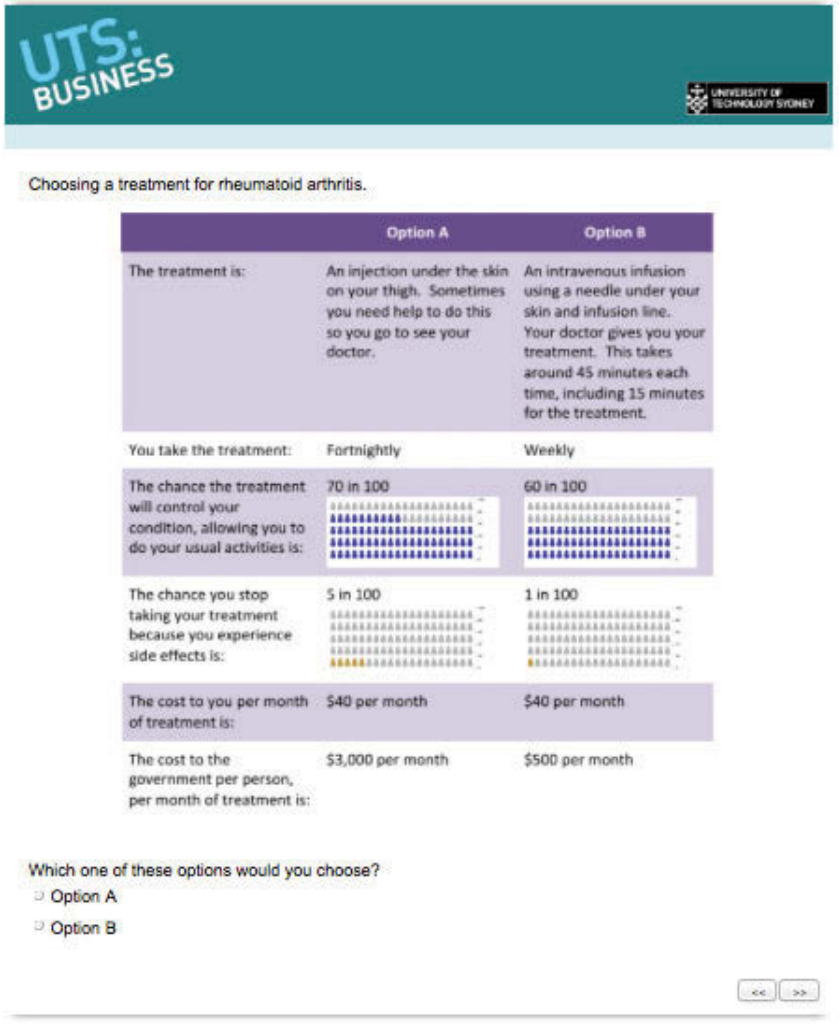
	Option A	Option B
The treatment is:	An injection under the skin on your thigh. Sometimes you need help to do this so you go to see your doctor.	A tablet. Sometimes the arthritis in your hands makes it difficult to handle the tablets.
You take the treatment:	Monthly	Weekly
The chance the treatment will control your condition, allowing you to do your usual activities is:	70 in 100	20 in 100
The chance you stop taking your treatment because you experience side effects is:	1 in 100	10 in 100
The cost to you per month of treatment is:	$500 per month	$500 per month
The cost to the government per person, per month of treatment is:	$1,500 per month	$3,000 per month

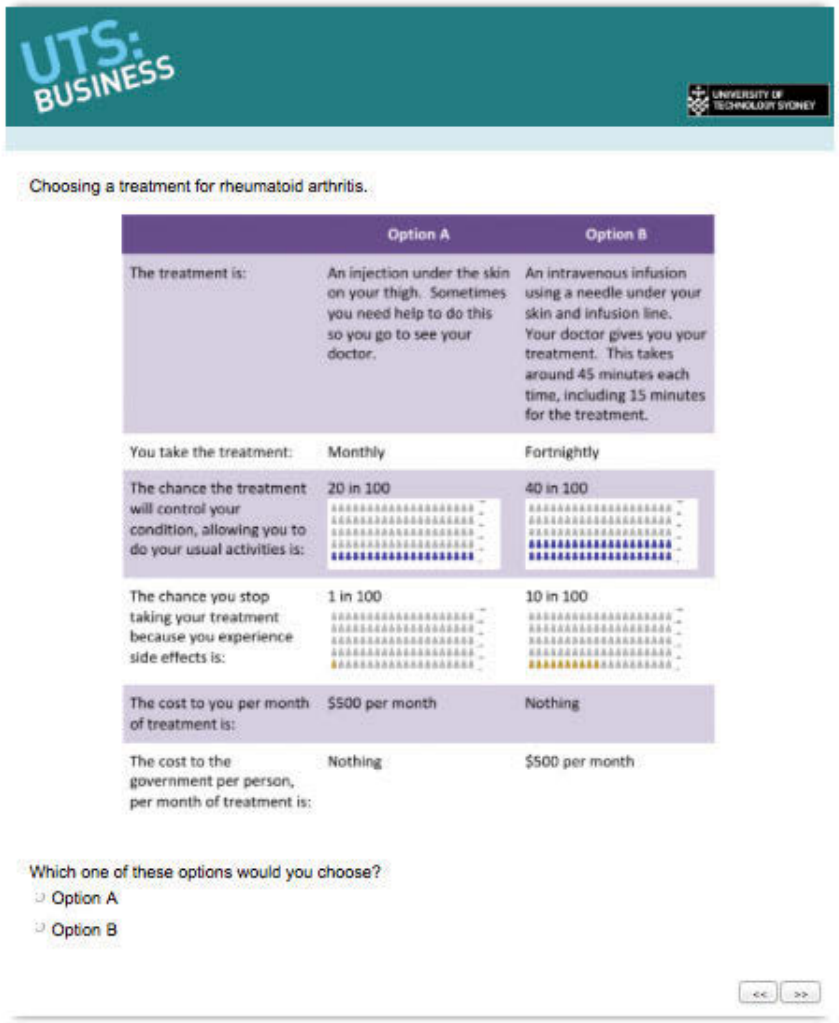
Which one of these options would you choose?

- Option A
- Option B

<< >>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Choosing a treatment for rheumatoid arthritis.
Option A
Option B
The treatment is:
An intravenous infusion using a needle under your skin and infusion line. Your doctor gives you your treatment. This takes around 45 minutes each time, including 15 minutes for the treatment.
An intravenous infusion using a needle under your skin and infusion line. Your doctor gives you your treatment. This takes around 45 minutes each time, including 15 minutes for the treatment.
You take the treatment:
Weekly
Monthly
The chance the treatment will control your condition, allowing you to do your usual activities is:
60 in 100
60 in 100
The chance you stop taking your treatment because you experience side effects is:
1 in 100
5 in 100
The cost to you per month of treatment is:
$40 per month
$500 per month
The cost to the government per person, per month of treatment is:
Nothing
Nothing
Which one of these options would you choose?
Option A
Option B
<< >>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Choosing a treatment for rheumatoid arthritis.
Option A
Option B
The treatment is:
An injection under the skin on your thigh. Sometimes you need help to do this so you go to see your doctor.
An intravenous infusion using a needle under your skin and infusion line. Your doctor gives you your treatment. This takes around 45 minutes each time, including 15 minutes for the treatment.
You take the treatment:
Fortnightly
Weekly
The chance the treatment will control your condition, allowing you to do your usual activities is:
70 in 100
60 in 100
The chance you stop taking your treatment because you experience side effects is:
5 in 100
1 in 100
The cost to you per month of treatment is:
$40 per month
$40 per month
The cost to the government per person, per month of treatment is:
$3,000 per month
$500 per month
Which one of these options would you choose?
Option A
Option B
<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

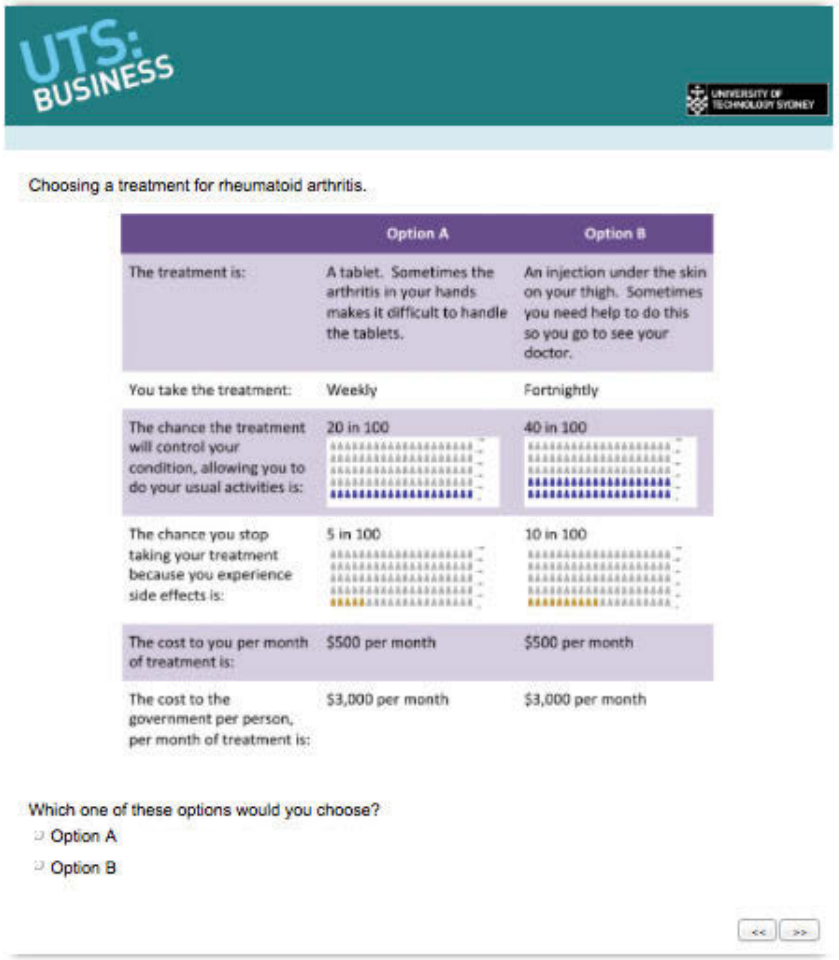
Choosing a treatment for rheumatoid arthritis.

	Option A	Option B
The treatment is:	An injection under the skin on your thigh. Sometimes you need help to do this so you go to see your doctor.	An intravenous infusion using a needle under your skin and infusion line. Your doctor gives you your treatment. This takes around 45 minutes each time, including 15 minutes for the treatment.
You take the treatment:	Monthly	Fortnightly
The chance the treatment will control your condition, allowing you to do your usual activities is:	20 in 100	40 in 100
The chance you stop taking your treatment because you experience side effects is:	1 in 100	10 in 100
The cost to you per month of treatment is:	$500 per month	Nothing
The cost to the government per person, per month of treatment is:	Nothing	$500 per month

Which one of these options would you choose?

- Option A
- Option B

<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

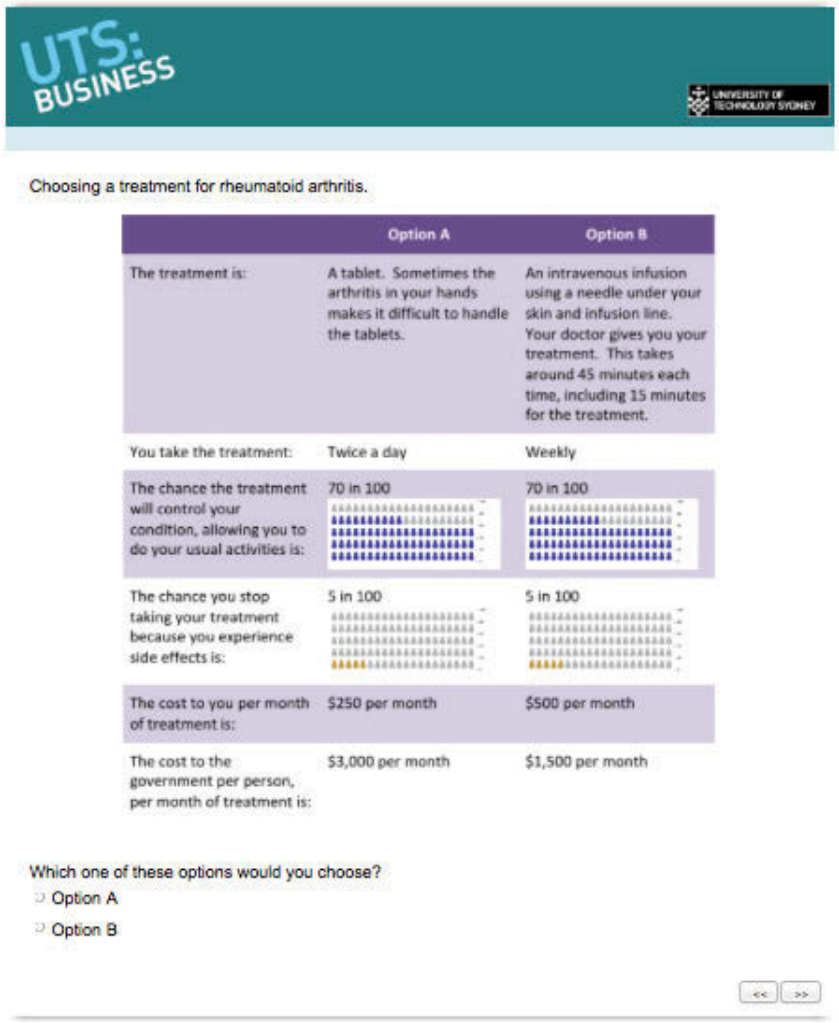
Choosing a treatment for rheumatoid arthritis.

	Option A	Option B
The treatment is:	A tablet. Sometimes the arthritis in your hands makes it difficult to handle the tablets.	An injection under the skin on your thigh. Sometimes you need help to do this so you go to see your doctor.
You take the treatment:	Weekly	Fortnightly
The chance the treatment will control your condition, allowing you to do your usual activities is:	20 in 100	40 in 100
The chance you stop taking your treatment because you experience side effects is:	5 in 100	10 in 100
The cost to you per month of treatment is:	$500 per month	$500 per month
The cost to the government per person, per month of treatment is:	$3,000 per month	$3,000 per month

Which one of these options would you choose?

- Option A
- Option B

<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

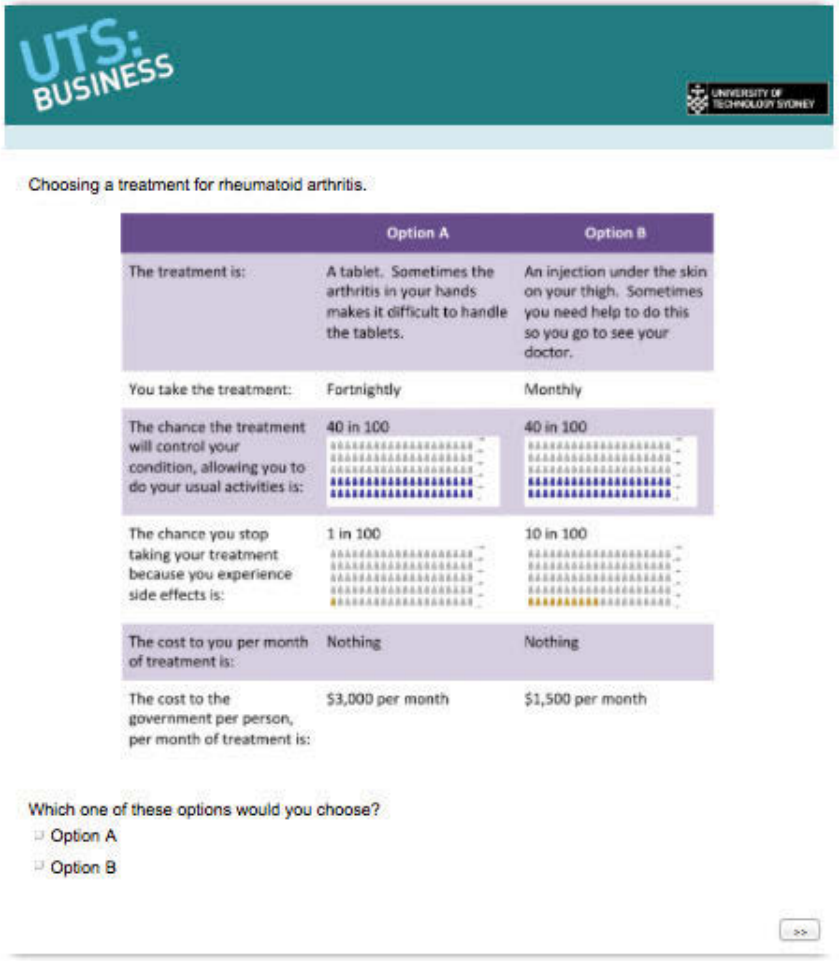
Choosing a treatment for rheumatoid arthritis.

	Option A	Option B
The treatment is:	A tablet. Sometimes the arthritis in your hands makes it difficult to handle the tablets.	An intravenous infusion using a needle under your skin and infusion line. Your doctor gives you your treatment. This takes around 45 minutes each time, including 15 minutes for the treatment.
You take the treatment:	Twice a day	Weekly
The chance the treatment will control your condition, allowing you to do your usual activities is:	70 in 100	70 in 100
The chance you stop taking your treatment because you experience side effects is:	5 in 100	5 in 100
The cost to you per month of treatment is:	$250 per month	$500 per month
The cost to the government per person, per month of treatment is:	$3,000 per month	$1,500 per month

Which one of these options would you choose?

- Option A
- Option B

<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

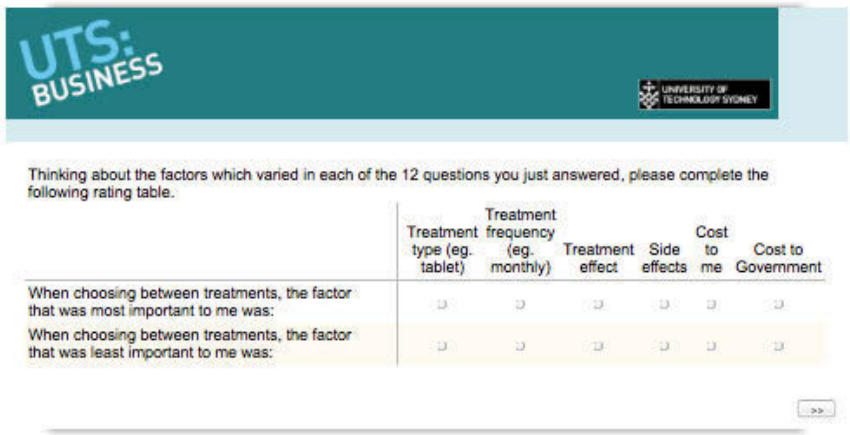
Choosing a treatment for rheumatoid arthritis.

	Option A	Option B
The treatment is:	A tablet. Sometimes the arthritis in your hands makes it difficult to handle the tablets.	An injection under the skin on your thigh. Sometimes you need help to do this so you go to see your doctor.
You take the treatment:	Fortnightly	Monthly
The chance the treatment will control your condition, allowing you to do your usual activities is:	40 in 100	40 in 100
The chance you stop taking your treatment because you experience side effects is:	1 in 100	10 in 100
The cost to you per month of treatment is:	Nothing	Nothing
The cost to the government per person, per month of treatment is:	$3,000 per month	$1,500 per month

Which one of these options would you choose?

- Option A
- Option B

>>

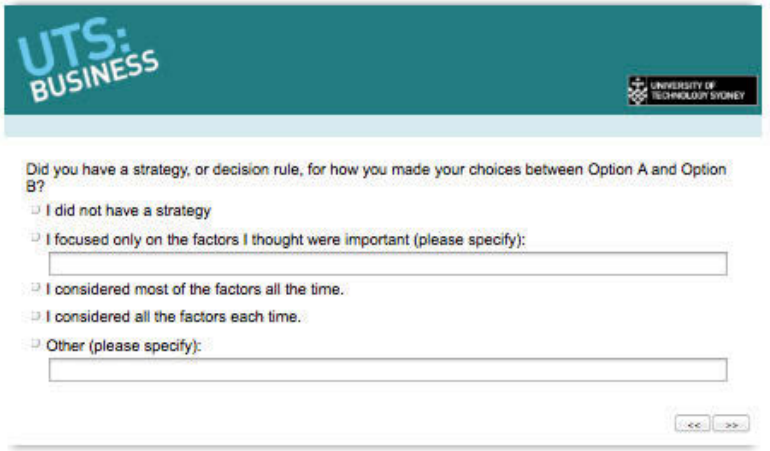

UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

Thinking about the factors which varied in each of the 12 questions you just answered, please complete the following rating table.

	Treatment type (eg. tablet)	Treatment frequency (eg. monthly)	Treatment effect	Side effects	Cost to me	Cost to Government
When choosing between treatments, the factor that was most important to me was:	☐	☐	☐	☐	☐	☐
When choosing between treatments, the factor that was least important to me was:	☐	☐	☐	☐	☐	☐

>>

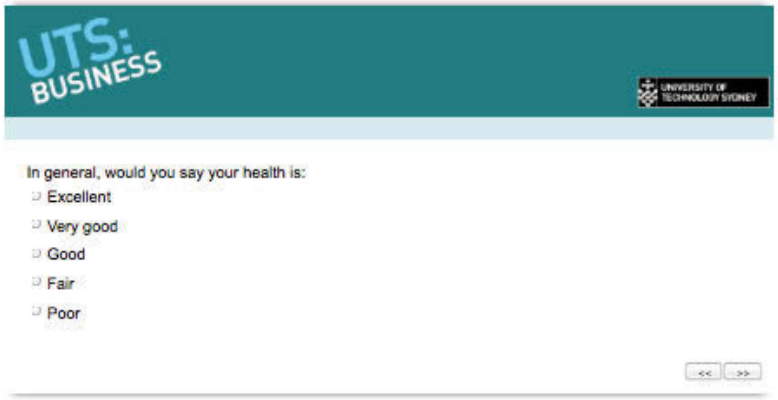

UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

Did you have a strategy, or decision rule, for how you made your choices between Option A and Option B?

- ☐ I did not have a strategy
- ☐ I focused only on the factors I thought were important (please specify):
- ☐ I considered most of the factors all the time.
- ☐ I considered all the factors each time.
- ☐ Other (please specify):

<< >>

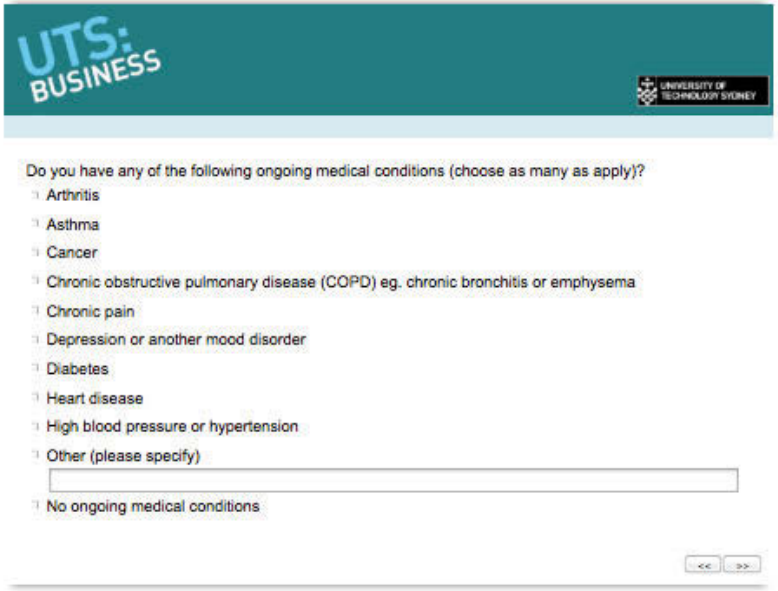

UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

In general, would you say your health is:

- ☐ Excellent
- ☐ Very good
- ☐ Good
- ☐ Fair
- ☐ Poor

<< >>

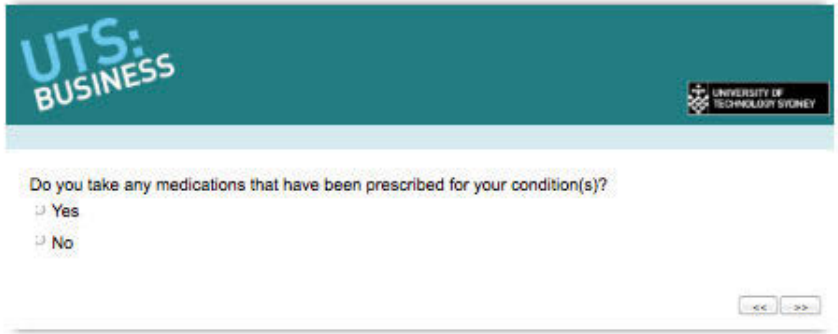

UTS: BUSINESS

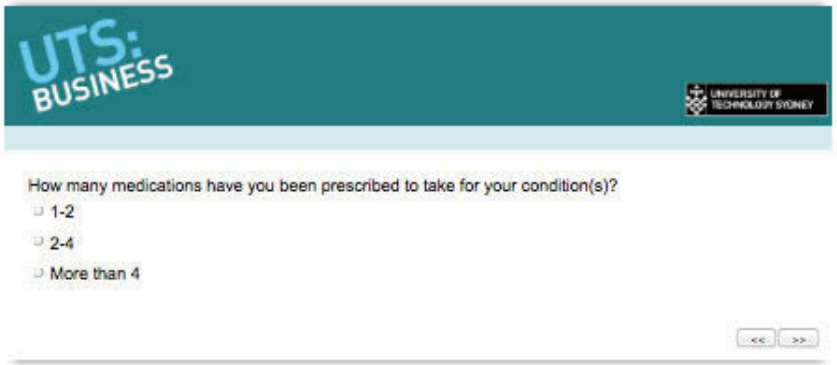
UNIVERSITY OF TECHNOLOGY SYDNEY

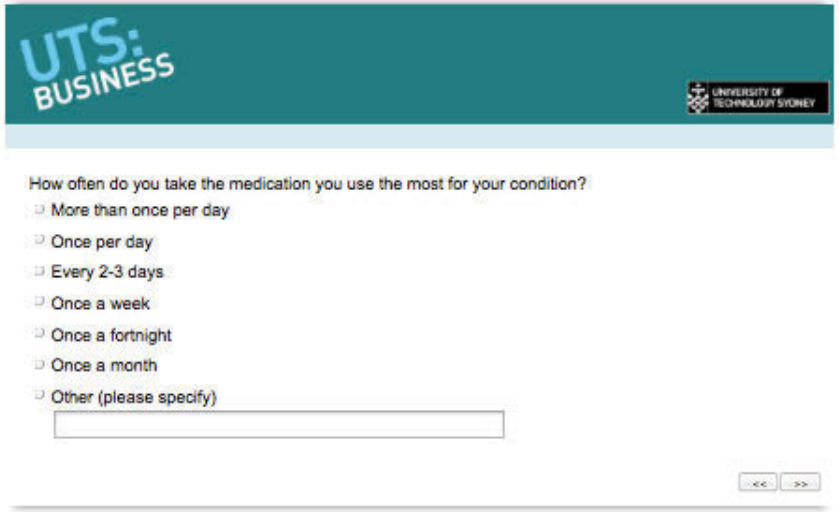
Do you have any of the following ongoing medical conditions (choose as many as apply)?

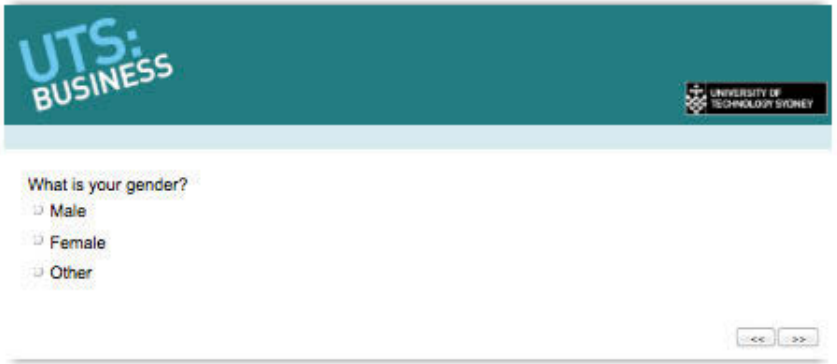
- ☐ Arthritis
- ☐ Asthma
- ☐ Cancer
- ☐ Chronic obstructive pulmonary disease (COPD) eg. chronic bronchitis or emphysema
- ☐ Chronic pain
- ☐ Depression or another mood disorder
- ☐ Diabetes
- ☐ Heart disease
- ☐ High blood pressure or hypertension
- ☐ Other (please specify)
- ☐ No ongoing medical conditions

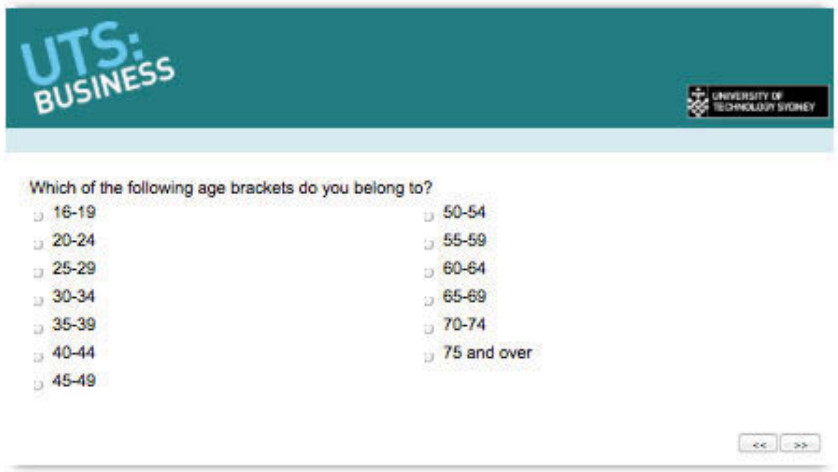
<< >>

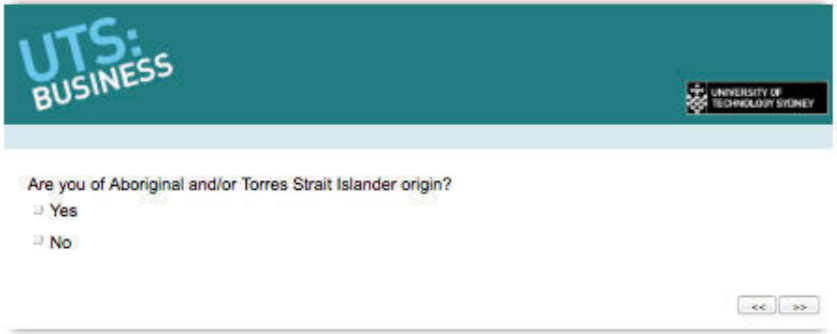

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Do you take any medications that have been prescribed for your condition(s)?
Yes
No
<< >>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
How many medications have you been prescribed to take for your condition(s)?
1-2
2-4
More than 4
<< >>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
How often do you take the medication you use the most for your condition?
More than once per day
Once per day
Every 2-3 days
Once a week
Once a fortnight
Once a month
Other (please specify)
<< >>

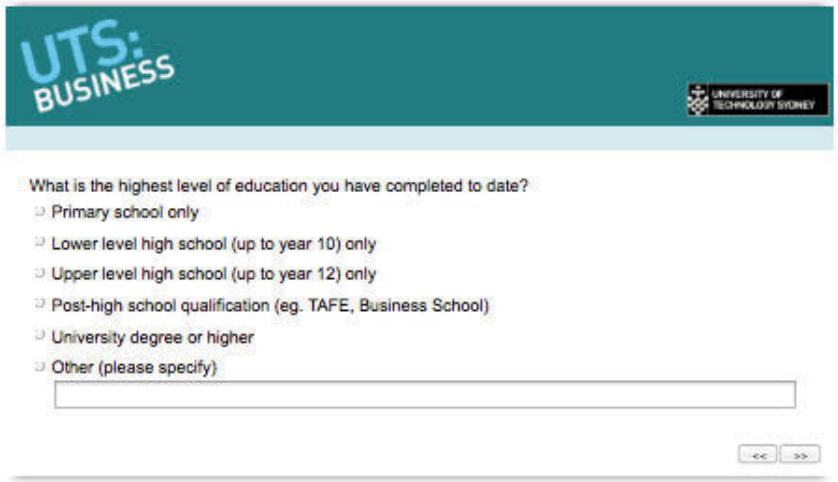

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
What is your gender?
Male
Female
Other
<< >>

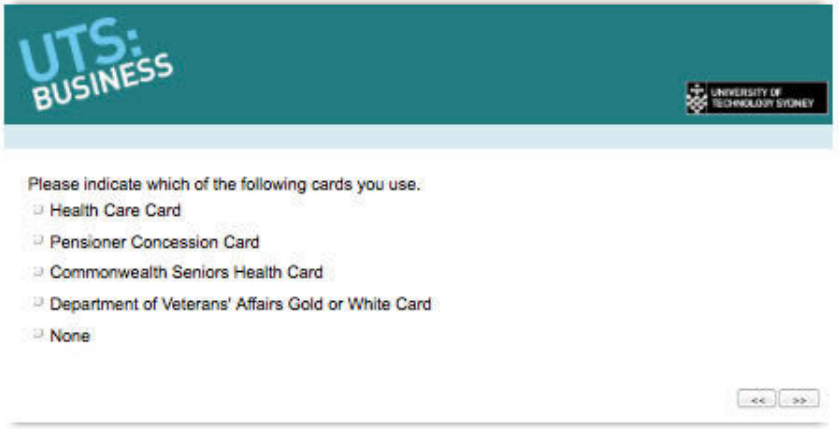

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Which of the following age brackets do you belong to?
16-19
20-24
25-29
30-34
35-39
40-44
45-49
50-54
55-59
60-64
65-69
70-74
75 and over
<< >>

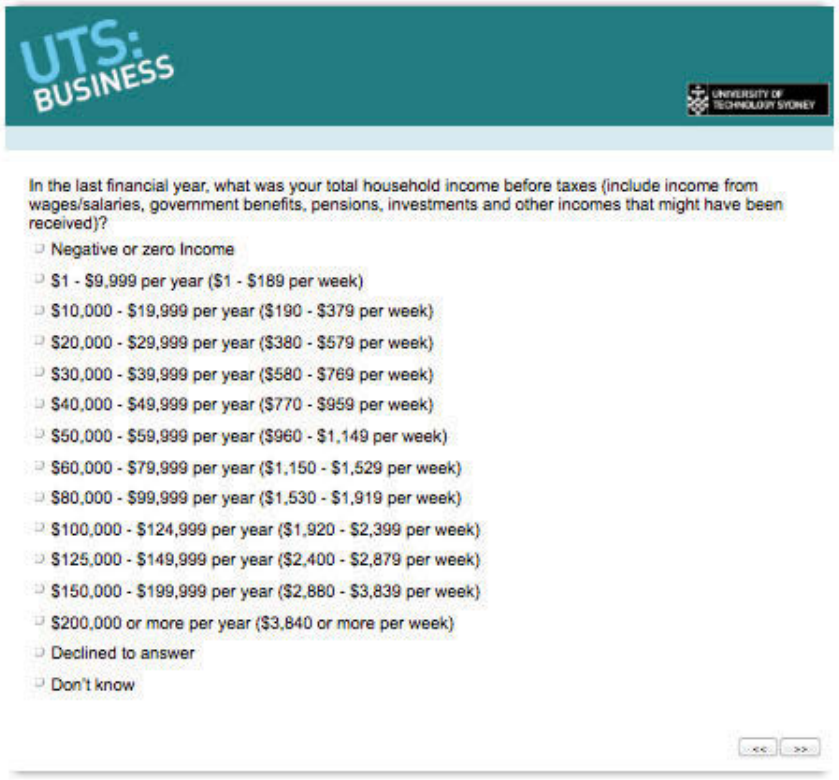

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Are you of Aboriginal and/or Torres Strait Islander origin?
Yes
No
<< >>

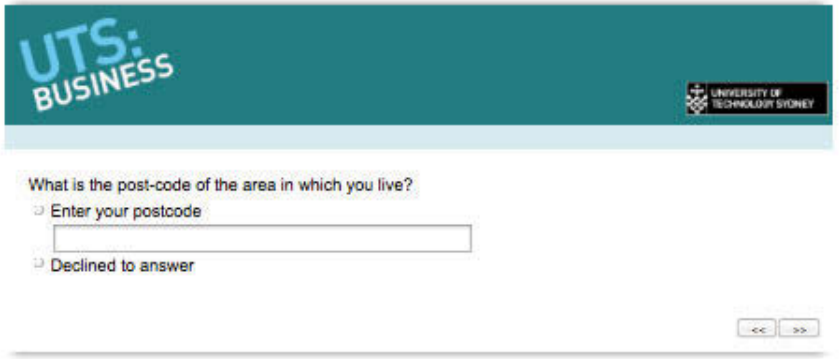

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Where were you born?
Australia
China
England
Germany
Greece
Italy
Lebanon
Netherlands
New Zealand
Scotland
Vietnam
Other
<< >>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Which language do you mainly speak at home?
Arabic
Cantonese
English
German
Greek
Hindi
Italian
Mandarin
Spanish
Vietnamese
Other
<< >>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
What is the highest level of education you have completed to date?
Primary school only
Lower level high school (up to year 10) only
Upper level high school (up to year 12) only
Post-high school qualification (eg. TAFE, Business School)
University degree or higher
Other (please specify)
<< >>

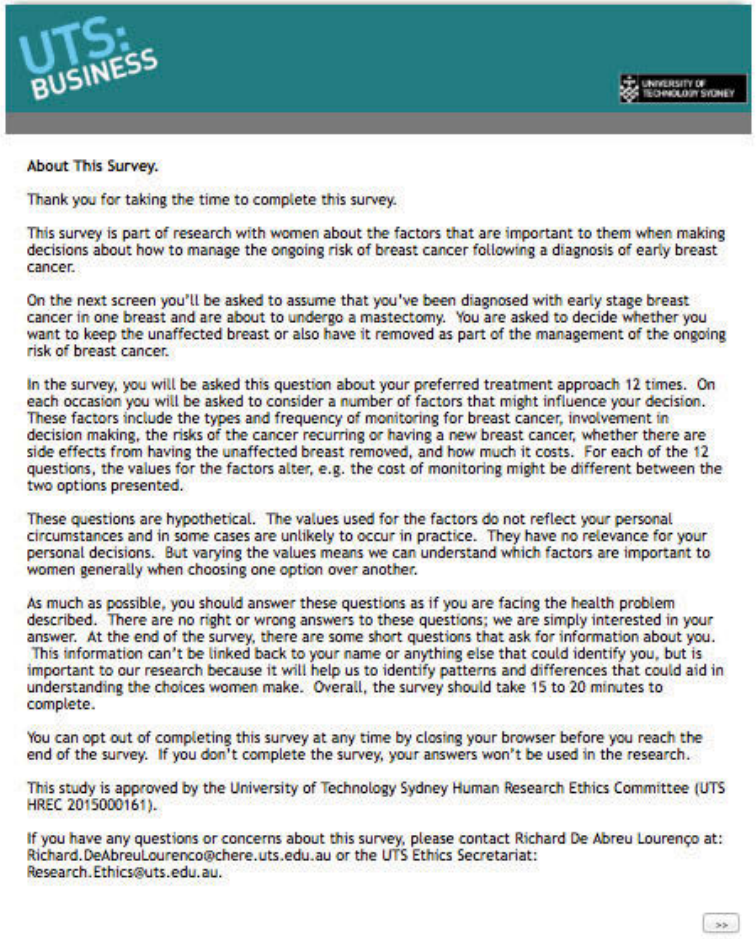

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Please indicate which of the following cards you use.
Health Care Card
Pensioner Concession Card
Commonwealth Seniors Health Card
Department of Veterans' Affairs Gold or White Card
None
<< >>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
In the last financial year, what was your total household income before taxes (include income from wages/salaries, government benefits, pensions, investments and other incomes that might have been received)?
Negative or zero Income
$1 - $9,999 per year ($1 - $189 per week)
$10,000 - $19,999 per year ($190 - $379 per week)
$20,000 - $29,999 per year ($380 - $579 per week)
$30,000 - $39,999 per year ($580 - $769 per week)
$40,000 - $49,999 per year ($770 - $959 per week)
$50,000 - $59,999 per year ($960 - $1,149 per week)
$60,000 - $79,999 per year ($1,150 - $1,529 per week)
$80,000 - $99,999 per year ($1,530 - $1,919 per week)
$100,000 - $124,999 per year ($1,920 - $2,399 per week)
$125,000 - $149,999 per year ($2,400 - $2,879 per week)
$150,000 - $199,999 per year ($2,880 - $3,839 per week)
$200,000 or more per year ($3,840 or more per week)
Declined to answer
Don't know
<< >>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
What is the post-code of the area in which you live?
Enter your postcode
Declined to answer
<< >>

Survey Screenshots: Managing the Ongoing Risks of Breast Cancer Recurrence

'Goal' frame: Design block 9 Left (randomisation block 17).



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

About This Survey.

Thank you for taking the time to complete this survey.

This survey is part of research with women about the factors that are important to them when making decisions about how to manage the ongoing risk of breast cancer following a diagnosis of early breast cancer.

On the next screen you'll be asked to assume that you've been diagnosed with early stage breast cancer in one breast and are about to undergo a mastectomy. You are asked to decide whether you want to keep the unaffected breast or also have it removed as part of the management of the ongoing risk of breast cancer.

In the survey, you will be asked this question about your preferred treatment approach 12 times. On each occasion you will be asked to consider a number of factors that might influence your decision. These factors include the types and frequency of monitoring for breast cancer, involvement in decision making, the risks of the cancer recurring or having a new breast cancer, whether there are side effects from having the unaffected breast removed, and how much it costs. For each of the 12 questions, the values for the factors alter, e.g. the cost of monitoring might be different between the two options presented.

These questions are hypothetical. The values used for the factors do not reflect your personal circumstances and in some cases are unlikely to occur in practice. They have no relevance for your personal decisions. But varying the values means we can understand which factors are important to women generally when choosing one option over another.

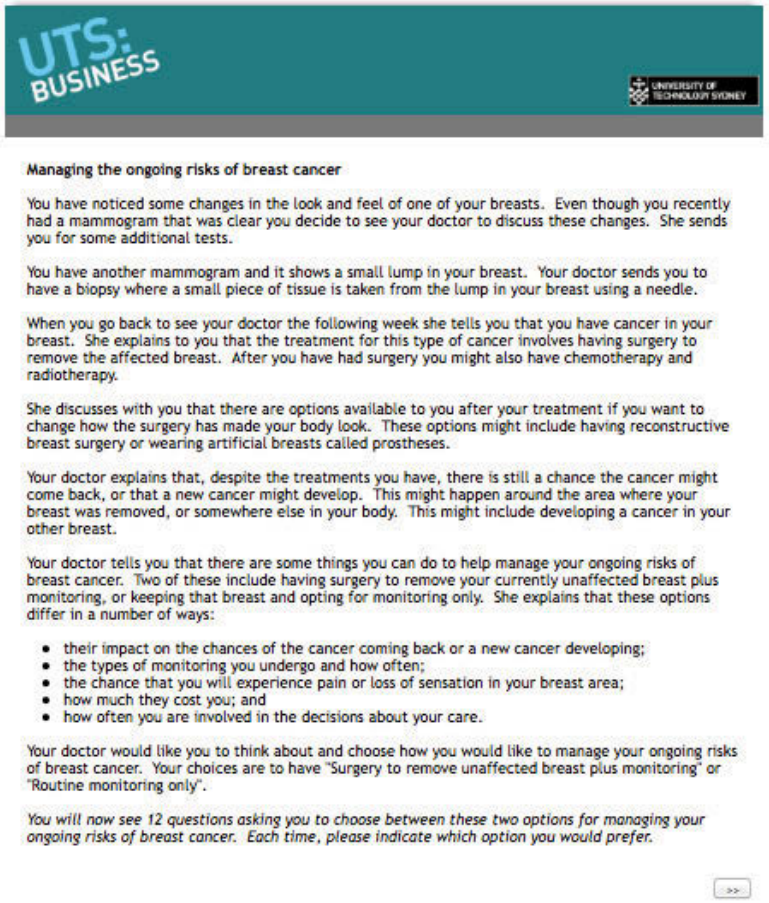
As much as possible, you should answer these questions as if you are facing the health problem described. There are no right or wrong answers to these questions; we are simply interested in your answer. At the end of the survey, there are some short questions that ask for information about you. This information can't be linked back to your name or anything else that could identify you, but is important to our research because it will help us to identify patterns and differences that could aid in understanding the choices women make. Overall, the survey should take 15 to 20 minutes to complete.

You can opt out of completing this survey at any time by closing your browser before you reach the end of the survey. If you don't complete the survey, your answers won't be used in the research.

This study is approved by the University of Technology Sydney Human Research Ethics Committee (UTS HREC 2015000161).

If you have any questions or concerns about this survey, please contact Richard De Abreu Lourenço at: Richard.DeAbreuLourenco@chere.uts.edu.au or the UTS Ethics Secretariat: Research.Ethics@uts.edu.au.

>>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

Managing the ongoing risks of breast cancer

You have noticed some changes in the look and feel of one of your breasts. Even though you recently had a mammogram that was clear you decide to see your doctor to discuss these changes. She sends you for some additional tests.

You have another mammogram and it shows a small lump in your breast. Your doctor sends you to have a biopsy where a small piece of tissue is taken from the lump in your breast using a needle.

When you go back to see your doctor the following week she tells you that you have cancer in your breast. She explains to you that the treatment for this type of cancer involves having surgery to remove the affected breast. After you have had surgery you might also have chemotherapy and radiotherapy.

She discusses with you that there are options available to you after your treatment if you want to change how the surgery has made your body look. These options might include having reconstructive breast surgery or wearing artificial breasts called prostheses.

Your doctor explains that, despite the treatments you have, there is still a chance the cancer might come back, or that a new cancer might develop. This might happen around the area where your breast was removed, or somewhere else in your body. This might include developing a cancer in your other breast.

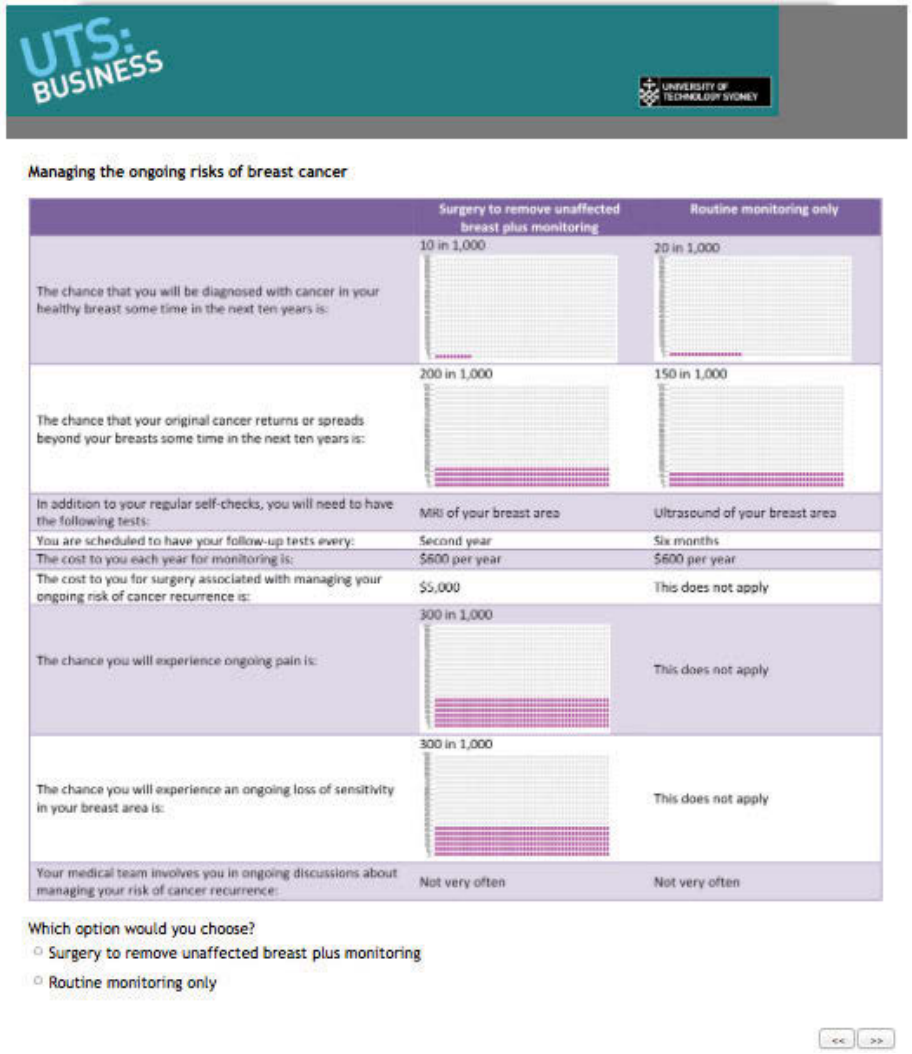
Your doctor tells you that there are some things you can do to help manage your ongoing risks of breast cancer. Two of these include having surgery to remove your currently unaffected breast plus monitoring, or keeping that breast and opting for monitoring only. She explains that these options differ in a number of ways:

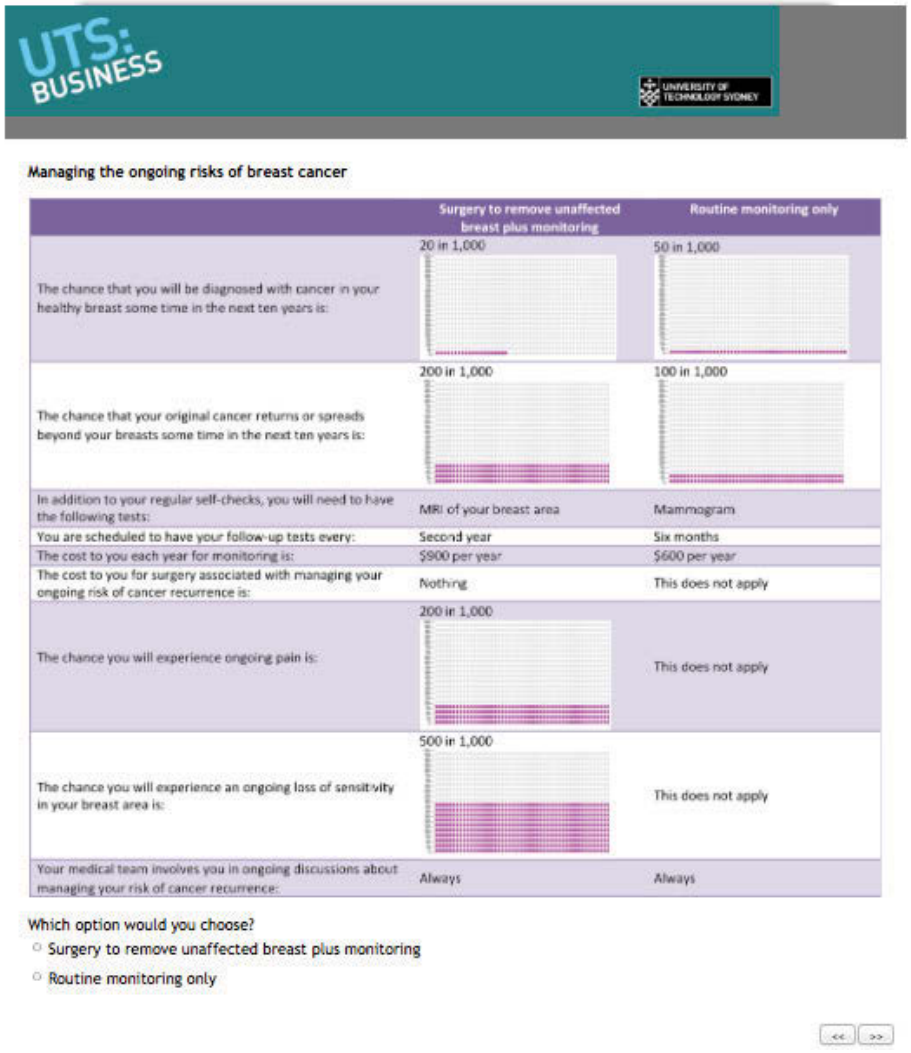
- their impact on the chances of the cancer coming back or a new cancer developing;
- the types of monitoring you undergo and how often;
- the chance that you will experience pain or loss of sensation in your breast area;
- how much they cost you; and
- how often you are involved in the decisions about your care.

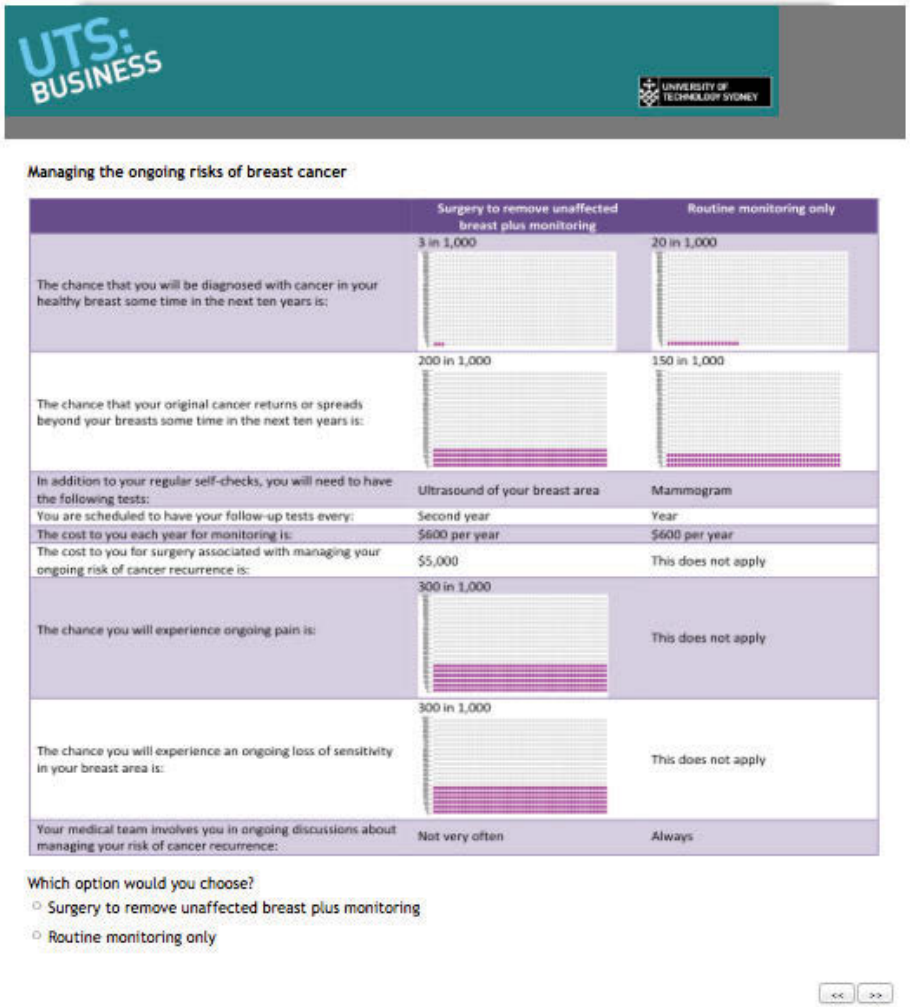
Your doctor would like you to think about and choose how you would like to manage your ongoing risks of breast cancer. Your choices are to have "Surgery to remove unaffected breast plus monitoring" or "Routine monitoring only".

You will now see 12 questions asking you to choose between these two options for managing your ongoing risks of breast cancer. Each time, please indicate which option you would prefer.

>>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Managing the ongoing risks of breast cancer
Surgery to remove unaffected breast plus monitoring
Routine monitoring only
The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:
10 in 1,000
20 in 1,000
The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:
200 in 1,000
150 in 1,000
In addition to your regular self-checks, you will need to have the following tests:
MRI of your breast area
Ultrasound of your breast area
You are scheduled to have your follow-up tests every:
Second year
Six months
The cost to you each year for monitoring is:
$600 per year
$600 per year
The cost to you for surgery associated with managing your ongoing risk of cancer recurrence is:
$5,000
This does not apply
The chance you will experience ongoing pain is:
300 in 1,000
This does not apply
The chance you will experience an ongoing loss of sensitivity in your breast area is:
300 in 1,000
This does not apply
Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence:
Not very often
Not very often
Which option would you choose?
Surgery to remove unaffected breast plus monitoring
Routine monitoring only
<< >>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Managing the ongoing risks of breast cancer
Surgery to remove unaffected breast plus monitoring
Routine monitoring only
The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:
20 in 1,000
50 in 1,000
The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:
200 in 1,000
100 in 1,000
In addition to your regular self-checks, you will need to have the following tests:
MRI of your breast area
Mammogram
You are scheduled to have your follow-up tests every:
Second year
Six months
The cost to you each year for monitoring is:
$900 per year
$600 per year
The cost to you for surgery associated with managing your ongoing risk of cancer recurrence is:
Nothing
This does not apply
The chance you will experience ongoing pain is:
200 in 1,000
This does not apply
The chance you will experience an ongoing loss of sensitivity in your breast area is:
500 in 1,000
This does not apply
Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence:
Always
Always
Which option would you choose?
Surgery to remove unaffected breast plus monitoring
Routine monitoring only
<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

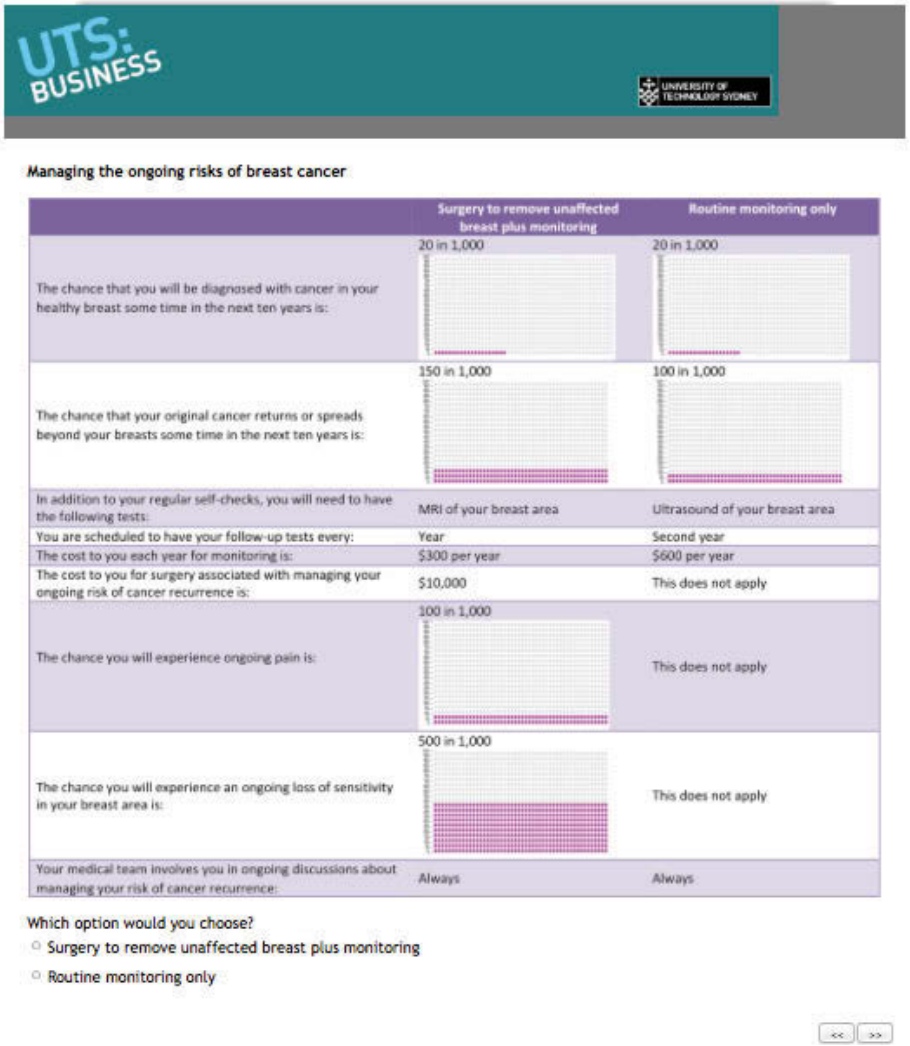
Managing the ongoing risks of breast cancer

	Surgery to remove unaffected breast plus monitoring	Routine monitoring only
The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:	3 in 1,000	20 in 1,000
The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:	200 in 1,000	150 in 1,000
In addition to your regular self-checks, you will need to have the following tests:	Ultrasound of your breast area	Mammogram
You are scheduled to have your follow-up tests every:	Second year	Year
The cost to you each year for monitoring is:	$600 per year	$600 per year
The cost to you for surgery associated with managing your ongoing risk of cancer recurrence is:	$5,000	This does not apply
The chance you will experience ongoing pain is:	300 in 1,000	This does not apply
The chance you will experience an ongoing loss of sensitivity in your breast area is:	300 in 1,000	This does not apply
Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence:	Not very often	Always

Which option would you choose?

- Surgery to remove unaffected breast plus monitoring
- Routine monitoring only

<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

Managing the ongoing risks of breast cancer

	Surgery to remove unaffected breast plus monitoring	Routine monitoring only
The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:	20 in 1,000	20 in 1,000
The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:	150 in 1,000	100 in 1,000
In addition to your regular self-checks, you will need to have the following tests:	MRI of your breast area	Ultrasound of your breast area
You are scheduled to have your follow-up tests every:	Year	Second year
The cost to you each year for monitoring is:	$300 per year	$600 per year
The cost to you for surgery associated with managing your ongoing risk of cancer recurrence is:	$10,000	This does not apply
The chance you will experience ongoing pain is:	100 in 1,000	This does not apply
The chance you will experience an ongoing loss of sensitivity in your breast area is:	500 in 1,000	This does not apply
Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence:	Always	Always

Which option would you choose?

- Surgery to remove unaffected breast plus monitoring
- Routine monitoring only

<< >>

Managing the ongoing risks of breast cancer

	Surgery to remove unaffected breast plus monitoring	Routine monitoring only
The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:	20 in 1,000	50 in 1,000
The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:	200 in 1,000	100 in 1,000
In addition to your regular self-checks, you will need to have the following tests:	Ultrasound of your breast area	MRI of your breast area
You are scheduled to have your follow-up tests every:	Second year	Six months
The cost to you each year for monitoring is:	$900 per year	$900 per year
The cost to you for surgery associated with managing your ongoing risk of cancer recurrence is:	$10,000	This does not apply
The chance you will experience ongoing pain is:	100 in 1,000	This does not apply
The chance you will experience an ongoing loss of sensitivity in your breast area is:	400 in 1,000	This does not apply
Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence:	Always	Not very often

Which option would you choose?

- Surgery to remove unaffected breast plus monitoring
- Routine monitoring only

<< >>

Managing the ongoing risks of breast cancer

	Surgery to remove unaffected breast plus monitoring	Routine monitoring only
The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:	20 in 1,000	20 in 1,000
The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:	100 in 1,000	100 in 1,000
In addition to your regular self-checks, you will need to have the following tests:	Ultrasound of your breast area	MRI of your breast area
You are scheduled to have your follow-up tests every:	Second year	Six months
The cost to you each year for monitoring is:	Nothing	$900 per year
The cost to you for surgery associated with managing your ongoing risk of cancer recurrence is:	Nothing	This does not apply
The chance you will experience ongoing pain is:	300 in 1,000	This does not apply
The chance you will experience an ongoing loss of sensitivity in your breast area is:	300 in 1,000	This does not apply
Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence:	Always	Always

Which option would you choose?

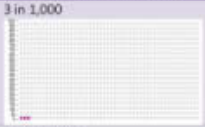
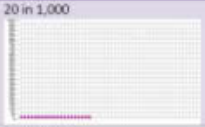
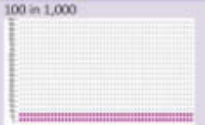
- Surgery to remove unaffected breast plus monitoring
- Routine monitoring only

<< >>

UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

Managing the ongoing risks of breast cancer

	Surgery to remove unaffected breast plus monitoring	Routine monitoring only
The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:	3 in 1,000 	20 in 1,000 
The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:	150 in 1,000 	150 in 1,000 
In addition to your regular self-checks, you will need to have the following tests:	MRI of your breast area	MRI of your breast area
You are scheduled to have your follow-up tests every:	Six months	Year
The cost to you each year for monitoring is:	$300 per year	Nothing
The cost to you for surgery associated with managing your ongoing risk of cancer recurrence is:	Nothing	This does not apply
The chance you will experience ongoing pain is:	100 in 1,000 	This does not apply
The chance you will experience an ongoing loss of sensitivity in your breast area is:	500 in 1,000 	This does not apply
Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence:	Not very often	Always

Which option would you choose?

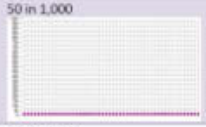
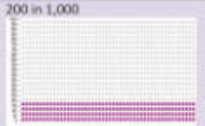
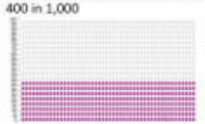
- ○ Surgery to remove unaffected breast plus monitoring
- ○ Routine monitoring only

<< >>

UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

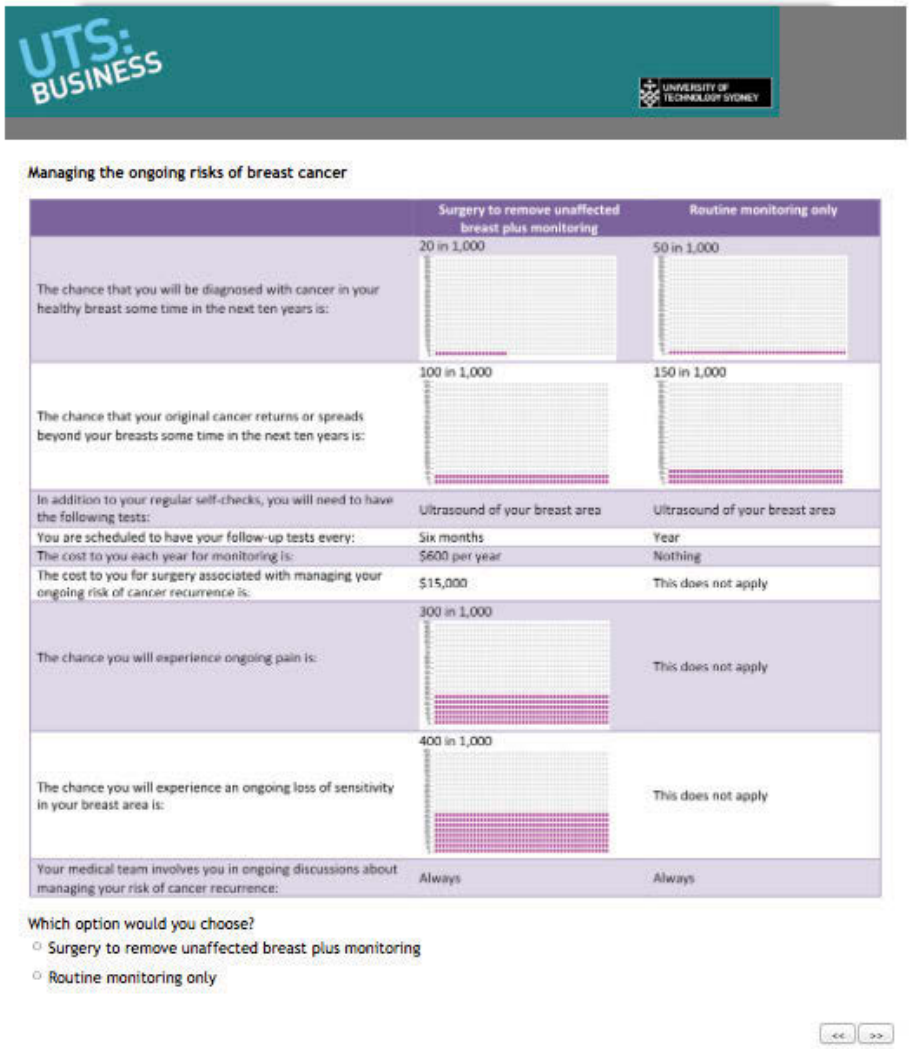
Managing the ongoing risks of breast cancer

	Surgery to remove unaffected breast plus monitoring	Routine monitoring only
The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:	3 in 1,000 	50 in 1,000 
The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:	100 in 1,000 	200 in 1,000 
In addition to your regular self-checks, you will need to have the following tests:	MRI of your breast area	Ultrasound of your breast area
You are scheduled to have your follow-up tests every:	Second year	Six months
The cost to you each year for monitoring is:	Nothing	$300 per year
The cost to you for surgery associated with managing your ongoing risk of cancer recurrence is:	$10,000	This does not apply
The chance you will experience ongoing pain is:	200 in 1,000 	This does not apply
The chance you will experience an ongoing loss of sensitivity in your breast area is:	400 in 1,000 	This does not apply
Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence:	Not very often	Always

Which option would you choose?

- ○ Surgery to remove unaffected breast plus monitoring
- ○ Routine monitoring only

<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

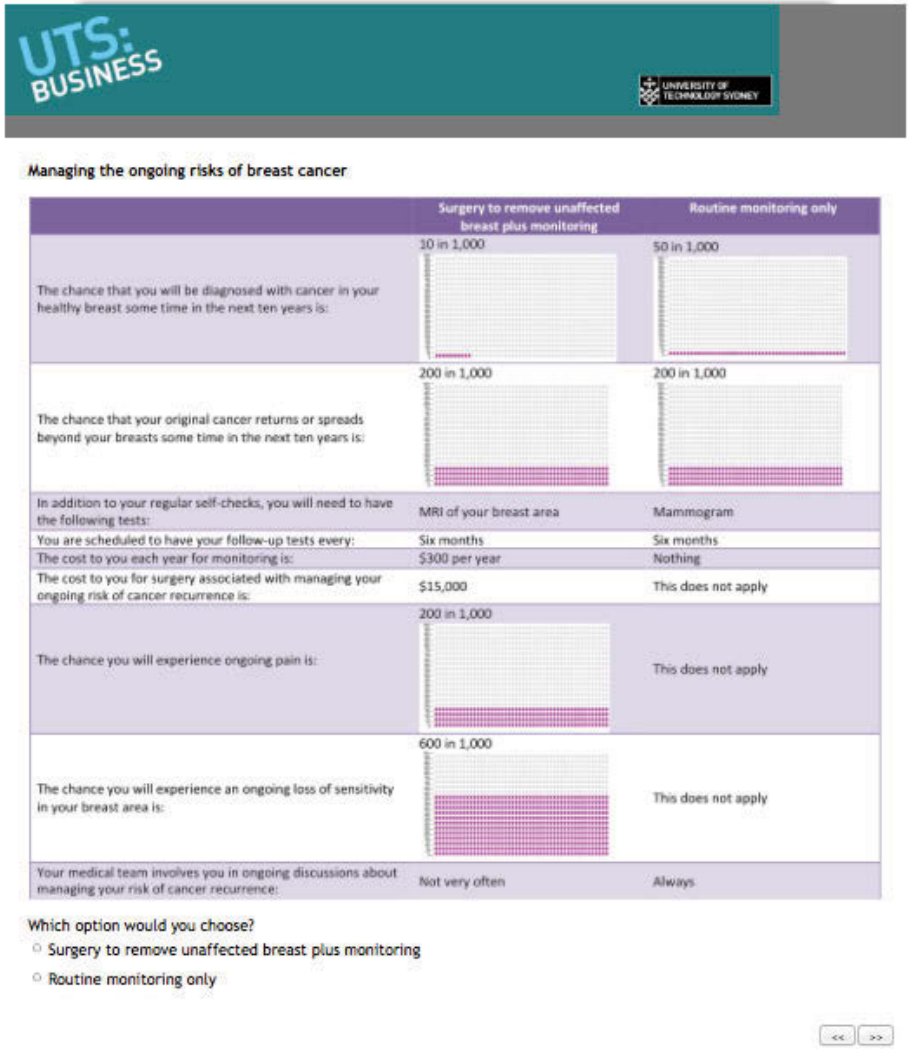
Managing the ongoing risks of breast cancer

	Surgery to remove unaffected breast plus monitoring	Routine monitoring only
The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:	20 in 1,000	50 in 1,000
The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:	100 in 1,000	150 in 1,000
In addition to your regular self-checks, you will need to have the following tests:	Ultrasound of your breast area	Ultrasound of your breast area
You are scheduled to have your follow-up tests every:	Six months	Year
The cost to you each year for monitoring is:	$600 per year	Nothing
The cost to you for surgery associated with managing your ongoing risk of cancer recurrence is:	$15,000	This does not apply
The chance you will experience ongoing pain is:	300 in 1,000	This does not apply
The chance you will experience an ongoing loss of sensitivity in your breast area is:	400 in 1,000	This does not apply
Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence:	Always	Always

Which option would you choose?

- Surgery to remove unaffected breast plus monitoring
- Routine monitoring only

<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

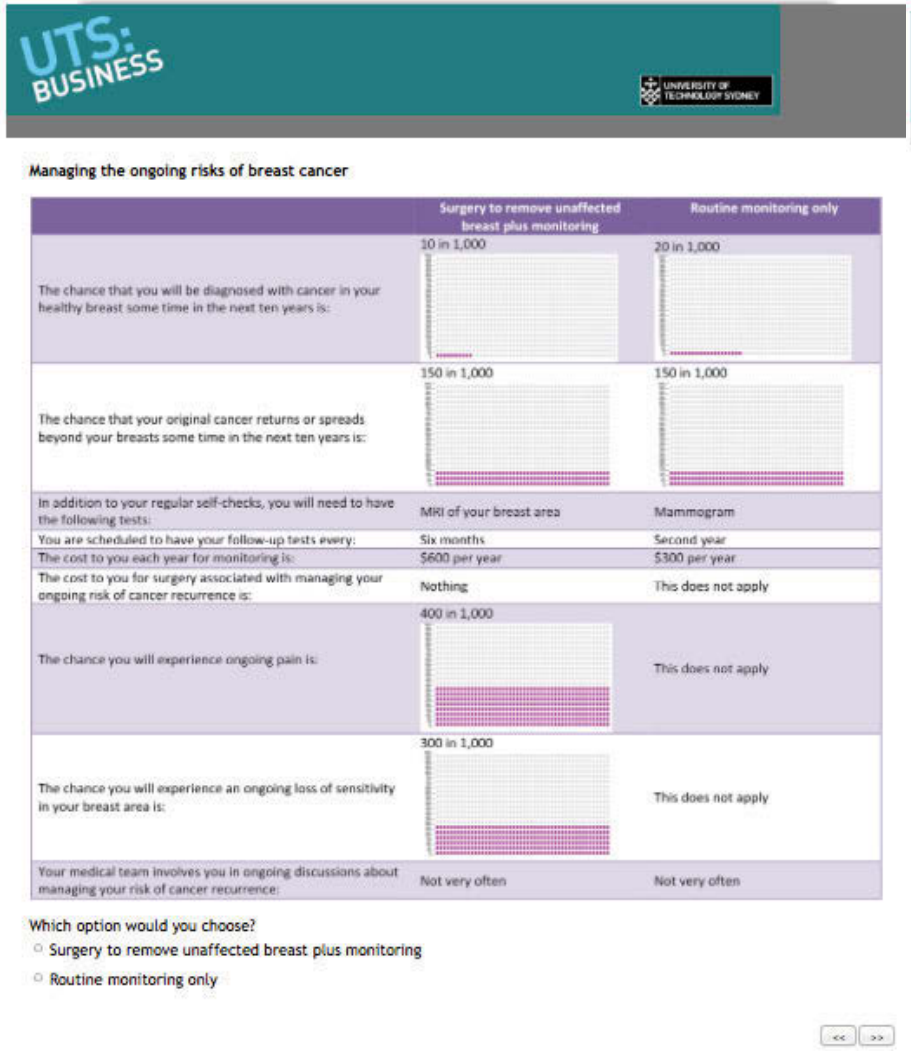
Managing the ongoing risks of breast cancer

	Surgery to remove unaffected breast plus monitoring	Routine monitoring only
The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:	10 in 1,000	50 in 1,000
The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:	200 in 1,000	200 in 1,000
In addition to your regular self-checks, you will need to have the following tests:	MRI of your breast area	Mammogram
You are scheduled to have your follow-up tests every:	Six months	Six months
The cost to you each year for monitoring is:	$300 per year	Nothing
The cost to you for surgery associated with managing your ongoing risk of cancer recurrence is:	$15,000	This does not apply
The chance you will experience ongoing pain is:	200 in 1,000	This does not apply
The chance you will experience an ongoing loss of sensitivity in your breast area is:	600 in 1,000	This does not apply
Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence:	Not very often	Always

Which option would you choose?

- Surgery to remove unaffected breast plus monitoring
- Routine monitoring only

<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

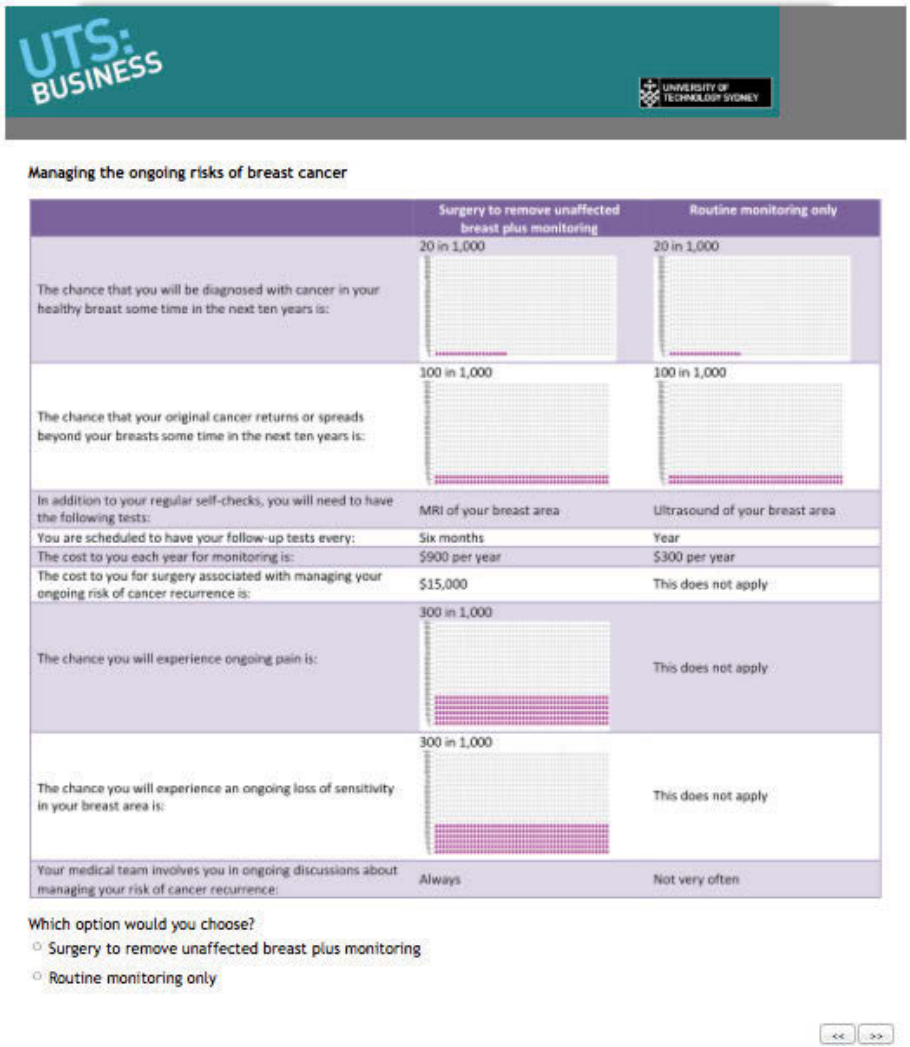
Managing the ongoing risks of breast cancer

	Surgery to remove unaffected breast plus monitoring	Routine monitoring only
The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:	10 in 1,000	20 in 1,000
The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:	150 in 1,000	150 in 1,000
In addition to your regular self-checks, you will need to have the following tests:	MRI of your breast area	Mammogram
You are scheduled to have your follow-up tests every:	Six months	Second year
The cost to you each year for monitoring is:	$600 per year	$300 per year
The cost to you for surgery associated with managing your ongoing risk of cancer recurrence is:	Nothing	This does not apply
The chance you will experience ongoing pain is:	400 in 1,000	This does not apply
The chance you will experience an ongoing loss of sensitivity in your breast area is:	300 in 1,000	This does not apply
Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence:	Not very often	Not very often

Which option would you choose?

- Surgery to remove unaffected breast plus monitoring
- Routine monitoring only

<< >>



UTS: BUSINESS

UNIVERSITY OF TECHNOLOGY SYDNEY

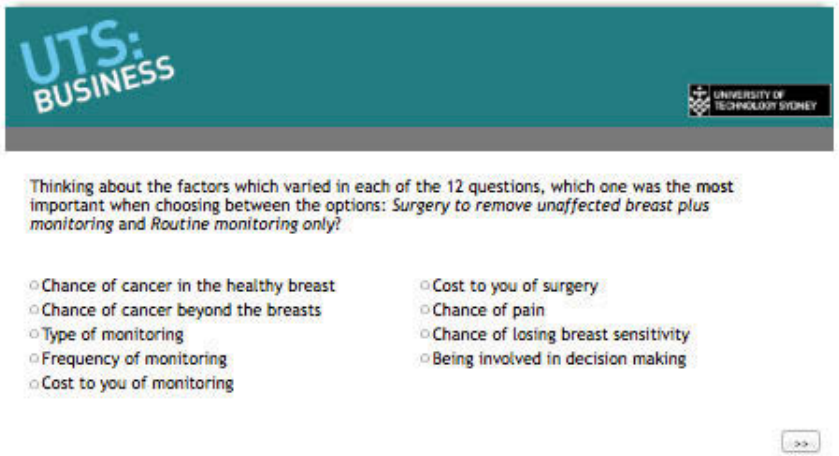
Managing the ongoing risks of breast cancer

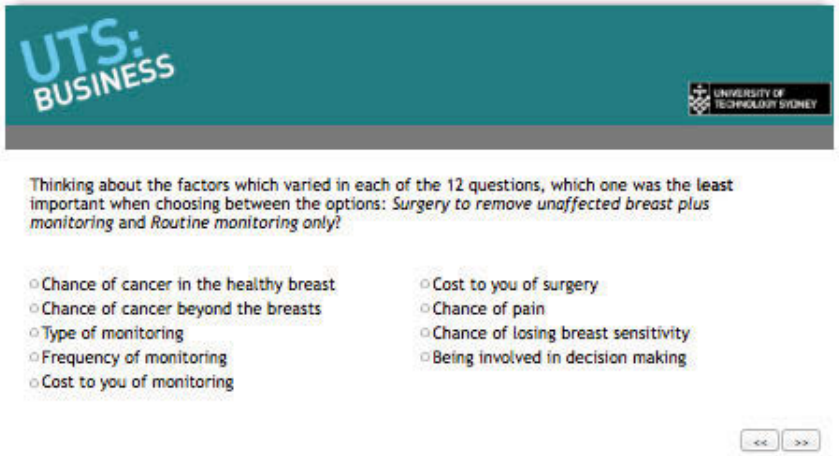
	Surgery to remove unaffected breast plus monitoring	Routine monitoring only
The chance that you will be diagnosed with cancer in your healthy breast some time in the next ten years is:	20 in 1,000	20 in 1,000
The chance that your original cancer returns or spreads beyond your breasts some time in the next ten years is:	100 in 1,000	100 in 1,000
In addition to your regular self-checks, you will need to have the following tests:	MRI of your breast area	Ultrasound of your breast area
You are scheduled to have your follow-up tests every:	Six months	Year
The cost to you each year for monitoring is:	$900 per year	$300 per year
The cost to you for surgery associated with managing your ongoing risk of cancer recurrence is:	$15,000	This does not apply
The chance you will experience ongoing pain is:	300 in 1,000	This does not apply
The chance you will experience an ongoing loss of sensitivity in your breast area is:	300 in 1,000	This does not apply
Your medical team involves you in ongoing discussions about managing your risk of cancer recurrence:	Always	Not very often

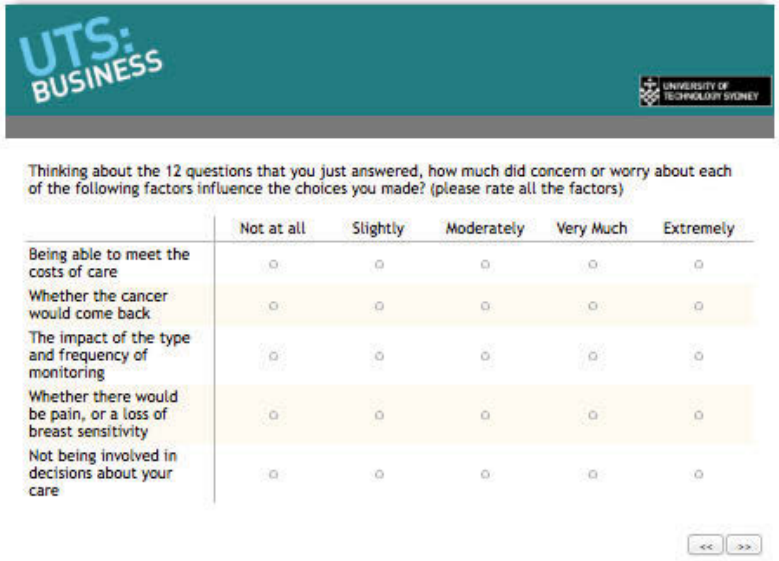
Which option would you choose?

- Surgery to remove unaffected breast plus monitoring
- Routine monitoring only

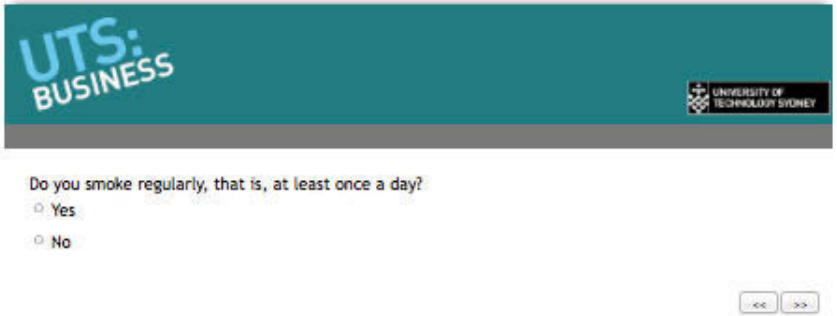
<< >>

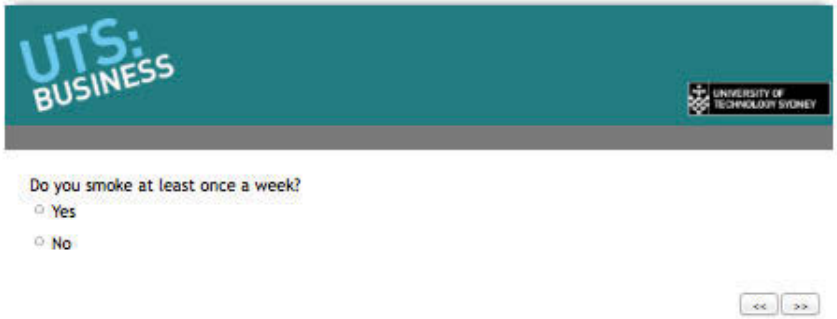

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Thinking about the factors which varied in each of the 12 questions, which one was the most important when choosing between the options: Surgery to remove unaffected breast plus monitoring and Routine monitoring only?
Chance of cancer in the healthy breast
Chance of cancer beyond the breasts
Type of monitoring
Frequency of monitoring
Cost to you of monitoring
Cost to you of surgery
Chance of pain
Chance of losing breast sensitivity
Being involved in decision making
>>

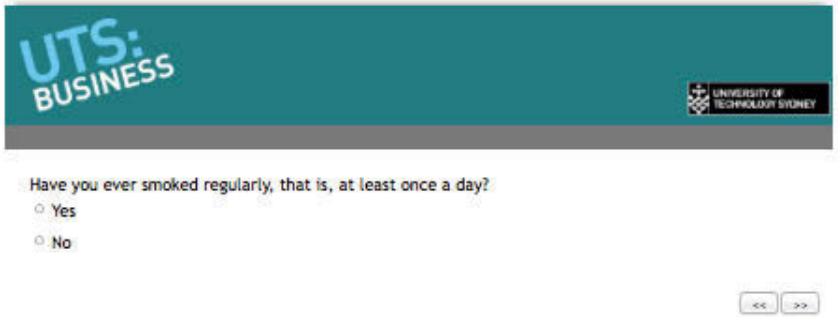

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Thinking about the factors which varied in each of the 12 questions, which one was the least important when choosing between the options: Surgery to remove unaffected breast plus monitoring and Routine monitoring only?
Chance of cancer in the healthy breast
Chance of cancer beyond the breasts
Type of monitoring
Frequency of monitoring
Cost to you of monitoring
Cost to you of surgery
Chance of pain
Chance of losing breast sensitivity
Being involved in decision making
<< >>

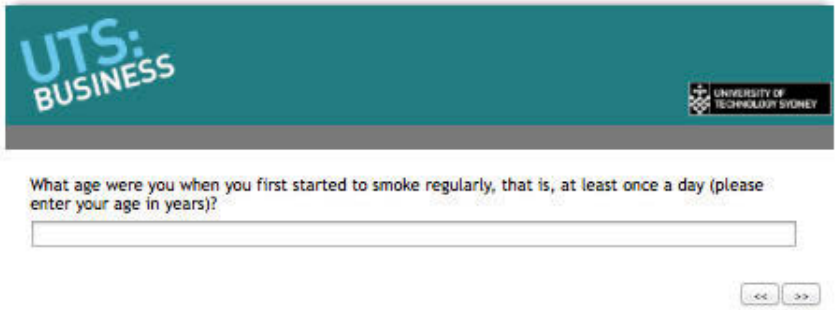

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Thinking about the 12 questions that you just answered, how much did concern or worry about each of the following factors influence the choices you made? (please rate all the factors)
Not at all
Slightly
Moderately
Very Much
Extremely
Being able to meet the costs of care
Whether the cancer would come back
The impact of the type and frequency of monitoring
Whether there would be pain, or a loss of breast sensitivity
Not being involved in decisions about your care
<< >>

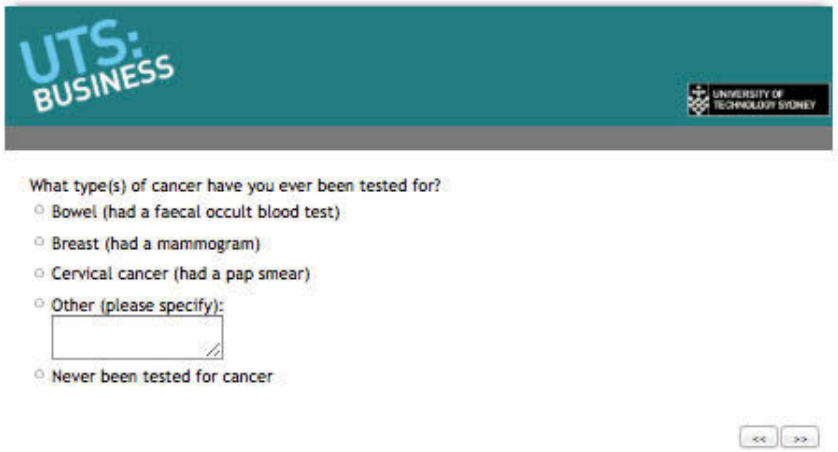

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Do you currently smoke?
Yes
No
<<
>>

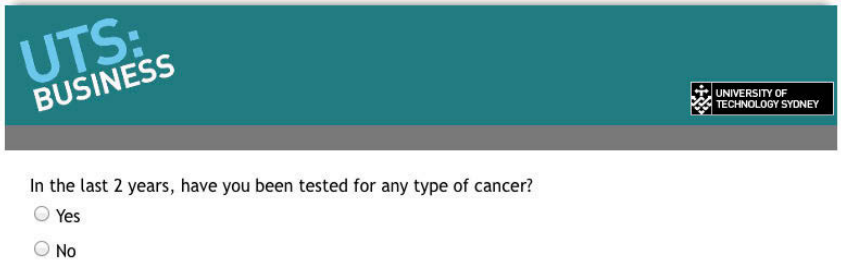

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Do you smoke regularly, that is, at least once a day?
Yes
No
<<
>>

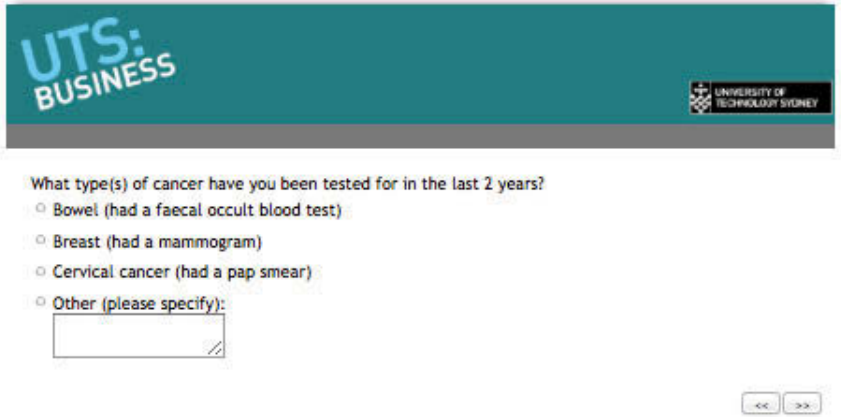

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Do you smoke at least once a week?
Yes
No
<<
>>

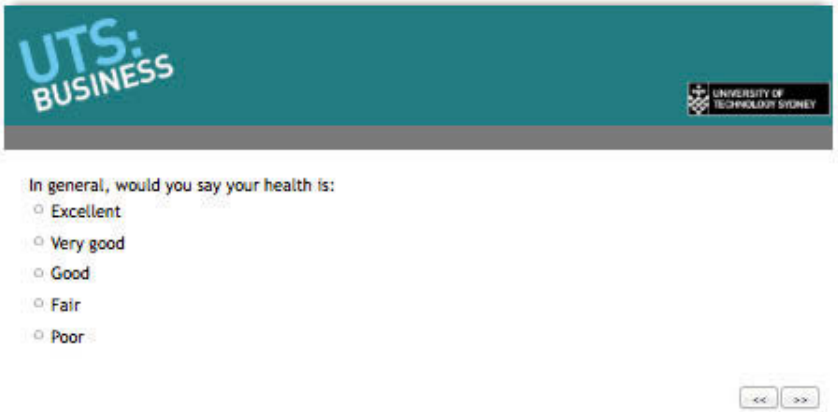

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Have you ever smoked regularly, that is, at least once a day?
Yes
No
<<
>>

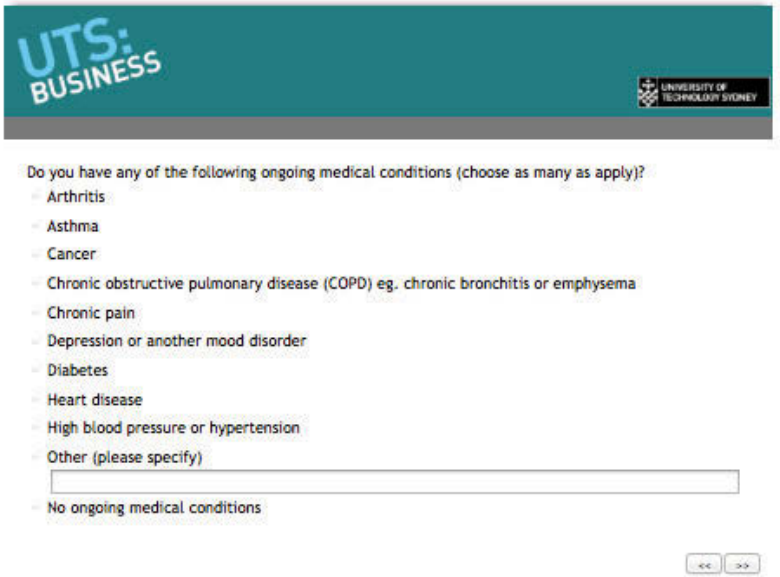

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
What age were you when you first started to smoke regularly, that is, at least once a day (please enter your age in years)?
<<
>>

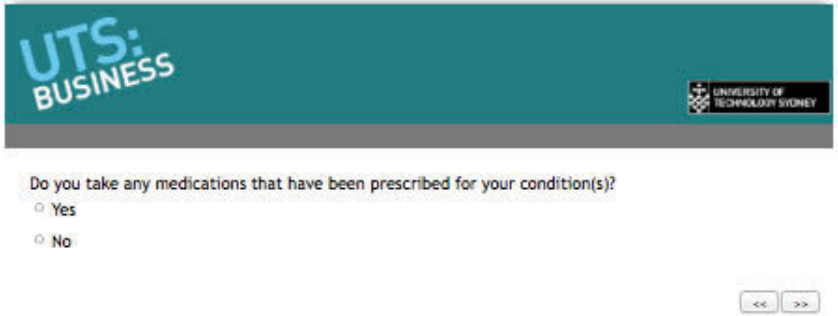

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
What type(s) of cancer have you ever been tested for?
Bowel (had a faecal occult blood test)
Breast (had a mammogram)
Cervical cancer (had a pap smear)
Other (please specify):
Never been tested for cancer
<<
>>

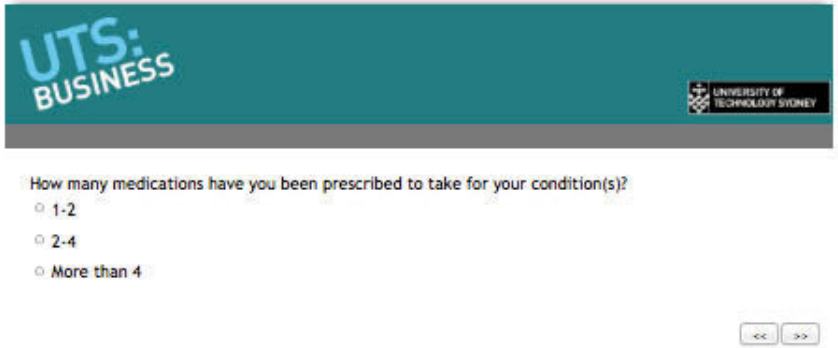

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
In the last 2 years, have you been tested for any type of cancer?
Yes
No

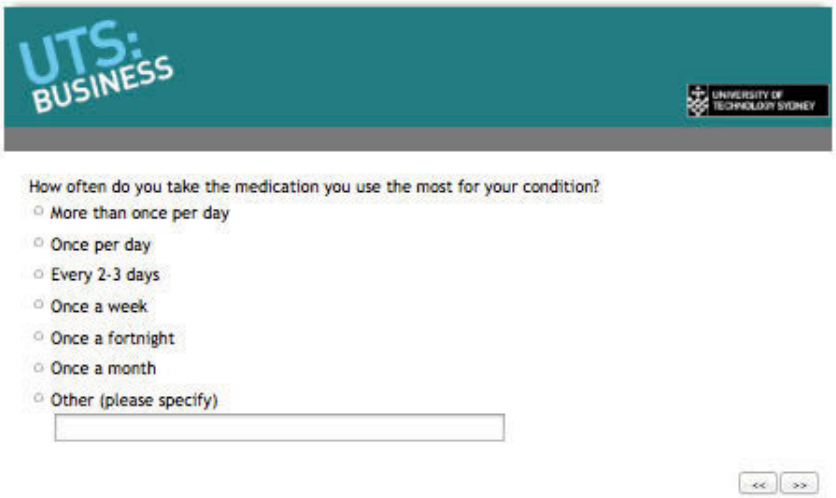

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
What type(s) of cancer have you been tested for in the last 2 years?
Bowel (had a faecal occult blood test)
Breast (had a mammogram)
Cervical cancer (had a pap smear)
Other (please specify):
<<
>>

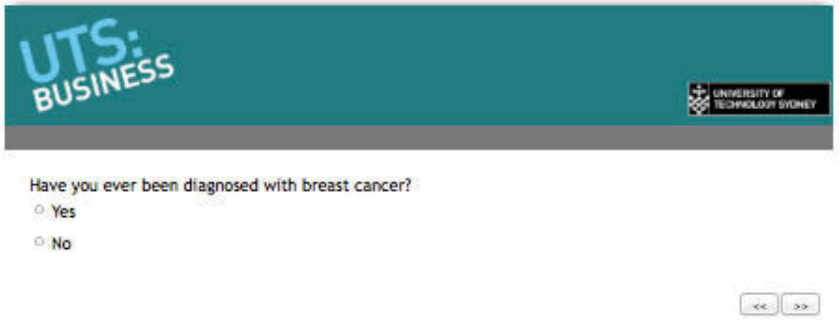

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
In general, would you say your health is:
Excellent
Very good
Good
Fair
Poor
<< >>

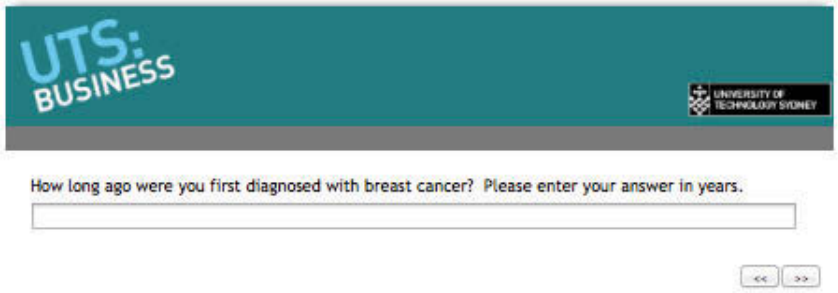

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Do you have any of the following ongoing medical conditions (choose as many as apply)?
Arthritis
Asthma
Cancer
Chronic obstructive pulmonary disease (COPD) eg. chronic bronchitis or emphysema
Chronic pain
Depression or another mood disorder
Diabetes
Heart disease
High blood pressure or hypertension
Other (please specify)
No ongoing medical conditions
<< >>

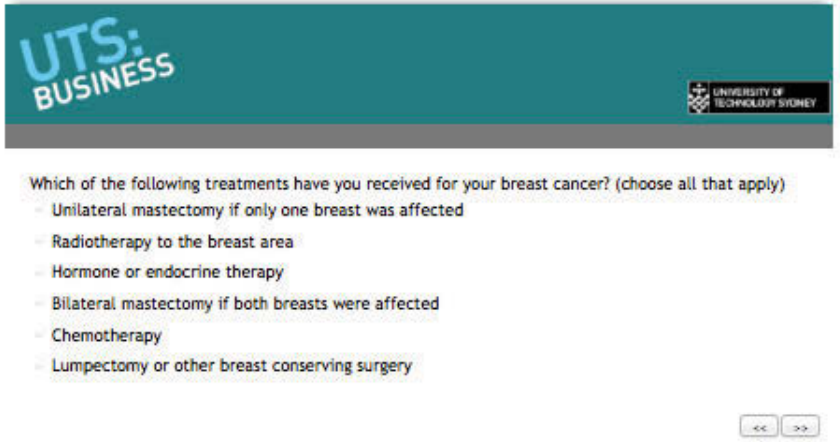

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Do you take any medications that have been prescribed for your condition(s)?
Yes
No
<< >>

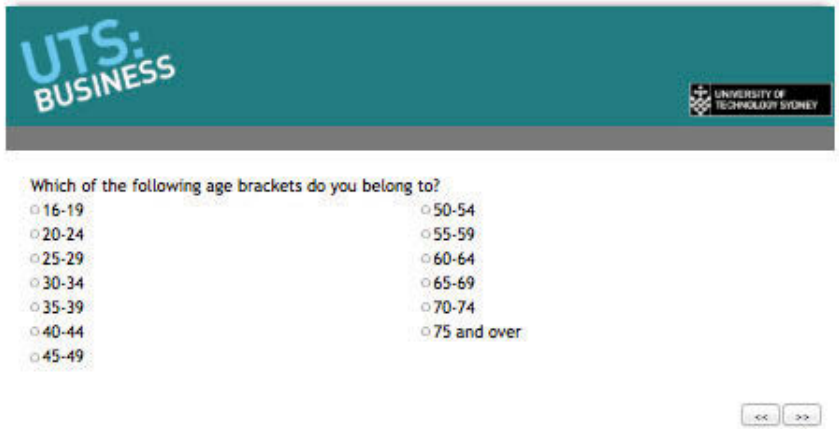

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
How many medications have you been prescribed to take for your condition(s)?
1-2
2-4
More than 4
<< >>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
How often do you take the medication you use the most for your condition?
More than once per day
Once per day
Every 2-3 days
Once a week
Once a fortnight
Once a month
Other (please specify)
<<
>>

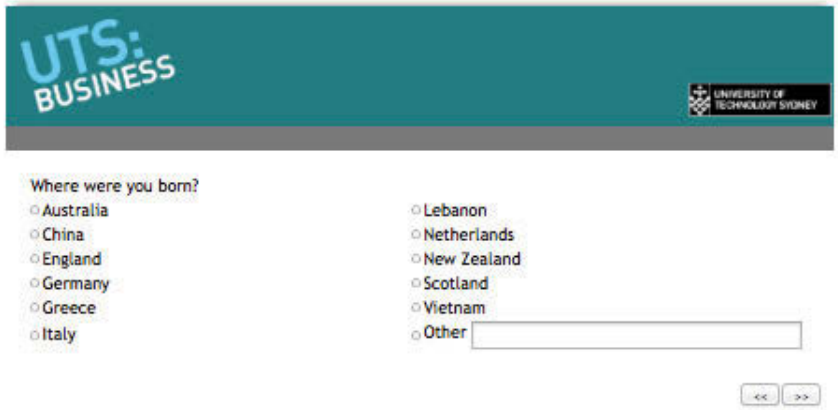

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Have you ever been diagnosed with breast cancer?
Yes
No
<<
>>

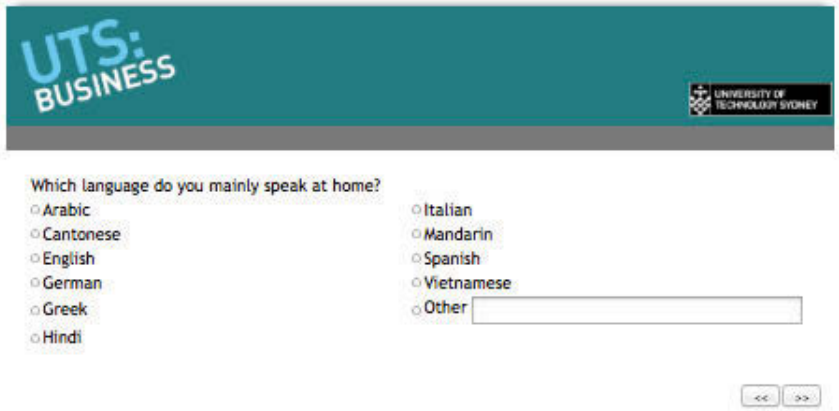

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
How long ago were you first diagnosed with breast cancer? Please enter your answer in years.
<<
>>

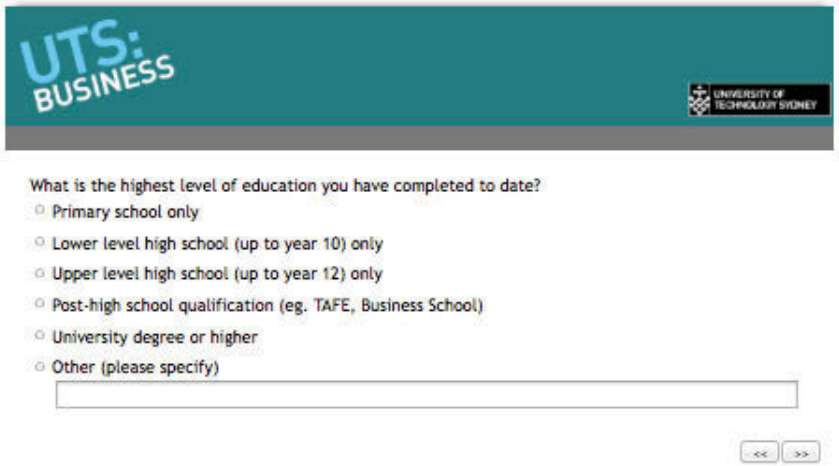

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Which of the following treatments have you received for your breast cancer? (choose all that apply)
Unilateral mastectomy if only one breast was affected
Radiotherapy to the breast area
Hormone or endocrine therapy
Bilateral mastectomy if both breasts were affected
Chemotherapy
Lumpectomy or other breast conserving surgery
<<
>>

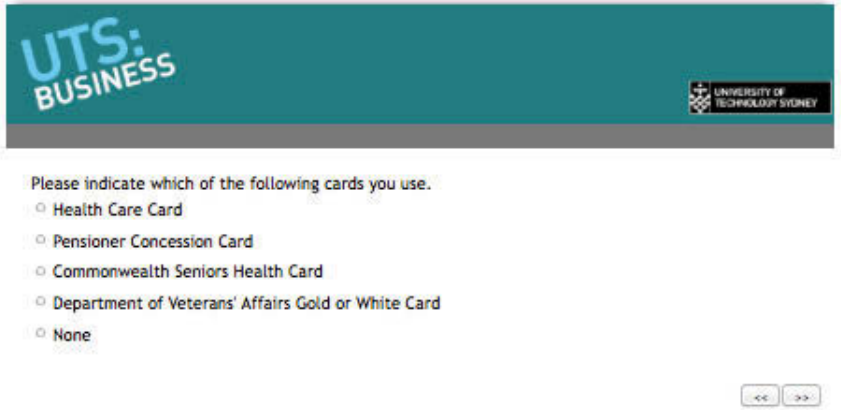

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Which of the following age brackets do you belong to?
16-19
20-24
25-29
30-34
35-39
40-44
45-49
50-54
55-59
60-64
65-69
70-74
75 and over
<<
>>

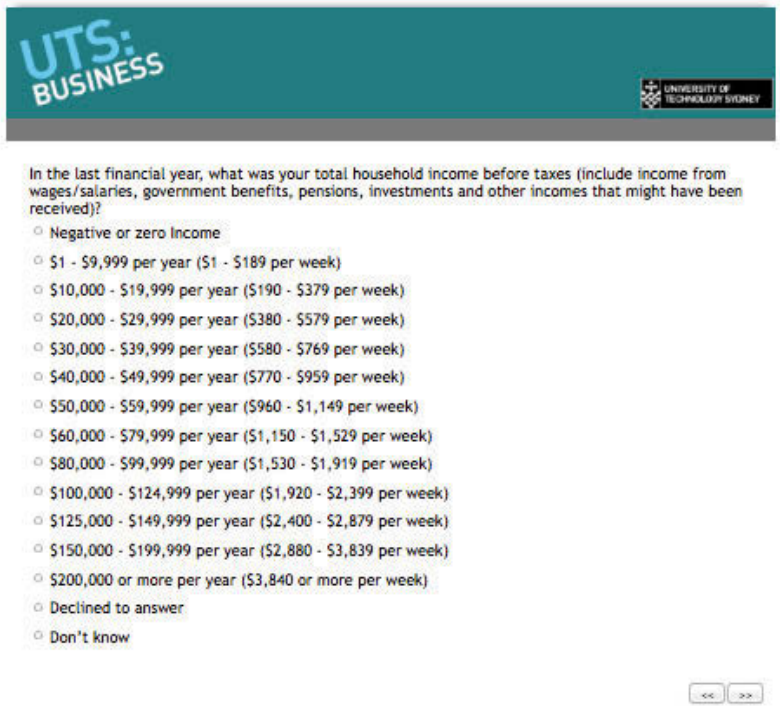

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Are you of Aboriginal and/or Torres Strait Islander origin?
Yes
No
<<
>>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Where were you born?
Australia
China
England
Germany
Greece
Italy
Lebanon
Netherlands
New Zealand
Scotland
Vietnam
Other
<<
>>

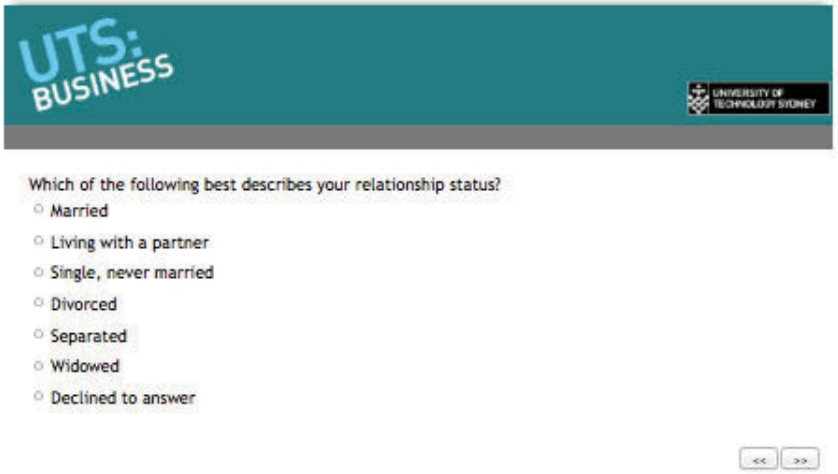

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Which language do you mainly speak at home?
Arabic
Cantonese
English
German
Greek
Hindi
Italian
Mandarin
Spanish
Vietnamese
Other
<<
>>

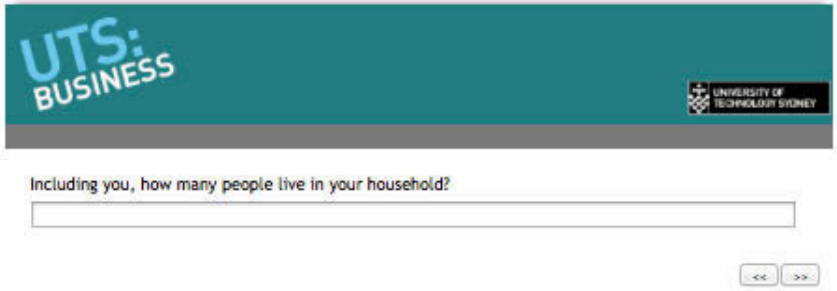

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
What is the highest level of education you have completed to date?
Primary school only
Lower level high school (up to year 10) only
Upper level high school (up to year 12) only
Post-high school qualification (eg. TAFE, Business School)
University degree or higher
Other (please specify)
<< >>

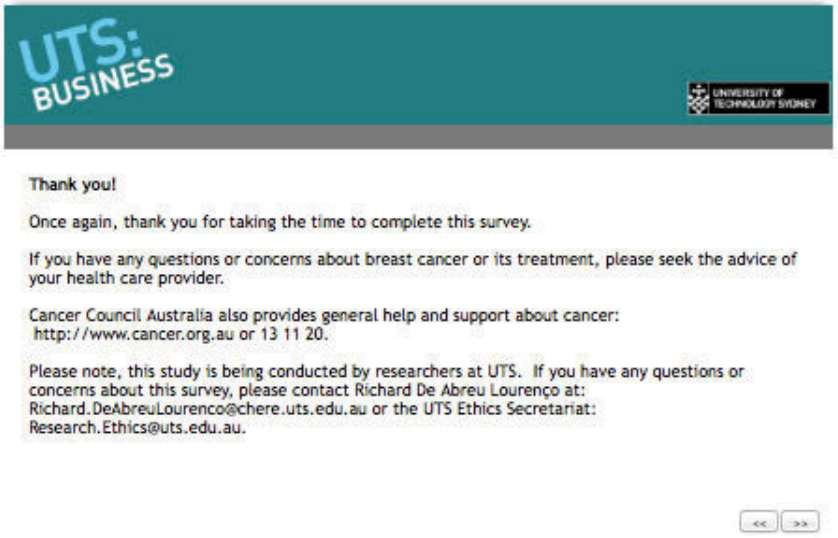

UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Please indicate which of the following cards you use.
Health Care Card
Pensioner Concession Card
Commonwealth Seniors Health Card
Department of Veterans' Affairs Gold or White Card
None
<< >>


UTS:
BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
In the last financial year, what was your total household income before taxes (include income from wages/salaries, government benefits, pensions, investments and other incomes that might have been received)?
Negative or zero income
$1 - $9,999 per year ($1 - $189 per week)
$10,000 - $19,999 per year ($190 - $379 per week)
$20,000 - $29,999 per year ($380 - $579 per week)
$30,000 - $39,999 per year ($580 - $769 per week)
$40,000 - $49,999 per year ($770 - $959 per week)
$50,000 - $59,999 per year ($960 - $1,149 per week)
$60,000 - $79,999 per year ($1,150 - $1,529 per week)
$80,000 - $99,999 per year ($1,530 - $1,919 per week)
$100,000 - $124,999 per year ($1,920 - $2,399 per week)
$125,000 - $149,999 per year ($2,400 - $2,879 per week)
$150,000 - $199,999 per year ($2,880 - $3,839 per week)
$200,000 or more per year ($3,840 or more per week)
Declined to answer
Don't know
<< >>


UTS:
BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
What is the post-code of the area in which you live?
Enter your postcode
Declined to answer
<< >>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Which of the following best describes your relationship status?
Married
Living with a partner
Single, never married
Divorced
Separated
Widowed
Declined to answer
<< >>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Including you, how many people live in your household?
<< >>


UTS: BUSINESS
UNIVERSITY OF TECHNOLOGY SYDNEY
Thank you!
Once again, thank you for taking the time to complete this survey.
If you have any questions or concerns about breast cancer or its treatment, please seek the advice of your health care provider.
Cancer Council Australia also provides general help and support about cancer: http://www.cancer.org.au or 13 11 20.
Please note, this study is being conducted by researchers at UTS. If you have any questions or concerns about this survey, please contact Richard De Abreu Lourenço at: Richard.DeAbreuLourenco@chere.uts.edu.au or the UTS Ethics Secretariat: Research.Ethics@uts.edu.au.
<< >>

References

1. Australian Institute of Health and Welfare (AIHW). Health Status 2017.
2. Mooney G. Beyond health outcomes: the benefits of health care. Health Care Anal 1998;6:99-105.
3. Monincx W, Birnie E. Fetal monitoring at home in high-risk pregnancy. An integrated clinical and economic evaluation: University of Amsterdam; 2000.
4. de Ruyter K, Wetzels M. On the complex nature of patient evaluations of general practice service. J Econ Psychol 1998;19:565-90.
5. Opmeer BC, De Borgie CAJM, Mol BWJ, Bossuyt PMM. Assessing preferences regarding healthcare interventions that involve non-health outcomes: An overview of clinical studies. Patient 2010;3:1-10.
6. Gandjour A. Is subjective well-being a useful parameter for allocating resources among public interventions? Health Care Anal 2001;9:437-47.
7. Lippert-Rasmussen K, Lauridsen S. Justice and the allocation of healthcare resources: should indirect, non-health effects count? Med Health Care Philos 2010;13:237-46.
8. Lorgelly PK, Lawson KD, Fenwick EA, Briggs AH. Outcome measurement in economic evaluations of public health interventions: a role for the capability approach? Int J Environ Res Public Health 2010;7:2274-89.
9. McIntosh E, Cairns J. A framework for the economic evaluation of telemedicine. J Telemed Telecare 1997;3:132-9.
10. Alayli-Goebbels AFG, Dellaert BGC, Knox SA, et al. Consumer preferences for health and nonhealth outcomes of health promotion: Results from a discrete choice experiment. Value Health 2013;16:114-23.
11. Mooney G. QALYs: Are they enough? A health economist's perspective. J Med Ethics 1989;15:148-52.
12. McGuire A, Henderson J, Mooney G. The Economics of Health Care. An Introductory Text. London: Routledge & Kegan Paul; 1988.
13. Payne K, McAllister M, Davies LM. Valuing the economic benefits of complex interventions: When maximising health is not sufficient. Health Economics (United Kingdom) 2013;22:258-71.
14. Benger J. Think about the care in healthcare. BMJ 2014;348.
15. Kitson A, Marshall A, Bassett K, Zeitz K. What are the core elements of patient-centred care? A narrative review and synthesis of the literature from health policy, medicine and nursing. J Adv Nurs 2013;69:4-15.
16. McMillan SS, Kendall E, Sav A, et al. Patient-centered approaches to health care: a systematic review of randomized controlled trials. Med Care Res Rev 2013;70:567-96.
17. Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. Med Care Res Rev 2013;70:351-79.
18. Benning T, Dellaert B, Severens J, Dirksen C. The effect of presenting information about invasive follow-up testing on individuals' noninvasive colorectal cancer screening participation decision: results from a discrete choice experiment. Value Health 2014;17:578-87.
19. Kilambi V, Reed Johnson F, Gonzalez J, Mohamed A. Valuations of genetic information for treatable conditions: the case of colorectal cancer screening. Value Health 2014;17:838-45.
20. Protiere C, Donaldson C, Luchini S, Moatti JP, Shackley P. The impact of information on non-health attributes on willingness to pay for multiple health care programmes. Soc Sci Med 2004;58:1257-69.

21. Hall J, Fiebig DG, King MT, Hossain I, Louviere JJ. What influences participation in genetic carrier testing? Results from a discrete choice experiment. J Health Econ 2006;25:520-37.
22. Centre for Drug Evaluation. Guidelines of Methodological Standards for Pharmacoeconomic Evaluations in Taiwan. 2006.
23. Institute for Quality and Efficiency in Health Care. General Methods. Cologne, Germany2015.
24. Pharmaceutical Benefits Advisory Committee. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee. In: Health Do, ed. Canberra: Commonwealth of Australia; 2013.
25. Teerawattananon Y, Thavorncharoensap M, Mohara A. Measurement of clinical effects. J Med Assoc Thai 2008;91 Suppl 2:S38-42.
26. Phantipa P. Measurement of clinical-effect: utility. J Med Assoc Thai 2008;91 Suppl 2:S43-52.
27. PHARMAC. Prescription for PharmacoEconomic Analysis: Methods for Cost-Utility Analysis. In: Agency PM, ed.: New Zealand Government; 2012.
28. Canadian Agency for Drugs and Technologies in Health. Guidelines for the economic evaluation of health technologies: Canada. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.
29. van der Meijden C, Grahlmann C. Dutch Assessment Procedures for the Reimbursement of Outpatient Medicines. In: Ministry of Health WaSCvZ, ed. The Netherlands2010.
30. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal. NICE; 2013.
31. Drummond F, O'Brien B, Stoddart G, Torrrance G. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 1997.
32. Brazier J, Tsuchiya A. Improving cross-sector comparisons: going beyond the health-related QALY. Applied health economics and health policy 2015;13:557-65.
33. Coast J, Flynn TN, Natarajan L, et al. Valuing the ICECAP capability index for older people. Soc Sci Med 2008;67:874-82.
34. Coast J, Smith R, Lorgelly P. Should the capability approach be applied in health economics? Health Econ 2008;17:667-70.
35. Kahneman D, Krueger A. Developmens in the measurement of subjective well-being. J Econ Perspect 2006;20:3-24.
36. Torrance G, Boyle M, Horwood S. Application of Multi-Attribute Utility Theory to Measure Social Preferences for Health States. Research and Working Paper Series. Hamilton Ontario: McMaster University; 1982.
37. Richardson J, Iezzi A, Khan MA. Why do multi-attribute utility instruments produce different utilities: the relative importance of the descriptive systems, scale and 'micro-utility' effects. Qual Life Res 2015;24:2045-53.
38. Kauf TL, Roskell N, Shearer A, et al. A predictive model of health state utilities for HIV patients in the modern era of highly active antiretroviral therapy. Value Health 2008;11:1144-53.
39. Brennan VK, Dixon S. Incorporating process utility into quality adjusted life years: A systematic review of empirical studies. Pharmacoeconomics 2013;31:677-91.
40. Bansback N, Brazier J, Tsuchiya A, Anis A. Using a discrete choice experiment to estimate societal health state utility values. J Health Econ 2012;31:306-18.
41. Norman R, Viney R, Brazier J, et al. Valuing SF-6D Health States Using a Discrete Choice Experiment. Med Decis Making 2013.
42. Torrance G, Thomas W, Sackett D. A utility maximization model for the evaluation of health care programs. Health Serv Res 1972;Summer:118-33.
43. Torrance G. Measurement of health state utilities for economic appraisal. J Health Econ 1986;5:1-30.

44. Clark MD, Determann D, Petrou S, Moro D, de Bekker-Grob EW. Discrete choice experiments in health economics: a review of the literature. Pharmacoeconomics 2014;32:883-902.
45. de Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. Health Econ 2012;21:145-72.
46. Lancaster K. A New Approach to Consumer Theory. The Journal of Political Economy 1966;74:132-57.
47. Johnson F, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Practices Research Task Force. Value Health 2013;16:3-13.
48. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics 2008;26:661-77.
49. Hall J, Viney R, Haas M, Louviere J. Using stated preference discrete choice modeling to evaluate health care programs. J Bus Res 2004;57:1026-32.
50. Amaya-Amaya M, Gerard K, Ryan M. Discrete choice experiments in a nutshell. In: Ryan M, Gerard K, Amaya-Amaya M, eds. Using Discrete Choice Experiments to Value Health and Health Care. The Netherlands: Springer; 2008:13-46.
51. Robinson A, Spencer A, Moffatt P. A Framework for estimating health state utility values within a discrete choice experiment: Modeling risky choices. Med Decis Making 2015 2015;35:341–50.
52. Adamowicz A, Louviere J, Williams M. Combining revealed and stated preference methods for valuing environmental amenities. J Environ Econ Manage 1994;26:271-92.
53. Diener A, O'Brien B, Gafni A. Health care contingent valuation studies: a review and classification of the literature. Health Econ 1998;7:313-26.
54. O'Brien B, Viramontes J. A valid and reliable measure of health state preference? Med Decis Making 1994;14:289-97.
55. Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. Applied health economics and health policy 2003;2:55-64.
56. Greene W. Discrete Choice Modeling. In: Mills T, Patterson K, eds. The Handbook of Econometrics. London: Palgrave; 2008.
57. Brouwer W, Culyer A, van Exel N, Rutten F. Welfarism vs. extra-welfarism. J Health Econ 2008;27:325 – 38.
58. Culyer A. The normative economics of health care finance and provision. Oxford Rev Econ Pol 1989;5:34-58.
59. Hall J. Health, health care and social welfare. The Australian Economic Review 2001;34:320–31.
60. Coast J, Smith RD, Lorgelly P. Welfarism, extra-welfarism and capability: the spread of ideas in health economics. Soc Sci Med 2008;67:1190-8.
61. Birch S, Donaldson C. Valuing the benefits and costs of health care programmes: where's the 'extra' in extra-welfarism? Soc Sci Med 2003;56:1121–33.
62. Johnson ES, Sullivan SD, Mozaffari E, Langley PC, Bodsworth NJ. A utility assessment of oral and intravenous ganciclovir for the maintenance treatment of AIDS-related cytomegalovirus retinitis. Pharmacoeconomics 1996;10:623-9.
63. Osborne RH, Dalton A, Hertel J, Schrover R, Kingsford Smith D. Health-related quality of life advantage of long-acting injectable antipsychotic treatment for schizophrenia: a time-trade off study. Health & Quality of Life Outcomes 2012;10:9.
64. Osborne RH, De Abreu Lourenco R, Dalton A, et al. Quality of life related to oral versus subcutaneous iron chelation: a time trade-off study. Value Health 2007;10:451-6.

65. Birch S, Melnikow J, Kuppermann M. Conservative versus aggressive follow up of mildly abnormal Pap smears: Testing for process utility. Health Econ 2003;12:879-84.
66. Swan J, Fryback D, Lawrence W, Sainfort F, Hagenauer M, Heisey D. A time-tradeoff method for cost-effectiveness models applied to radiology. Med Decis Making 2000;20:79-88.
67. Swan JS, Lawrence WF, Roy J. Process utility in breast biopsy. Med Decis Making 2006;26:347-59.
68. Swan JS, Sainfort F, Lawrence WF, Kuruchittham V, Kongnakorn T, Heisey DM. Process utility for imaging in cerebrovascular disease. Acad Radiol 2003;10:266-74.
69. Ryan M. Valuing psychological factors in the provision of assisted reproductive techniques using the economic instrument of willingness to pay. J Econ Psychol 1998;19:179-204.
70. Dixon S, Shackley P. Estimating the benefits of community water fluoridation using the willingness-to-pay technique: results of a pilot study. Community Dent Oral Epidemiol 1999;27:124-9.
71. Borghi J, Jan S. Measuring the benefits of health promotion programmes: Application of the contingent valuation method. Health Policy 2008;87:235-48.
72. Neumann PJ, Cohen JT, Hammitt JK, et al. Willingness-to-pay for predictive tests with no immediate treatment implications: A survey of US residents. Health Econ 2012;21:238-51.
73. Lakdawalla DN, Romley JA, Sanchez Y, Maclean JR, Penrod JR, Philipson T. How cancer patients value hope and the implications for cost-effectiveness assessments of high-cost cancer therapies. Health Aff (Millwood) 2012;31:676-82.
74. Cookson R. Willingness to pay methods in health care: a sceptical view. Health Econ 2003;12:891-4.
75. Dolan P, Kahneman D. Interpretations of utility and their implications for the valuation of health. The Economic Journal 2008;118:215-34.
76. Kahneman D, Sugden R. Experienced utility as a standard of policy evaluation. Environmental & Resource Economics 2005;32:161-81.
77. Cookson R. QALYs and the capability approach. Health Econ 2005;14:817-29.
78. Baker R, Robinson A. Responses to standard gambles: are preferences 'well constructed'? Health Econ 2004;13:37-48.
79. Kahneman D, Wakker P, Sarin R. Back to Bentham? Explorations of experiened utility. The Quarterly Journal of Economics 1997;May.
80. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. Econometrica 1979;47:263-91.
81. Kahneman D. Thinking, Fast and Slow. Australia: Penguin Books; 2011.
82. de Botton A. Religion for Atheists. London: Penguin Books; 2012.
83. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science 1981;211:453-8.
84. Kragt M, Bennett J. Designing choice experiments to test for anchoring and framing effects. Canberra: ANU; 2008.
85. Woodhead E, Lynch E, Edelstein B. Decisional strategy determines whether frame influences treatment preferences for medical decisions. Psychol Aging 2011;26:285-94.
86. Froberg D, Kane R. Methodology for measuring health-state preferences-III: population and context effects. J Clin Epidemiol 1989;42:585-92.
87. Hallahan K. Seven models of framing: implications for public relations. Journal of Public Relations Research 1999;11:205-42.
88. Kahneman D, Tversky A. Choices, values and frames. Am Psychol 1984;39:341-50.
89. Levin I, Schneider S, Gaeth G. All frames are not created equal: a typology and critical analysis of framing effects. Organizational Behavior and Human Decision Process 1998;76:149-88.

90. Augustovski F, Beratarrechea A, Irazola V, et al. Patient preferences for biologic agents in rheumatoid arthritis: a discrete-choice experiment. Value Health 2013;16:385-93.
91. Boye K, Matza L, Walter K, Van Brunt K, Palsgrove A, Tynan A. Utilities and disutilities for attributes of injectable treatments for type 2 diabetes. Eur J Health Econ 2011;12:219-30.
92. de Bekker-Grob E, Essink-Bot M, Meerding W, Pols H, Koes B, Steyerberg E. Patients' preferences for osteoporsosis drug treatment: a discrete choice experiment. Osteoporosis Int 2008;19:1029-37.
93. Higgins A, Barnett J, Meads C, Singh J, Longworth L. Does convenience matter in health care delivery? A systematic review of convenience-based aspects of process utility. Value Health 2014;17:877-87.
94. Gerard K, Lattimer V, Surridge H, et al. The introduction of integrated out-of-hours arrangements in England: a discrete choice experiment of public preferences for alternative models of care. Health Expectations 2006;9:60-9.
95. Rubin G, Bate A, George A, Shackley P, Hall N. Preferences for access to the GP: a discrete choice experiment. Br J Gen Pract 2006;56:743-8.
96. Gerard K, Salisbury C, Street D, Pope C, Baxter H. Is fast access to general practice all that should matter? A discrete choice experiment of patients' preferences. J Health Serv Res Policy 2008;13 Suppl 2:3-10.
97. Phillips K, Johnson F, Maddala T. Measuring what people value: a comparison of "attitude" and "preference" surveys. HSR: Health Services Research 2002;37:1659-79.
98. Haas M. The benefits of health care beyond health: an exploration of non-health outcomes of health care [Doctoral]. Sydney: University of Sydney; 2002.
99. Salkeld G, Quine S, Cameron ID. What constitutes success in preventive health care? a case study in assessing the benefits of hip protectors. Soc Sci Med 2004;59:1593-601.
100. Koszegi B. Health anxiety and patient behavior. J Health Econ 2003;22:1073-84.
101. Bunge EM, de Bekker-Grob EW, van Biezen FC, Essink-Bot ML, de Koning HJ. Patients' preferences for scoliosis brace treatment: a discrete choice experiment. Spine 2010;35:57-63.
102. Ratcliffe J, Brazier J, Tsuchiya A, Symonds T, Brown M. Using DCE and ranking data to estimate cardinal values for health states for deriving a preference-based single index from the sexual quality of life questionnaire. Health Econ 2009;18:1261-76.
103. Sadique MZ, Legood R. Women's preferences regarding options for management of atypical, borderline or low-grade cervical cytological abnormalities: A review of the evidence. Cytopathology 2012;23:161-6.
104. Howard K, Salkeld G, McCaffery K, Irwig L. HPV triage testing or repeat pap smear for the management of a typical squamous cells (ASCUS) on pap smear: Is there evidence of process utility? Health Econ 2008;17:593-605.
105. Philips Z, Avis M, Whynes D. Introducing HPV triage into the English Cervical Cancer Screening Program: consequences for participation. Women Health 2006;43:17-34.
106. Cairns J, Shackley P, Hundley V. Decision making with respect to diagnostic testing: a method of valuing the benefits of antenatal screening. Med Decis Making 1996;16:161-8.
107. Chancellor J, Aballea S, Lawrence A, et al. Preferences of patients with diabetes mellitus for inhaled versus injectable insulin regimens. Pharmacoeconomics 2008;26:217-37.
108. Cook J, Richardson JR, Street A. A cost utility analysis of treatment options for gallstone disease: methodological issues and results. Health Econ 1994;3:157-68.
109. Polster M, Zanutto E, McDonald S, Conner C, Hammer M. A comparison of preferences for two GLP-1 products - liraglutide and exenatide - for the treatment of type 2 diabetes. J Med Econ 2010;13:655-61.
110. Schmier J, Palmer C, Flood E, Gourlay G. Utility assessments of opioid treatment for chronic pain. Pain Med 2002;3:218-30.

111. Gidengil C, Lieu T, Payne K, Rusinak D, Messonnier M, Prosser L. Parental and societal values for the risks and benefit of childhood combination vaccines. Vaccine 2012;30:3445-52.
112. Phillips K, Maddala T, Johnson F. Conjoint Analysis: An application to HIV testing. HSR: Health Services Research 2002;37:1681-705.
113. Prosser LA, Kuntz KM, Bar-Or A, Weinstein MC. Patient and community preferences for treatments and health states in multiple sclerosis. Mult Scler 2003;9:311-9.
114. Louviere J, Flynn T, Carson R. Discrete choice experiments are not conjoint analysis. Journal of Choice Modelling 2010;3:57-72.
115. Fiebig DG, Knox S, Viney R, Haas M, Street DJ. Preferences for new and existing contraceptive products. Health Econ 2011;20 Suppl 1:35-52.
116. Brown TM, Pashos CL, Joshi AV, Lee WC. The perspective of patients with haemophilia with inhibitors and their care givers: preferences for treatment characteristics. Haemophilia 2011;17:476-82.
117. Brett Hauber A, Mohamed AF, Beam C, Medjedovic J, Mauskopf J. Patient preferences and assessment of likely adherence to hepatitis C virus treatment. J Viral Hepat 2011;18:619-27.
118. Waschbusch DA, Cunningham CE, Pelham WE, et al. A discrete choice conjoint experiment to evaluate parent preferences for treatment of young, medication naive children with ADHD. J Clin Child Adolesc Psychol 2011;40:546-61.
119. Sung L, Alibhai SM, Ethier MC, et al. Discrete choice experiment produced estimates of acceptable risks of therapeutic options in cancer patients with febrile neutropenia. J Clin Epidemiol 2012;65:627-34.
120. Guimaraes C, Marra CA, Colley L, et al. Socioeconomic differences in preferences and willingness-to-pay for insulin delivery systems in type 1 and type 2 diabetes. Diabetes Technol Ther 2009;11:567-73.
121. Hechmati G, Hauber AB, Arellano J, et al. Patients' preferences for bone metastases treatments in France, Germany and the United Kingdom. Support Care Cancer 2015;23:21-8.
122. Torbica A, Fattore G, Ayala F. Eliciting preferences to inform patient-centred policies: the case of psoriasis. Pharmacoeconomics 2014;32:209-23.
123. Wilson L, Loucks A, Bui C, et al. Patient centered decision making: use of conjoint analysis to determine risk-benefit trade-offs for preference sensitive treatment choices. J Neurol Sci 2014;344:80-7.
124. Porzsolt F, Clouth J, Deutschmann M, Hippler HJ. Preferences of diabetes patients and physicians: a feasibility study to identify the key indicators for appraisal of health care values. Health & Quality of Life Outcomes 2010;8:125.
125. Kauf TL, Mohamed AF, Hauber AB, Fetzer D, Ahmad A. Patients' willingness to accept the risks and benefits of new treatments for chronic hepatitis C virus infection. The Patient: Patient-Centered Outcomes Research 2012;5:265-78.
126. Bryan S, Jowett S. Hypothetical versus real preferences: results from an opportunistic field experiment. Health Econ 2010;19:1502-9.
127. Matza LS, Sapra SJ, Dillon JF, et al. Health state utilities associated with attributes of treatments for hepatitis C. Eur J Health Econ 2015;16:1005-18.
128. Hauber AB, Gonzalez JM, Schenkel B, Lofland JH, Martin S. The value to patients of reducing lesion severity in plaque psoriasis. Journal of Dermatological Treatment 2011;22:266-75.
129. Chan YM, Sahota DS, Leung TY, Choy KW, Chan OK, Lau TK. Chinese women's preferences for prenatal diagnostic procedure and their willingness to trade between procedures. Prenat Diagn 2009;29:1270-6.
130. Gerard K, Shanahan M, Louviere J. Using Stated Preference Discrete Choice Modelling to Inform Health Care Decision-Making: A Pilot Study of Breast Screening Participation. Appl Econ 2003;35:1073-85.

131. Oteng B, Marra F, Lynd LD, Ogilvie G, Patrick D, Marra CA. Evaluating societal preferences for human papillomavirus vaccine and cervical smear test screening programme. Sex Transm Infect 2011;87:52-7.
132. Benning TM, Dellaert BG, Dirksen CD, Severens JL. Preferences for potential innovations in non-invasive colorectal cancer screening: A labeled discrete choice experiment for a Dutch screening campaign. Acta Oncol 2014;53:898-908.
133. Yasunaga H, Ide H, Imamura T, Ohe K. Benefit evaluation of mass screening for prostate cancer: Willingness-to-pay measurement using contingent valuation. Urology 2006;68:1046-50.
134. Naik-Panvelkar P, Armour C, Rose J, Saini B. Patients' value of asthma services in Australian pharmacies: the way ahead for asthma care. J Asthma 2012;49:310-6.
135. Davison SN, Kromm SK, Currie GR. Patient and health professional preferences for organ allocation and procurement, end-of-life care and organization of care for patients with chronic kidney disease using a discrete choice experiment. Nephrology Dialysis Transplantation 2010;25:2334-41.
136. Franken M, Koolman X. Health system goals: a discrete choice experiment to obtain societal valuations. Health Policy 2013;112:28-34.
137. de Bekker-Grob EW, Bliemer MC, Donkers B, et al. Patients' and urologists' preferences for prostate cancer treatment: a discrete choice experiment. Br J Cancer 2013;109:633-40.
138. Kaambwa B, Lancsar E, McCaffrey N, et al. Investigating consumers' and informal carers' views and preferences for consumer directed care: A discrete choice experiment. Soc Sci Med 2015;140:81-94.
139. Dwight-Johnson M, Lagomasino IT, Hay J, et al. Effectiveness of collaborative care in addressing depression treatment preferences among low-income Latinos. Psychiatr Serv 2010;61:1112-8.
140. Dwight Johnson M, Apesoa-Varano C, Hay J, Unutzer J, Hinton L. Depression treatment preferences of older white and Mexican origin men. Gen Hosp Psychiatry 2013;35:59-65.
141. Marsidi N, van den Bergh MW, Luijendijk RW. The best marketing strategy in aesthetic plastic surgery: evaluating patients' preferences by conjoint analysis. Plast Reconstr Surg 2014;133:52-7.
142. Palumbo A, De La Fuente P, Rodriguez M, et al. Willingness to pay and conjoint analysis to determine women's preferences for ovarian stimulating hormones in the treatment of infertility in Spain. Hum Reprod 2011;26:1790-8.
143. Shackley P, Slack R, Michaels J. Vascular patients' preferences for local treatment: An application of conjoint analysis. J Health Serv Res Policy 2001;6:151-7.
144. Snoek GJ, van Til JA, Krabbe PF, Ijzerman MJ. Decision for reconstructive interventions of the upper limb in individuals with tetraplegia: the effect of treatment characteristics. Spinal Cord 2008;46:228-33.
145. Van der Pol M, Cairns J. Establishing patient preferences for blood transfusion support: An application of conjoint analysis. J Health Serv Res Policy 1998;3:70-6.
146. Donaldson C, Shackley P. Does 'process utility' exist? A case study of willingness to pay for laparoscopic cholecystectomy. Soc Sci Med 1997;44:699-707.
147. Scotland GS, McNamee P, Cheyne H, Hundley V, Barnett C. Women's preferences for aspects of labor management: results from a discrete choice experiment. Birth 2011;38:36-46.
148. Hendrix M, Pavlova M, Nieuwenhuijze MJ, Severens JL, Nijhuis JG. Differences in preferences for obstetric care between nulliparae and their partners in the Netherlands: a discrete-choice experiment. Journal of Psychosomatic Obstetrics & Gynecology 2010;31:243-51.
149. Ryan M. Using conjoint analysis to take account of patient preferences and go beyond health outcomes: An application to in vitro fertilization. Soc Sci Med 1999;48:535-46.

150. Bijlenga D, Birnie E, Bonsel GJ. Feasibility, reliability, and validity of three health-state valuation methods using multiple-outcome vignettes on moderate-risk pregnancy at term. Value Health 2009;12:821-7.
151. Yee LM, Kaimal AJ, Houston KA, et al. Mode of delivery preferences in a diverse population of pregnant women. Am J Obstet Gynecol 2015;212:377 e1-24.
152. Carroll FE, Al-Janabi H, Flynn T, Montgomery AA. Women and their partners' preferences for Down's syndrome screening tests: a discrete choice experiment. Prenat Diagn 2013;33:449-56.
153. Kan VY, Marquez Azalgara V, Ford JA, Peter Kwan WC, Erb SR, Yoshida EM. Patient preference and willingness to pay for transient elastography versus liver biopsy: A perspective from British Columbia. Can J Gastroenterol Hepatol 2015;29:72-6.
154. Marti J. A best-worst scaling survey of adolescents' level of concern for health and non-health consequences of smoking. Soc Sci Med 2012;75:87-97.
155. Burton CR, Fargher E, Plumpton C, Roberts GW, Owen H, Roberts E. Investigating preferences for support with life after stroke: a discrete choice experiment. BMC Health Serv Res 2014;14:63.
156. Nafees B, Lloyd A, Elkin E, Porret T. To explore preferences and willingness to pay for attributes regarding stoma appliances amongst patients in the UK, France and Germany. Curr Med Res Opin 2015;31:687-95.
157. Haas M. The impact of non-health attributes of care on patients' choice of GP. Australian Journal of Primary Health 2005;11:40-6.
158. Park H, Chon Y, Lee J, Choi Ie J, Yoon KH. Service design attributes affecting diabetic patient preferences of telemedicine in South Korea. Telemed J E Health 2011;17:442-51.
159. Ghibjen P, Lancsar E, Zavarsek S. Preferences for oral anticoagulants in atrial fibrilation: a best-best discrete choice experiment. Pharmacoeconomics 2014;32:1115-27.
160. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Second ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1988.
161. Bleichrodt H. A new explanation for the difference between time trade-off utilities and standard gamble utilities. Health Econ 2002;11:447-56.
162. Drew WL, Ives D, Lalezari JP, et al. Oral Ganciclovir as Maintenance Treatment for Cytomegalovirus Retinitis in Patients with AIDS. N Engl J Med 1995;333:615-20.
163. Shani Y, Tykocinski OE, Zeelenberg M. When ignorance is not bliss: How feelings of discomfort promote the search for negative information. J Econ Psychol 2008;29:643-53.
164. Standing Council on Health. National Primary Health Care Strategic Framework. In: Ageing DoHa, ed.: Commonwealth of Australia; 2013.
165. Robinson JC. Theory and practice in the design of physician payment incentives. Milbank Q 2001;79:149-77, III.
166. Christakis D. Continuity of care: process or outcome? Ann Fam Med 2003;1:131-3.
167. Fotaki M. Is patient choice the future of health care systems? International Journal of Health Policy and Management 2013;1:121-3.
168. Victoor A, Delnoij D, Friele R, Rademakers J. Determinants of patient choice of healthcare providers: a scoping review. BMC Health Serv Res 2012;12:1-16.
169. Jatrana S, Crampton P. Affiliation with a primary care provider in New Zealand: Who is, who isn't. Health Policy 2009;91:286-96.
170. McRae I, Yen L, Gillespie J, Douglas K. Patient affiliation with GPs in Australia—Who is and who is not and does it matter? Health Policy 2011;103:16-23.
171. Rundle-Thiele S, Russell-Bennett R. Patient Influences on Satisfaction and Loyalty for GP Services. Health Mark Q 2010;27:195-214.
172. Viera AJ, Pathman DE, Garrett JM. Adults' lack of a usual source of care: a matter of preference? Ann Fam Med 2006;4:359-65.

173. Dreiher J, Comaneshter D, Rosenbluth Y, Battat E, Bitterman H, Cohen A. The association between continuity of care in the community and health outcomes: a population-based study. Israel Journal of Health Policy Research 2012;1:1-12.
174. Weiss L, Blustein J. Faithful patients: the effect of long-term physician-patient relationships on the costs and use of health care by older Americans. Am J Public Health 1996;86:1742-7.
175. Beckman A, Anell A. Changes in health care utilisation following a reform involving choice and privatisation in Swedish primary care: a five year follow-up of GP-visits. BMC Health Serv Res 2013;13:1-9.
176. Freeman G, Richards S. Is personal continuity of care compatible with free choice of doctor? Patients' views on seeing the same doctor. Br J Gen Pract 1993;43:493-7.
177. Campbell BH, Marbella A, Layde PM. Quality of life and recurrence concern in survivors of head and neck cancer. Laryngoscope 2000;110:895-906.
178. Kuzmanovic M, Vujosevic M, Martic M. Using Conjoint Analysis to Elicit Patients' Preferences for Public Primary Care Service in Serbia. HealthMed 2012;6:496-504.
179. Mengoni A, Seghieri C, Nuti S. Heterogeneity in Preferences for Primary Care Consultations: Results from a Discrete Choice Experiment. International Journal of Statistics in Medical Research 2013;2:67-75.
180. Longo MF, Cohen DR, Hood K, et al. Involving patients in primary care consultations: assessing preferences using discrete choice experiments. Br J Gen Pract 2006;56:35-42.
181. Pedersen LB, Kjaer T, Kragstrup J, Gyrd-Hansen D. Does the inclusion of a cost attribute in forced and unforced choices matter? Results from a web survey applying the discrete choice experiment. Journal of Choice Modelling 2011;4:88-109.
182. Vick S, Scott A. Agency in health care. Examining patients' preferences for attributes of the doctor-patient relationship. J Health Econ 1998;17:587-605.
183. O'Brien R. A Caution Regarding Rules of Thumb for Variance Inflation Factors. Quality & Quantity 2007;41:673–90.
184. Cheraghi-Sohi S, Hole AR, Mead N, et al. What Patients Want From Primary Care Consultations: A Discrete Choice Experiment to Identify Patients' Priorities. Ann Fam Med 2008;6:107-15.
185. Turner D, Tarrant C, Windridge K, et al. Do patients value continuity of care in general practice? An investigation using stated preference discrete choice experiments. J Health Serv Res Policy 2007;12:132-7.
186. Scott A, Watson MS, Ross S. Eliciting preferences of the community for out of hours care provided by general practitioners: a stated preference discrete choice experiment. Soc Sci Med 2003;56:803-14.
187. Seghieri C, Mengoni A, Nuti S. Applying discrete choice modelling in a priority setting: an investigation of public preferences for primary care models. Eur J Health Econ 2014;15:773-85.
188. Hole AR. Modelling heterogeneity in patients' preferences for the attributes of a general practitioner appointment. J Health Econ 2008;27:1078-94.
189. Zickafoose JS, DeCamp LR, Prosser LA. Parents' preferences for enhanced access in the pediatric medical home: a discrete choice experiment. JAMA Pediatr 2015;169:358-64.
190. Hjelmgren J, Anell A. Population preferences and choice of primary care models: A discrete choice experiment in Sweden. Health Policy 2007;83:314-22.
191. Boonen LH, Donkers B, Schut FT. Channeling consumers to preferred providers and the impact of status quo bias: does type of provider matter? Health Serv Res 2011;46:510-30.
192. Pedersen LB, Kjaer T, Kragstrup J, Gyrd-Hansen D. Do general practitioners know patients' preferences? An empirical study on the agency relationship at an aggregate level using a discrete choice experiment. Value Health 2012;15:514-23.

193. Caldow J, Bond C, Ryan M, et al. Treatment of minor illness in primary care: a national survey of patient satisfaction, attitudes and preferences regarding a wider nursing role. Health Expect 2007;10:30-45.
194. Haas M, De Abreu Lourenco R, Hall J. The REFinE-PHC Consumer Survey: Overview of Initial Results. Sydney: University of Technology Sydney; 2014.
195. Australian Bureau of Statistics (ABS). Patient Experiences in Australia: Summary of Findings, 2012-13. Canberra: Australian Bureau of Statistics; 2013.
196. International Health Policy Survey. 2013. (Accessed January 13, 2014, at http://www.commonwealthfund.org/publications/surveys/2013/2013-commonwealth-fund-international-health-policy-survey.)
197. Kenny P, De Abreu Lourenco R, Wong CY, Haas M, Goodall S. Community preferences in general practice: important factors for choosing a general practitioner. Health Expect 2015.
198. PureProfile. 2014. (Accessed January 13, 2014, at http://www.pureprofile.com/au.)
199. Qaltrics. 2014. (Accessed January 13, 2014, at http://www.qualtrics.com/.)
200. Winkelmann R, Boes S. Analysis of microdata. Second ed. Berlin: Springer; 2009.
201. Veale BM, McCallum J, Saltman DC, Lonergan J, Wadsworth YJ, Douglas RM. Consumer use of multiple general practitioners: an Australian epidemiological study. Fam Pract 1995;12:303-8.
202. Doiron D, Jones G, Savage E. Healthy, wealthy and insured? The role of self-assessed health in the demand for private health insurance. Health Econ 2008;17:317-34.
203. Buchmueller T, Fiebig D, Jones G, Savage E. Preference heterogeneity and selection in private health insurance: The case of Australia. J Health Econ 2013;32:757-67.
204. Nutting P, Goodwin M, Flocke S, Zyzanski S, Stange K. Continuity of primary care: to whom does it matter and when? Ann Fam Med 2003;1:149-55.
205. Nagraj S, Abel G, Paddison C, et al. Changing practice as a quality indicator for primary care: analysis of data on voluntary disenrollment from the English GP Patient Survey. BMC Fam Pract 2013;14:1-8.
206. Train K. Probit. Discrete choice methods with simulation. Cambridge: Cambridge University Press; 2003:101-37.
207. Keane MP. A Note on Identification in the Multinomial Probit Model. Journal of Business & Economic Statistics 1992;10:193-200.
208. Belkar R, Fiebig DG, Haas M, Viney R. Why worry about awareness in choice problems? Econometric analysis of screening for cervical cancer. Health Econ 2006;15:33-47.
209. Montmarquette C, Mahseredjian S, Houle R. The determinants of university dropouts: a bivariate probability model with sample selection. Econ Educ Rev 2001;20:475-84.
210. Fiebig D, Uldry P. Sensitivity bounds for use with flawed data. Mathematics and Computers in Simulation 1999;48:479-86.
211. Donders A, van der Heijden G, Stijnen T, Moons K. Review: a gentle introduction to imputation of missing values. J Clin Epidemiol 2006;59:1087-91.
212. Dow JK, Endersby JW. Multinomial probit and multinomial logit: a comparison of choice models for voting research. Electoral Studies 2004;23:107-22.
213. Stata Corp LP. Stata. College Station, USA: Stata Corp LP; 2012.
214. Australian Bureau of Statistics (ABS). Census. 2011. (Accessed January 13, 2014, at http://www.censusdata.abs.gov.au/census_services/getproduct/census/2011/quickstat/0?opendocument&navpos=220.)
215. Australian Bureau of Statistics (ABS). Australian Health Survey: First results, 2011-12. 2013. (Accessed January 13, 2014, at http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0012011-12?OpenDocument.)

216. Australian Bureau of Statistics (ABS). 4839.0 - Patient experiences in Australia: Summary of findings, 2012-13. 2013. (Accessed January 13, 2014, at http://www.abs.gov.au/ausstats/abs@.nsf/mf/4839.0.)
217. Mazumdar S, Konings P, Butler D, McRae IS. General practitioner (family physician) workforce in Australia: comparing geographic data from surveys, a mailing list and medicare. BMC Health Serv Res 2013;13:1-15.
218. Commonwealth of Australia. Budget Overview. 2014. (Accessed 24 June, 2015, at http://www.budget.gov.au/2014-15/content/overview/download/Budget_Overview.pdf.)
219. Bridges JF, Hauber AB, Marshall D, et al. Conjoint analysis applications in health--a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value Health 2011;14:403-13.
220. Hauber AB, Gonzalez JM, Groothuis-Oudshoorn CG, et al. Statistical Methods for the Analysis of Discrete Choice Experiments: A Report of the ISPOR Conjoint Analysis Good Research Practices Task Force. Value Health 2016;19:300-15.
221. Hundley V, Ryan M. Are women's expectations and preferences for intrapartum care affected by the model of care on offer? BJOG 2004;111:550-60.
222. Singh J, Cuttler L, Shin M, Silvers JB, Neuhauser D. Medical decision-making and the patient: understanding preference patterns for growth hormone therapy using conjoint analysis. Med Care 1998;36:AS31-45.
223. Poulos C, Hauber A, Gonzalez J, Turpcu A. Patients' willingness to trade off between the duration and frequency of rheumatoid arthritis treatments. Arthritis Care Res 2014;66:1008-15.
224. Kimman ML, Dellaert BG, Boersma LJ, Lambin P, Dirksen CD. Follow-up after treatment for breast cancer: one strategy fits all? An investigation of patient preferences using a discrete choice experiment. Acta Oncol 2010;49:328-37.
225. Bessen T, Chen G, Street J, et al. What sort of follow-up services would Australian breast cancer survivors prefer if we could no longer offer long-term specialist-based care? A discrete choice experiment. Br J Cancer 2014;110:859-67.
226. Street D, Burgess L, Viney R, Louviere J. Designing discrete choice experiments for health care. In: Ryan M, Gerard K, Amaya-Amaya M, eds. Using Discrete Choice Experiments to Value Health and Health Care. The Netherlands: Springer; 2008:47-72.
227. Zwarina K, Huber J, Kuhfeld W. A general method for constructing efficient choice designs. In: Kuhfeld W, ed. Marketing Research Methods in SAS Experimental Design, Choice, Conjoint and Graphical Techniques. Cary, NC, USA: SAS; 2010:265-83.
228. Huber J, Zwarina K. The importance of utility balance in choice designs. J Marketing Res 1996;33:307-17.
229. Rose J, Collins A, Bliemer M, Hensher D. Ngene. The cutting edge of experimental design for stated choice experiments.: ChoiceMetrics Pty Ltd; 2012.
230. Flynn TN. Valuing citizen and patient preferences in health: recent developments in three types of best-worst scaling. Expert Rev Pharmacoecon Outcomes Res 2010;10:259-67.
231. Ratcliffe J, Huynh E, Stevens K, Brazier J, Sawyer M, Flynn T. Nothing about us without us? A comparison of adolescent and adult health-state values for the Child Health Utility-9D using profile case best-worst scaling. Health Econ 2015;25:486-96.
232. Flynn TN. Using conjoint analysis and choice experiments to estimate QALY values: issues to consider. Pharmacoeconomics 2010;28:711-22.
233. Yoo H, Doiron D. The use of alternative preference elicitation methods in complex discrete choice experiments. J Health Econ 2013;32:1166-79.
234. de Bekker-Grob E, Hol L, Donkers B, et al. Labeled versus unlabeled discrete choice experiments in health economics: an application to colorectal cancer screening. Value Health 2010;13:315-23.

235. McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka P, ed. Frontiers in Econometrics. New York: New York: Academic Press.; 1974:105-42.
236. Keane M, Wasi N. Comparing alternative models of heterogeneity in consumer choice behavior. J Appl Econometrics 2013;28:1018–45.
237. Bhat C. An Endogenous Segmentation Mode Choice Model with an Application to Intercity Travel. Transportation Science 1997;31:34-48.
238. Pacifico D, Yoo H. lclogit: A Stata command for fitting latent-class conditional logit models via the expectation-maximization algorithm. The Stata Journal 2013;Volume 13 Number 3: pp. :625-39.
239. Fiebig D, Keane M, Louviere J, Wasi N. The generalized multinomial logit model: Accounting for scale and coefficient heterogeneity. Marketing Sci 2010;29:393–421.
240. Davis KJ, Burton M, Kragt ME. Discrete choice models: scale heterogeneity and why it matters. Working Paper 1602, School of Agricultural and Resource Economics 2016;University of Western Australia, Crawley, Australia.
241. Louviere J, Eagle T. Confound it! That pesky little scale constant messes up our convenient assumptions. Proceedings of the Sawtooth Software Conference 2006; 2006. p. 211-28.
242. Hess S, Rose JM. Can scale and coefficient heterogeneity be separated in random coefficients models? Transportation 2012;39:1225-39.
243. Gu Y, Hole A, Knox S. Fitting the generalized multinomial logit model in Stata. The Stata Journal 2013;13:382–97.
244. Carson R, Czajkowski M. A new baseline model for estimating willingness to pay from discrete choice models. 2013. (Accessed 18 Feb 2015, at http://www.icmconference.org.uk/index.php/icmc/ICMC2013/paper/viewFile/730/354.)
245. Hensher D, Greene W. The mixed logit model: the state of practice. Transportation 2003;30:133-76.
246. Daly A, Hess S, Train K. Assuring finite moments for willingness to pay in random coefficients models. Transportation 2012;39:19-31.
247. Daly A, Hess S, de Jong G. Calculating errors for measures derived from choice modelling estimates. Transportation Research Part B Methodological 2012;46:333-41.
248. Hole A. A comparison of approaches to estimating confidence intervals for willingness to pay measures. Health Econ 2007;16:827–40.
249. Bliemer M, Rose J. Confidence intervals of willingness-to-pay for random coefficient logit models. Transportation Research Part B 2013;58:199-214.
250. de Bekker-Grob EW, Rose J, Bliemer M. A closer look at decision and analyst error by incuding nonlinearities in discrete choice models: implications on willingness-to-pay estimates derived from discrete choice data in healthcare. Pharmacoeconomics 2013;31:1169-83.
251. Lancsar E, Louviere J, Flynn T. Several methods to investigate relative attribute impact in stated preference experiments. Soc Sci Med 2007;64:1738-53.
252. Rose J, Bliemer M. Designing efficient data for stated choice experiments: accounting for socio-demographic and contextual effects in designing stated choice experiments. 11th International Conference on Travel Behaviour Research. Kyoto, Japan2006.
253. Meier F, Frerix M, Hermann W, Muller-Ladner U. Current immunotherapy in rheumatoid arthritis. Immunotherapy 2013;5:955-74.
254. Harrison M, Marra C, Shojania K, Bansback N. Societal preferences for rheumatoid arthritis treatments: evidence from a discrete choice experiment. Rheumatology (Oxford) 2015;54:1816-25.
255. Louder AM, Singh A, Saverno K, et al. Patient Preferences Regarding Rheumatoid Arthritis Therapies: A Conjoint Analysis. American health & drug benefits 2016;9:84-93.
256. Hiligsmann M, Dellaert B, Dirksen C, et al. Patients' preferences for osteoporosis drug treatment: a discrete-choice experiment. Arthritis Res Ther 2014;16:1-9.

257. Hole A. Fitting mixed logit models by using maximum simulated likelihood. The Stata Journal 2007;7:388-401.
258. Kragt M, Bennett J. Attribute framing in choice experiments: how do attribute level descriptions affect value estimates? Environ Resource Econ 2012;51:43-59.
259. Rolfe J, Bennett J, Louviere J. Stated values and reminders of substitute goods: testing for framing effects with choice modelling. The Australian Journal of Agricultural and Resource Economics 2002;46:1-20.
260. Orme ME, Macgilchrist KS, Mitchell S, Spurden D, Bird A. Systematic review and network meta-analysis of combination and monotherapy treatments in disease-modifying antirheumatic drug-experienced patients with rheumatoid arthritis: analysis of American College of Rheumatology criteria scores 20, 50, and 70. Biologics 2012;6:429-64.
261. Nam JL, Ramiro S, Gaujoux-Viala C, et al. Efficacy of biological disease-modifying antirheumatic drugs: a systematic literature review informing the 2013 update of the EULAR recommendations for the management of rheumatoid arthritis. Ann Rheum Dis 2014;73:516-28.
262. Ramiro S, Gaujoux-Viala C, Nam JL, et al. Safety of synthetic and biological DMARDs: a systematic literature review informing the 2013 update of the EULAR recommendations for management of rheumatoid arthritis. Ann Rheum Dis 2014;73:529-35.
263. Pfizer Australia Pty Ltd. Product Information. Enbrel. Etanercept TGA2015.
264. Abbvie Pty Ltd. Product Information. Humira. Adalimumab. 2016. (Accessed Access Date, Access 2016, at https://http://www.ebs.tga.gov.au/ebs/picmi/picmirepository.nsf/pdf?OpenAgent&id=CP-2010-PI-03779-3.)
265. Janssen Cilag Pty Ltd. Product Information. Remicade. Infliximab. In: TGA, ed.2015.
266. Bristol-Myers Squibb Australia Pty Ltd. Product Information. Orencia. Abatacept. In: TGA, ed.2015.
267. Pfizer Australia Pty Ltd. Product Information. Methotrexate. In: TGA, ed.2011.
268. Ryan M, Wordsworth S. Sensitivity of willingness to pay estimates to the level of attributes in discrete choice experiments. Scot J Polit Economy 2000;47:504-24.
269. Rose J, Bliemer M. Stated choice experimental design theory: the who, the what and the why. In: Hess S, Daly A, eds. Handbook of choice modelling. Cheltenham, UK: Edward Elgar Publishing; 2014:152-77.
270. Harrison M, Rigby D, Vass C, Flynn T, Louviere J, Payne K. Risk as an Attribute in Discrete Choice Experiments: A Systematic Review of the Literature. Patient 2014;February:1-20.
271. Knox SA, Viney RC, Gu Y, et al. The effect of adverse information and positive promotion on women's preferences for prescribed contraceptive products. Soc Sci Med 2013;83:70-80.
272. Corp S. Stata/MP 12.1 for Mac In: LP S, ed. College Station, Texas: Stata Corp; 2014.
273. Swait J, Louviere J. The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. J Marketing Res 1993;30:305–14.
274. Edrdem S, Campbell D, Hole A. Accounting for attribute-level non-attendance in a health choice experiment: does it matter? Health Econ 2015;24:773-89.
275. Johnson FR, Mohamed AF, Ozdemir S, Marshall DA, Phillips KA. How does cost matter in health-care discrete-choice experiments? Health Econ 2011;20:323-30.
276. Department of Health (DOH). 5.24 Tofacitinib. Public Summary Document - March 2015 PBAC Meeting. Canberra: Commonwealth Government 2015.
277. van Vollenhoven RF, Fleischmann R, Cohen S, et al. Tofacitinib or adalimumab versus placebo in rheumatoid arthritis. N Engl J Med 2012;367:508-19.
278. Gyrd-Hansen D. Using the stated preference technique for eliciting valuations: the role of the payment vehicle. Pharmacoeconomics 2013;31:853-61.

279. Brazier J, Akehurst R, Brennan A, et al. Should patients have a greater role in valuing health states? Applied health economics and health policy 2005;4:201-8.
280. Ubel PA, Loewenstein G, Jepson C. Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public. Qual Life Res 2003;12:599-607.
281. Cormier O'Connor A, Boyd N, Tritchler D, Kirukov Y, Sutherland H, Till J. Eliciting preferences for alternative cancer drug treatments. The influence of framing, medium, and rater variables. Med Decis Making 1985;5:453-63.
282. Howard K, Salkeld G. Does attribute framing in discrete choice experiments influence willingness to pay? Results from a discrete choice experiment in screening for colorectal cancer. Value Health 2009;12:354-63.
283. Muhlbacher AC, Bethge S. Reduce mortality risk above all else: a discrete-choice experiment in acute coronary syndrome patients. Pharmacoeconomics 2015;33:71-81.
284. Fisher CS, Martin-Dunlap T, Ruppel MB, Gao F, Atkins J, Margenthaler JA. Fear of recurrence and perceived survival benefit are primary motivators for choosing mastectomy over breast-conservation therapy regardless of age. Ann Surg Oncol 2012;19:3246-50.
285. Rosenberg SM, Tracy MS, Meyer ME, et al. Perceptions, knowledge, and satisfaction with contralateral prophylactic mastectomy among young women with breast cancer: a cross-sectional survey. Ann Intern Med 2013;159:373-81.
286. Boughey J, Hoskin T, Degnim A, et al. Contralateral prophylactic mastectomy is associated with a survival advantage in high=-risk women with a personal history of breast cancer. Ann Surg Oncol 2010;17:2702-9.
287. Fayanju O, Stoll C, Fowler S, Colditz G, Margenthaler JA. Contralateral prophylactic mastectomy after unilateral breast cancer. Ann Surg 2014;260:1000-10.
288. Lostumbo L, Carbine N, Wallace J. Prophylactic mastectomy for the prevention of breast cancer. Cochrane Database of Sytematic Reviews 2010;11.
289. Tuttle T, Abbott A, Arrington A, Rueth N. The Increasing Use of Prophylactic Mastectomy in the Prevention of Breast Cancer. Curr Oncol Rep 2010;12:16-21.
290. Nelson JA, Tchou J, Domchek S, Sonnad SS, Serletti JM, Wu LC. Breast reconstruction in bilateral prophylactic mastectomy patients: factors that influence decision making. J Plast Reconstr Aesthet Surg 2012;65:1481-9.
291. Soran A, Kamali Polat A, Johnson R, McGuire KP. Increasing trend of contralateral prophylactic mastectomy: What are the factors behind this phenomenon? Surgeon 2014.
292. Sepucha K, Ozanne E, Silvia K, Partridge A, Mulley AG, Jr. An approach to measuring the quality of breast cancer decisions. Patient Educ Couns 2007;65:261-9.
293. Tate PS, McGee EM, Hopkins SF, Rogers EL, Page GV. Breast conservation versus mastectomy: patient preferences in a community practice in Kentucky. J Surg Oncol 1993;52:213-6.
294. Wei JP, Sherry RM, Baisden BL, Peckel J, Lala G. Prospective hospital-based survey of attitudes of Southern women toward surgical treatment of breast cancer. Ann Surg Oncol 1995;2:360-4.
295. Ananian P, Houvenaeghel G, Protiere C, et al. Determinants of patients' choice of reconstruction with mastectomy for primary breast cancer. Ann Surg Oncol 2004;11:762-71.
296. Lee CN, Hultman CS, Sepucha K. What are patients' goals and concerns about breast reconstruction after mastectomy? Ann Plast Surg 2010;64:567-9.
297. Deshmukh AA, Cantor SB, Crosby MA, et al. Cost of contralateral prophylactic mastectomy. Ann Surg Oncol 2014;21:2823-30.
298. Caldon LJ, Walters SJ, Reed MW. Changing trends in the decision-making preferences of women with early breast cancer. Br J Surg 2008;95:312-8.

299. Damen TH, de Bekker-Grob EW, Mureau MA, et al. Patients' preferences for breast reconstruction: a discrete choice experiment. J Plast Reconstr Aesthet Surg 2011;64:75-83.
300. Osoba D, Hsu MA, Copley-Merriman C, et al. Stated preferences of patients with cancer for health-related quality-of-life (HRQOL) domains during treatment. Qual Life Res 2006;15:273-83.
301. Ali K, Roffe C, Crome P. What patients want: consumer involvement in the design of a randomized controlled trial of routine oxygen supplementation after acute stroke. Stroke 2006;37:865-71.
302. Hanley B, Truesdale A, King A, Elbourne D, Chalmers I. Involving consumers in designing, conducting, and interpreting randomised controlled trials: questionnaire survey. BMJ 2001;322:519-23.
303. Muhlbacher A, Johnson F. Choice experiments to quantify preferences for health and healthcare: state of the practice. Applied Health Economics and Policy 2016;14:1-16.
304. Yi M, Hunt KK, Arun BK, et al. Factors affecting the decision of breast cancer patients to undergo contralateral prophylactic mastectomy. Cancer Prev Res (Phila) 2010;3:1026-34.
305. Yao K, Winchester D, Czechura T, Huo D. Contralateral prophylactic mastectomy and survival: report from the National Cancer Data Base, 1998-2002. Breast Cancer Res Treat 2013;142:465-76.
306. Barton M, West C, Liu I, et al. Complications following bilateral prophylactic mastectomy. Journal of the National Cancer Institute Monographs 2005;35:61-6.
307. Gahm J, Hansson P, Brandberg Y, Wickman M. Breast sensibility after bilateral risk-reducing mastectomy and immediate breast reconstruction: A prospective study. J Plast Reconstr Aesthet Surg 2013;66:1521-7.
308. Gahm J, Wickman M, Brandberg Y. Bilateral prophylactic mastectomy in women with inherited risk of breast cancer-prevalence of pain and discomfort, impact on sexuality, quality of life and feelings of regret two years after surgery. The Breast 2010;19:462-9.
309. Osman F, Saleh F, Jackson T, Corrigan M, Cil T. Increased postoperative complications in bilateral mastectomy patients compared to unilateral mastectomy: an analysis of the NSQIP Database. Ann Surg Oncol 2013;20:3212-7.
310. Frost M, Hoskin T, Hartmann L, Degnim A, Johnson J, Boughey J. Contralateral prophylactic mastectomy: long-term consistency of satisfaction and adverse effects and the significance of informed decision-making, quality of life, and personality traits. Ann Surg Oncol 2011;18:3110-6.
311. Davies K, Cantor S, Brewster A. Better contralateral breast cancer risk estimation and alternative options to contralateral prophylactic mastectomy. International Journal of Women's Health 2015;7:181-7.
312. Cancer Australia. Recommendations for follow-up of women with early breast cancer 2011.
313. Isern A, Tengrup I, Loman N, Olsson H, Ringberg A. Aesthetic outcome, patient satisfaction, and health-related quality of life in women at high risk undergoing prophylactic mastectomy and immediate breast reconstruction. J Plast Reconstr Aesthet Surg 2008;61:1177-87.
314. Rolnick S, Altschuler A, Nekhlyudov L, et al. What women wish they knew before prophylactic mastectomy. Cancer Nurs 2007;30:285-91.
315. Semple J, Metcalfe KA, Lynch HT, et al. International rates of breast reconstruction after prophylactic mastectomy in BRCA1 and BRCA2 mutation carriers. Ann Surg Oncol 2013;20:3817-22.
316. Breast Cancer Network Australia (BCNA). Breast Reconstruction Survey2010.
317. Cancer Australia. Recommendations for the identification and management of fear of cancer recurrence in adult cancer survivors2014.

318. Custers JA, van den Berg SW, van Laarhoven HW, Bleiker EM, Gielissen MF, Prins JB. The Cancer Worry Scale: detecting fear of recurrence in breast cancer survivors. Cancer Nurs 2014;37:E44-50.
319. Thewes B, Butow P, Zachariae R, Christensen S, Simard S, Gotay C. Fear of cancer recurrence: a systematic literature review of self-report measures. Psychooncology 2012;21:571-87.
320. Spencer S, Lehman J, Wynings C, et al. Concerns about breast cancer and relations to psychosocial well-being in a multiethnic sample of early-stage patients. Health Psychol 1999;18:159-68.
321. Australian Bureau of Statistics (ABS). Australian Demographic Statistics. 2015. (Accessed February 11, 2016, at http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3101.0Jun 2015?OpenDocument.)
322. Australian Bureau of Statistics (ABS). Education and Work, Australia, May 2015. 2015. (Accessed February 11, 2016, at http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/6227.0May 2015?OpenDocument.)
323. Australian Bureau of Statistics (ABS). Measures of Australia's Progress, 2013 2015. (Accessed February 11, 2016, at http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by Subject/1370.0~2013~Download Summary~Data and downloads~10.)
324. Australian Institute of Health and Welfare (AIHW). Australian Institute of Health and Welfare (AIHW). Cancer in Australia: an overview 2012. Canberra: AIHW; 2012.
325. Australian Institute of Health and Welfare (AIHW). National Bowel Cancer Screening Program Canberra: AIHW; 2014.
326. Australian Institute of Health and Welfare (AIHW). Cancer screening programs in Australia. 2016. (Accessed 11 February, 2016, at http://www.aihw.gov.au/cancer/screening/.)
327. Altschuler A, Nekhlyudov L, Rolnick SJ, et al. Positive, negative, and disparate--women's differing long-term psychosocial experiences of bilateral or contralateral prophylactic mastectomy. Breast J 2008;14:25-32.
328. Hess S, Rose JM, Polak J. Non-trading, lexicographic and inconsistent behaviour in stated choice data. Transportation Research Part D: Transport and Environment 2010;15:405-17.
329. Barsky A, Orav E, Bates D. Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. Arch Gen Psychiatry 2005;62:903-10.
330. Bermingham SL, Cohen A, Hague J, Parsonage M. The cost of somatisation among the working‐age population in England for the year 2008‐2009. Mental Health in Family Medicine 2010;7:71-84.
331. Mewton L, Andrews G. Poor self-rated health and its associations with somatisation in two Australian national surveys. BMJ Open 2013;3:e002965.
332. Puri P, Dimsdale J. Healthcare utilization and poor reassurance: potential predictors of somatoform disorders. Psychiatr Clin North Am 2011;34:525-44.
333. Ahlert M, Breyer F, Schwettmann L. How you ask is what you get: framing effects in willingness-to-pay for a QALY. Soc Sci Med 2016;150:40-8.
334. Skedgel C, Wailoo A, Akehurst R. Societal preferences for distributive justice in the allocation of health care resources: a latent class discrete choice experiment. Med Decis Making 2015;35:94-105.
335. Rowen D, Brazier J, Keetharuth A, Tsuchiya A, Mukuria C. Comparison of modes of administration and alternative formats for eliciting societal preferences for burden of illness. Applied health economics and health policy 2016;14.
336. Hicks J. The rehabilitation of consumers' surplus. The Review of Economic Studies 1941;8:108-16.

337. Currie J, Murphy J, Schnitz A. The concept of economic surplus and its use in economic analysis. The Economic Journal 1971;324:741-99.
338. Walzer S, Nuitjem M, Wiesner C, Kaier K, Johansson P, Oertel S. Microeconomic surplus in health care: applied economic theory in health care in four European countries. Front Pharmacol 2013;4:1-6.
339. Train K. Consumer surplus. In: Train K, ed. Discrete choice methods with simulation: Cambridge University Press; 2002:64-7.
340. Whynes D, Frew E, Wolstenholme J. Willingness-t-pay and demand curves: a comparison of results obtained using different elicitation formats. International Journal of Health Care Finance and Economics 2005;5:369-86.
341. Office of Best Practice Regulation. Cost-benefit analysis. Guidance note. In: Cabinet DotPMa, ed.: Australian Government; 2014.
342. Commonwealth of Australia. Handbook of Cost Benefit Analysis. 2006.
343. Smith RD, Sach TH. Contingent valuation: (still) on the road to nowhere? Health Econ 2009;18:863-6.
344. Mortimer D. The value of thinly spread QALYs. Pharmacoeconomics 2006;24:845-53.
345. Salkeld G. What are the benefits of preventive health care? Health Care Anal 1998;6:106-12.
346. Bliemer M, Rose J. Efficiency and sample size requirements for stated choice studies. ITLS Working Paper. Sydney: Institute of Transport and Logistic Studies, University of Sydney; 2005.
347. Hensher D, Rose J, Greene W. Processes in setting up stated choice experiments. Applied Choice Analysis A primer. Cambridge: Cambridge University Press; 2005:100-60.
348. Bech M, Gyrd-Hansen D. Effects coding in discrete choice experiments. Health Econ 2005;14:1079-83.
349. Rose J, Bliemer M. Sample optimality in the design of stated choice experiments. ITLS Working Paper. Sydney: Institute of Transport and Logistic Studies, University of Sydney; 2005.