Longitudinal and Life Course Studies • vol 13 • no 4 • 647–666
© Authors 2022 • Accepted for publication 17 January 2022/First published online 07 February 2022
Online ISSN 1757-9597 • https://doi.org/10.1332/175795921X16424353247247

RESEARCH ARTICLE

Health, social and economic implications of adolescent risk behaviours/states: protocol for Raine Study Gen2 cohort data linkage study

Jennifer L. Marino, jennifer.marino@unimelb.edu.au
Royal Women's Hospital, Murdoch Children's Research Institute and University of Melbourne, Australia

Robert J. Tait, robert.tait@curtin.edu.au
Leon M. Straker, L.Straker@curtin.edu.au
Curtin University, Australia

Deborah J. Schofield, deborah.schofield@mq.edu.au
Macquarie University, Australia

Dorota A. Doherty, dorota.doherty@uwa.edu.au
University of Western Australia, Australia

Rebecca Q. Ivers, rebecca.ivers@unsw.edu.au
University of New South Wales, Australia

Petra L. Graham, petra.graham@mq.edu.au
Macquarie University, Australia

Katharine Steinbeck, kate.steinbeck@health.nsw.gov.au
University of Sydney and Children's Hospital at Westmead, Australia

Sharyn Lymer, Sharyn.Lymer@health.nsw.gov.au
University of Sydney, Australia

Lena A. Sanci, l.sanci@unimelb.edu.au
University of Melbourne, Australia

George C. Patton, george.patton@mcri.edu.au
Murdoch Children's Research Institute, Australia

Bette Liu, bette.liu@unsw.edu.au
University of New South Wales, Australia

Fiona M. Brooks, Fiona.Brooks@uts.edu.au
Melissa S. Kang, Melissa.Kang@uts.edu.au
University of Technology Sydney, Australia

Martha Hickey, hickeym@unimelb.edu.au
Royal Women's Hospital and University of Melbourne, Australia

Michelle Cunich, michelle.cunich@sydney.edu.au
University of Sydney and Sydney Local Health District, Australia

Sarita Bista, sarita.bista@sydney.edu.au
S. Rachel Skinner, rachel.skinner@health.nsw.gov.au
University of Sydney and Children's Hospital at Westmead, Australia

Background: Risk-taking behaviours are a major contributor to youth morbidity and mortality. Vulnerability to these negative outcomes is constructed from individual behaviour including risk-taking, and from social context, ecological determinants, early life experience, developmental capacity and mental health, contributing to a state of higher risk. However, although risk-taking is part of normal adolescent development, there is no systematic way to distinguish young people with a high probability of serious adverse outcomes, hindering the capacity to screen and intervene. This study aims to explore the association between risk behaviours/states in adolescence and negative health, social and economic outcomes through young adulthood.
Methods: The Raine Study is a prospective cohort study which recruited pregnant women in 1989–91, in Perth, Western Australia. The offspring cohort (N = 2,868) was followed up at regular intervals from 1 to 27 years of age. These data will be linked to State government health and welfare administrative data.
We will empirically examine relationships across multiple domains of risk (for example, substance use, sexual behaviour, driving) with health and social outcomes (for instance, road-crash injury, educational underachievement). Microsimulation models will measure the impact of risk-taking on educational attainment and labour force outcomes.
Discussion: Comprehensive preventive child health programmes and policy prioritise a healthy start to life. This is the first linkage study focusing on adolescence to adopt a multi-domain approach, and to integrate health economic modelling. This approach captures a more complete picture of health and social impacts of risk behaviour/states in adolescence and young adulthood.

Key words prenatal cohort • young adult • development • substance use • sexual health

Key messages

- The first linkage study of adolescent multi-domain risk-taking using health economic modelling.
- This study links a prenatal cohort to 15 Western Australian government and health databases.
- The most extensive and diverse use of the oldest and most comprehensive data linkage collaboration in Australia.

To cite this article: Marino, J., Tait, R., Straker, L., Schofield, D., Doherty, D., Ivers, R., Graham, P., Steinbeck, K., Lymer, S., Sanci, L., Patton, G., Liu, B., Brooks, F., Kang, M., Hickey, M., Cunich, M., Bista, S., Skinner, S. (2022) Health, social and economic implications of adolescent risk behaviours/·states: protocol for Raine Study Gen2 cohort data linkage study, *Longitudinal and Life Course Studies*, 13(4): 647–666, DOI: 10.1332/175795921X16424353247247

Introduction

Risk-taking behaviours, such as substance use, sexual risk-taking and unsafe driving are a major contributor to youth mortality and morbidity including road-crash injury, sexually transmissible infections (STI), and interpersonal violence, yet these adverse outcomes are largely preventable. Globally, 70% of preventable deaths in adulthood are linked to long-standing patterns of behaviour originating in childhood and adolescence (United Nations Secretary-General, 2015). Alcohol and other drug use are leading risk factors for disease burden in adolescents and young adults (Mokdad et al, 2016; AIHW, 2019). In Australia, road crash is a leading cause of death in adolescents and young adults, with injury of all types accounting for 68% of years of life lost among males and 59% among females (AIHW, 2019). Despite these observations, only recently have adolescence and young adulthood become a focus in global health, most notably with the addition of adolescents to the UN Global Strategy for Health in late 2015, and the *Lancet* Commission on Adolescent Health and Wellbeing in 2016 (Patton et al, 2016). Adolescents have largely been overlooked in preventive programmes worldwide, due in part to lack of longitudinal data, and possibly the widespread view that by adolescence, life trajectories are entrenched and intervention unlikely to be cost-effective. Adolescents are also subject to age-related stigma, which attributes negative outcomes to destructive choices.

This perspective does not consider that vulnerability to these negative health outcomes is constructed, over the life course, from not only individual behaviour including risk-taking, but from past and contemporary social context, ecological determinants, early life experience, developmental capacity, and mental health issues such as depression and suicidality, all of which may contribute to a state of higher risk. Here we present the protocol for a study, using Australian data, to estimate the independent contribution of risk behaviours/states to serious harms and longer-term health and societal costs over the lifespan, considering early life factors and social context more comprehensively than has been done before. This work has been co-designed with the Wellbeing Health & Youth (WH&Y) Commission, a diverse group of young people (14–25 years), uniquely trained in health research ethics, methods and outcome translation (Nguyen et al, 2021). We will develop an evidence base for the costs and benefits of intervention during adolescence. We propose a nuanced model of risk, which acknowledges the tendency of risk behaviours/states and negative outcomes to cluster in particular individuals. We will specifically consider the contribution of gender, a critical social and biological variable, to this model.

The *Lancet* Commission (Patton et al, 2016) argues that adolescent health is critical for three reasons: good health underpins normal adolescent development; adolescence is a period during which the foundations are laid for good health in later life; and

adolescents will become the next generation of parents within a decade, so protecting and enhancing their health provides a dual benefit by supporting a good start for their children (that is, intergenerational impacts).

Risk-taking is part of normal adolescent development. However, in public health dialogue, adolescent risk-taking has been framed chiefly in a deficit model, wherein risk-taking is treated as inherently negative, and as equally damaging to all adolescents. American psychologists Duell and Steinberg have recently argued that adolescent risk-taking may be positive or negative, with the critical differentiating elements the potential benefit to well-being, and the potential severity of cost (Duell and Steinberg, 2019; 2020). However, some young people, such as those who are 'vulnerable', may be more likely to experience a range of negative psychosocial outcomes of risk-taking (Moffitt et al, 2011; Fergusson et al, 2013).

The social and environmental context plays a key antecedent and ongoing role in risk-taking behaviours/states across the life course, a point that has been emphasised to us by the WH&Y Commissioners. But an important aspect of adolescent risk that has been neglected in health is co-occurrence or clustering of risky behaviours in individuals. Risk behaviours do not occur in isolation; the same individual engages in multiple risk behaviours, reflecting the general vulnerability implied by developmental predisposition, and the continued influence of an adverse social context (Noble et al, 2015; Morris et al, 2016). They have many common antecedents in childhood leading to multiple adverse health and social outcomes in adolescence and in adulthood (Akasaki et al, 2019; Mewton et al, 2019). However, epidemiological surveillance and public health generally considers risk-taking in a single domain (for example, sexual behaviour, substance use, driving).

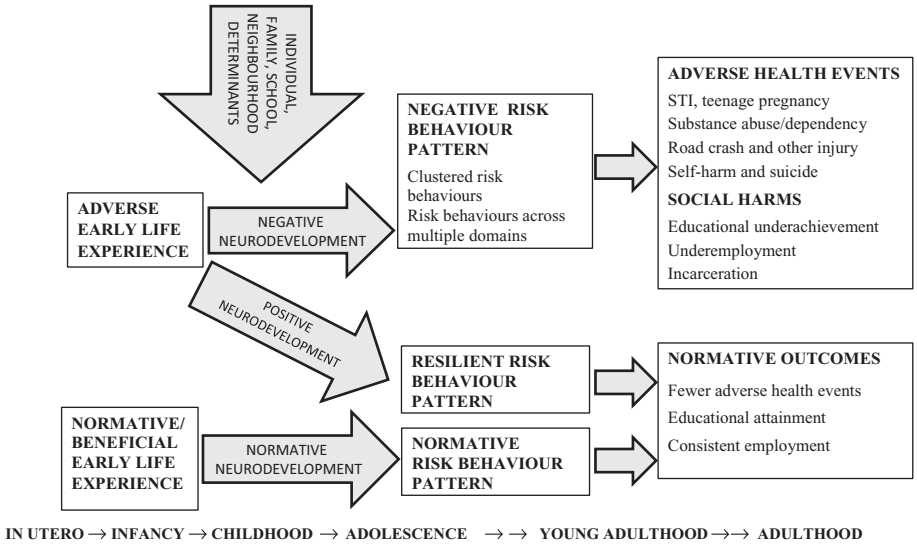
We propose a theoretical model (Figure 1) that recognises the complexity of risk behaviours/states and their outcomes.

We need to be able to identify adolescents at risk of adverse outcomes, the factors that place and keep them on pathways to poorer outcomes, and what distinguishes these adolescents from those who do well. To inform funding for evidence-based interventions and policy, we also need to quantify the societal burden associated with these patterns of risk behaviours/states. We will apply our model to a longitudinal Australian cohort study with records-based outcomes ascertained in routinely collected hospitalisation, educational, licensing and vital statistics data from the Western Australian Data Linkage System (WADLS) (https://www.datalinkage-wa.org.au/; Figure 2). Microsimulation modelling will be undertaken to quantify the impact of patterns of adolescent risk behaviours/states on educational attainment and labour force outcomes.

The aims of this study are:

1. To identify those who experience negative health events and social harms from risk behaviours/states in young adulthood, then characterise their gender-specific patterns of risk behaviours/states in adolescence.
2. To identify the predictors of these risk behaviours/states in early life and childhood, and contemporary ecological determinants, in a data set that follows individuals in the Raine Study prenatal cohort through to young adulthood.
3. To use statistical and microsimulation modelling methods to assess the association between risk behaviours/states, educational outcomes and labour force participation.

Figure 1: Model of adolescent patterns of risk behaviours/states and outcomes in adulthood



The data sets and variables required to ascertain risk behaviours/states and outcomes, as well as contextual variables, are described in more detail in Table 1.

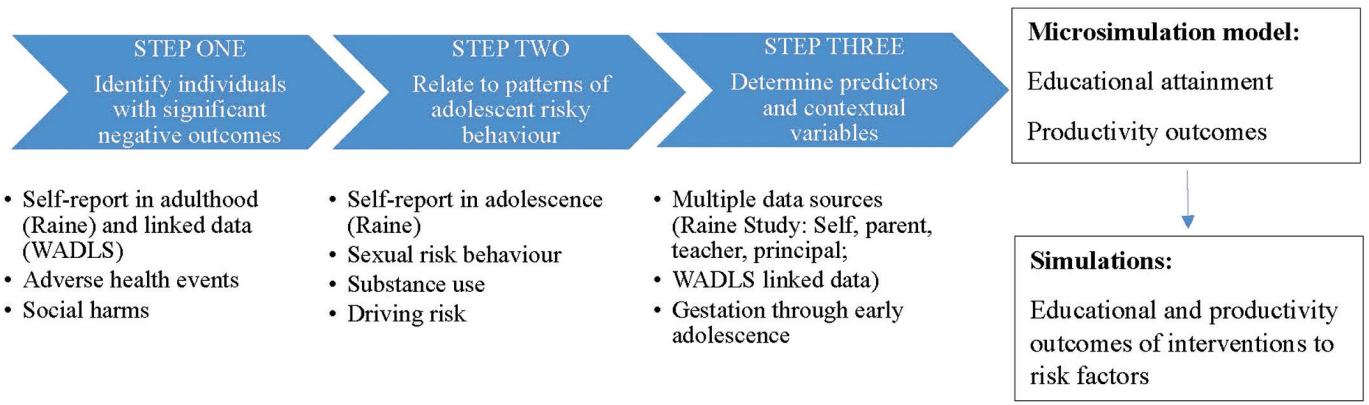
Design

The Raine Study (https://rainestudy.org.au/) is a prospective cohort study t recruited 2,968 pregnant women (Generation 1; Gen1) between May 1989 to November 1991, from a tertiary maternity hospital and nearby clinics in Perth, Western Australia (WA). Although a denominator of all eligible women could not be directly calculated, the participation rate among eligible patients in the clinics was estimated at 90% (Newnham et al, 1993). The 2,868 live births (1,405 female, 1,463 male) comprise the offspring cohort, Generation 2 (Gen2). Follow-up on Gen2 was conducted when the average ages were 1, 2, 3, 5, 8, 10, 14, 17, 18, 20, 22 and 27 years, with extensive clinical and questionnaire data collected. Cohort profile papers have been published for Gen1 (Dontje et al, 2019) and Gen2 (McKnight et al, 2012; Straker et al, 2015; Straker et al, 2017).

We will link all Gen2 cohort members to routinely collected data, allowing us to examine relationships of early life events to outcomes in adulthood for the full cohort. Prevalence estimates of primary outcomes in young adulthood ascertained for the whole cohort will be the basis for most of the microsimulation modelling. Patterns of risk behaviours/states will be assessed based on the 14- and 17-year follow-ups, which were completed by 1,860 (65%) and 1,754 (61%) participants, respectively.

Nearly all (99%) of births in Western Australia in 1989 occurred in a hospital, and 74% took place in Perth hospitals (Gee, 1990). Comparison of the participants at birth, during childhood, and in the 14-, 17-, 20- and 22-year follow-ups with Australian Bureau of Statistics Census data on a range of characteristics has

Figure 2: Steps in statistical and microsimulation modelling of risk behaviours/states in adolescence and outcomes in adulthood



shown consistent representativeness of the retained sample in terms of educational attainment, labour force status, occupation, income level, family structure, area of residence, socio-economic index for areas (SEIFA) indices of relative advantage and disadvantage, and birthplace of parents/carers, compared to the WA population of a similar age (Straker et al, 2017; White et al, 2017). Similarly, in Census data from 1991 (Castles, 1993a-c), when the sample was established, the State was representative of the nation as a whole, with similar median age and income (31 years in WA versus 30 years nationally; annual individual income WA $13,817 versus Australia $13,950; annual household income WA $29,164 versus Australia $29,337). The State and nation had similar cultural diversity, with Western Australians slightly less likely to be Australian-born (69% versus 76%) but more likely to speak English at home (86% versus 83%). The most common languages spoken at home in WA in 1991 were Italian (2.9%) and Chinese languages (1.5%), as in the whole country (Italian 2.6% and Chinese languages 1.6%), but with lower proportions of the third and fourth most common languages in Australia, Greek and Arabic (Greek, WA 0.4% versus Australia 1.8%; Arabic, WA 0.2% versus Australia 0.9%). While Aboriginal and Torres Strait Islander people represented only 1.6% of the Australian population, they comprised 2.6% of the WA population, but of the Australian Aboriginal and Torres Strait Islander population, 15.7% lived in WA (Castles, 1993a-c).

The Western Australian Data Linkage System (WADLS)

WA has the oldest and most comprehensive record linkage system in Australia. This system has already enabled other significant contributions to health research (for example, nutrition (Nyaradi et al, 2015), reproductive health (Hansen et al, 2002), ageing and dementia (Almeida et al, 2012), mental health (Almeida et al, 2018), cerebral palsy (Stanley et al, 1994)). Multiple data sets with records as far back as 1974 are available for routine linkage for epidemiological studies (Hodges et al, 2019). Importantly, this means that for this study, follow-up will be possible for those who have not completed questionnaires at every data collection wave. Data sets that are available for linkage and which will provide valid and comprehensive information for this study are described extensively in Table 1. Records will be linked using probabilistic linkage of name, date of birth, sex and other identifying details. Each record that is found to be likely to belong to the same person is added to a list for that person, called a record chain. Record chains for members of Raine Gen2 would have been established on registration of the birth. Record chains are dynamic, responding to the addition of new data to the collections, and becoming more stable over time. The rate of shifting chains (error rate) for data older than ten years is 0.05% (Eitelhuber, 2016).

Linkage to WADLS resources

Raine Study participants' questionnaire and study data will be linked to a range of data sets managed by the WA Data Linkage Branch (DLB), to establish records-based outcomes (Table 1). The Raine Study data have previously been linked to Department of Education educational attainment, attendance and testing data (Nyaradi et al, 2015). For this new linkage, the procedure will be the same. Briefly,

Table 1: Data sets to be linked to the cohort, main variables to be extracted and time period of extract

Data set	Description	Variables to be extracted	Use	Period
Western Australian Pregnancy Cohort (Raine) Study	Prospective cohort study of 2,868 live births enrolled in 1989–91, followed up to the present	See Table 2	Risk data Augmentation of injury and pregnancy outcomes by self-report Most contextual variables	Birth (1989) to adulthood (most current)
Notifiable Infectious Diseases Database	All diagnoses of infectious diseases that are required to be notified under public health legislation	For each notification of chlamydia, gonorrhoea, syphilis, hepatitis B and C: disease, disease category, organism name, date of onset, hospitalised, confirmation of diagnosis, recent overseas travel, travel country, where infection acquired	Outcomes: sexually transmitted infections	Puberty (1999) to adulthood (most current)
Hospital Morbidity Data	All public and private hospital separations in WA. Diagnoses (up to 20) and procedures (up to 10) coded using International Classification of Diseases (ICD)	All hospital separations from 1989 to current: admission date, separation date, principal procedure date, additional procedure dates, mental health legal status, days of psychiatric care, days of qualified newborn care, days of hospital in the home care, days in ICU, hours in ICU, principal diagnosis, co-diagnosis, additional diagnoses, E-codes, procedure codes, additional procedures	Outcomes: road-related injury, other unintentional injury, self-harm, suicide, substance dependency, substance overdose, pregnancy, chronic illness and mental health diagnoses	Birth (1989) to adulthood (most current)
Emergency Department Data Collection	All emergency department presentations in WA with a code for the primary diagnosis	All emergency department visits from 2002 to current. Admission date, departure destination, departure date, presenting problem, principal diagnosis, triage category, type of visit, external cause of injury, human intent of injury, episode end status	Outcomes: road-related injury, other unintentional injury, self-harm, suicide, substance dependency, substance overdose, pregnancy, chronic illness and mental health diagnoses	Childhood (2002, earliest available) to adulthood (most current)
Midwives Records	All hospital and planned home births in WA of at least 20 weeks' gestation or 400 g birthweight	Height, previous pregnancies, outcomes of previous pregnancies, details of current pregnancy and delivery for babies born to participants	Outcomes: teenage pregnancy, pregnancy complications Contextual variables: reproductive history	Puberty (1999) to adulthood (most current)

(*Continued*)

Table 1: (*Continued*)

Data set	Description	Variables to be extracted	Use	Period
Birth registrations	All registered births in WA	Details of pregnancy and delivery for babies born to participants	Outcomes: teenage pregnancy Contextual variables: reproductive history	Puberty (1999) to adulthood (most current)
Death registrations	All registered deaths in WA	Date of death, pregnant within 6 weeks of death, pregnant within 1 year of death, post-mortem status, causes of death codes and text, antecedent causes, other significant conditions, operations, injury and injury description, place of occurrence of death, activity code, firearm type	Outcomes: road-related injury, other unintentional injury, self-harm, suicide, substance dependency, substance overdose, pregnancy, exacerbations of chronic illness	Birth (1989) to adulthood (most current)
Mental Health Information System	Public and private inpatient and public outpatient psychiatric episodes in WA	Primary diagnosis code, contact type, venue of contact, referral at discharge, referral at admission, date of contact, episode start and end dates, health professional type, group or single, source of data, legal status, first inpatient date, first outpatient date	Outcomes: self-harm, suicide, substance dependency, substance overdose Contextual variables: mental health diagnoses, mental health care	Birth (1989) to adulthood (most current)
Monitoring Drugs of Dependence System (MODD)	Records of Community Program for Opioid Pharmacotherapy, prescribed stimulants and schedule 8 drugs in WA	Duration, dose and indication for prescription of stimulants, opiates, cannabinoids and some benzodiazepines	Outcomes: substance dependency Contextual variables: mental health diagnoses, mental health care	Puberty (1999) to adulthood (most current)
Main Roads WA	Reported road-crash data, including serious injury and fatality crashes, in WA	Date, accident severity, speed factor, road condition, atmospheric condition, accident type, damage, persons in vehicle, injury, protection worn, inattention, fatigue, alcohol, loaded vehicle	Outcomes: road-related injury	Puberty (1999) to adulthood (most current)
Insurance Commission WA Data	Motor vehicle crash injury claims paid by insurance in WA	Date, accident type, ambulance attendance, police attendance, accident role, fatality status, injury code, injury description, drivers licence status, vehicle registration status, injury number, injury source, AIS injury codes, injury severity	Outcomes: road-related injury	Puberty (1999) to adulthood (most current)

(*Continued*)

Table 1: (*Continued*)

Data set	Description	Variables to be extracted	Use	Period
Department of Transportation Licensing Data		Licence dates, demerits, suspensions and cancellations	Risky driving behaviour (demerits and sanctions, stalling in progression from learner's to open licence, for moped, motorcycle and car licence)	Puberty (1999) to adulthood (most current)
Department of Education		WA Literacy and Numeracy Assessment standardised scores in reading, numeracy, spelling and writing, year assessed; school attendance and suspensions; WA Certificate of Education (12th year completion)	Outcomes: educational underachievement Contextual variables: school attendance, suspensions	Puberty (1999) to latest likely date of graduation (2007)
Department of Corrections		Reception date, reception facility, remand date, discharge date, release type, offence code	Outcomes: incarceration	Puberty (1999) to adulthood (most current)
Child Protection Notifications		Substantiated case (including category)	Contextual variables: abuse and maltreatment	Birth (1989) to last turning 18 (2009)
WA Police data		Police cautions, arrests and charges	Outcomes: arrest	Puberty (1999) to adulthood (most current)

a file containing identifying information from the Raine Study cohort (names, dates of birth, sex, addresses), and a unique record number for each participant ('Raine ID'), is sent directly to the DLB. All WADLS databases contain two types of data: identifying information, and clinical or service data (for instance, visit date, test result). DLB staff will use probabilistic matching to compare the identifiers in the Raine Study file to those in the core databases and generate a unique linkage key matching the records in the Raine Study file to the records in the databases. This key replaces all identifiers and is furnished to the data custodian, who will extract the relevant clinical or service data, and return that file, identified only with the data linkage key, to the DLB staff. The DLB staff will replace the data linkage key with the original Raine ID and return the concatenated Raine Study and WADLS data to the researchers. At no point will the researchers have access to the names or addresses of Raine Study participants, nor of any other individuals whose records are held by WADLS – all data handled by the researchers will be de-identified (and unidentifiable by the researchers), and the Raine Study Data Manager will have access only to identifiable data collected by the Raine Study, not to linked data.

Risk behaviours/states

Risk data collected in the Raine Study during adolescence and young adulthood include both behaviours (such as substance use, sexual behaviour) and circumstances/states that increase vulnerability or propensity to risk (for example, mental health symptoms, bullying). Relevant standardised questionnaires used in the Raine Study with Gen2 are listed in Table 2. Participants also reported:

- substance use (14-, 17-, 20-, 22-, 27-year follow-ups): frequency and quantity of alcohol use; tobacco use; illicit drug use, age of first use;
- sexual behaviour, sexual orientation, and gender identity (14-, 17-, 20-, 27-year follow-ups): age of first sexual activity, sexual behaviour, contraception/condom use; gender identity; sexual orientation (attraction, behaviour, identity);
- mental health (14-, 17-, 20-, 22-, 27-year follow-ups): internalising and externalising behaviours, depression and anxiety symptoms, self-harm, suicidal ideation, self-reported suicide attempts, eating disorder symptoms, bullying experience;
- driving behaviour (23-, 27-year follow-ups): seat belt use, sleepy driving, mobile phone use, texting while driving, average kilometres driven;
- school (14-, 17-year follow-ups): attendance, attitudes towards school; and
- exercise and physical activity (14-, 17-, 20-, 22-, 27-year follow-ups): self-reported participation in organised sports, physical activity, screen time.

Study outcomes

Negative outcomes, defined as adverse health events and social harms, will be identified in the linked data (Table 1).

Adverse health events include:

- sexually transmitted infection / bloodborne virus
- adolescent pregnancy and parenthood
- substance use, dependence and overdose (fatal and nonfatal; including tobacco and alcohol)
- road-crash injury in adolescence or adulthood
- lifetime unintentional injury
- self-harm and suicide
- chronic/noncommunicable disease

Social harms include:

- educational underachievement
- unemployment
- arrest and incarceration

Where outcomes are rare events, composite indices will be constructed within the relevant domain based on the number of outcomes and their severity (for example, using disability weights from the Global Burden of Disease Study (Global Burden of Disease Collaborative Network, 2017)).

Table 2: Standardised questionnaires or assessments to be used in analyses

Measure level	Characteristic measured	Questionnaire or assessment
Individual	Mental health	Ages and Stages Questionnaire,[1] Short Temperament Scale for Toddlers,[2] Child Behavior Checklist,[3, 4] Youth Self-Report,[5] Teacher Report Form,[6] Cowen Self-Efficacy Scale,[7] Harter Self-Perception Profile,[8] Beck Depression Inventory for Youth,[9] Depression Anxiety Stress Scale[10]
	Cognition	Peabody Picture Vocabulary Test,[11] Raven Coloured Progressive Matrices,[12] Clinical Evaluation of Language Fundamentals,[13] Symbol Digit Modalities Test[14]
	Nutrition	Dietary Questionnaire for Epidemiological Studies Version 2[15]
Family	Parental mental health	Depression Anxiety Stress Scale[10]
	Parental relationship quality	Abbreviated Dyadic Adjustment Scale,[16] Parent Problem Checklist[17]
	Parenting style	Parenting Scale[18]
	Family function	McMaster Family Assessment Device[19]

Notes:

1. Bricker, D. and Squires, J. (1989) The effectiveness of parental screening of at-risk infants: the Infant Monitoring Questionnaires, *Topics in Early Childhood Special Education*, 9(3): 67–85.

2. Pedlow, R., Sanson, A., Prior, M. and Oberklaid, F. (1993) Stability of maternally reported temperament from infancy to 8 years, *Developmental Psychology*, 29(6): 998–1007.

3. Achenbach, T.M. (1991) *Manual for the Child Behavior Checklist/4–18 and 1991 Profile*, Burlington, VT: Department of Psychiatry, University of Vermont.

4. Achenbach, T.M. (1992) *Manual for the Child Behavior Checklist/2–3 and 1992 Profile*, Burlington, VT: Department of Psychiatry, University of Vermont.

5. Achenbach, T.M. (1991) *Manual for the Youth Self-Report and 1991 Profile*, Burlington, VT: Department of Psychiatry, University of Vermont.

6. Achenbach, T.M. (1991) *Manual for the Teacher's Report Form and 1991 Profile*, Burlington, VT: University of Vermont.

7. Cowen, E.L., Work, W.C., Hightower, A.D., Wyman, P.A., Parker, G.R. and Lotyczewski, B.S. (1991) Toward the development of a measure of perceived self-efficacy in children, *Journal of Clinical Child Psychology*, 20(2): 169–78.

8. Harter, S. (1982) The Perceived Competence Scale for Children, *Child Development*, 53(1): 87–97.

9. Beck, J.S., Beck, A.T. and Jolly, J. (2001) *Manual for the Beck Youth Inventories of Emotional and Social Adjustment*, San Antonio, TX: Psychological Corporation.

10. Lovibond, S.H. and Lovibond, P.F. (1995) *Manual for the Depression Anxiety Stress Scales*, 2nd edn, Sydney, NSW: Psychology Foundation.

11. Dunn, L.M. (1981) *Peabody Picture Vocabulary Test-Revised (PPVT-R)*, Circle Pines, MN: American Guidance Service.

12. Raven, J.C., Court, J.H. and Raven, J.C. (1990) *Manual for Raven's Progressive Matrices and Vocabulary Scales*, Oxford: Oxford Psychologists Press.

13. Semel, E., Wiig, E. and Secord, W. (2003) Clinical Evaluation of Language Fundamentals, 4th edn, San Antonio, TX: Psychological Corporation.

14. Smith, A. (1968) The symbol-digit modalities test, in J. Helmuth (ed) *Learning Disorders*, Seattle, WA, US: Special Child Publications, pp 83–91.

15. Giles, G.G. and Ireland, P.D. (1996) *Dietary Questionnaire for Epidemiological Studies (Version 2)*, Melbourne: Cancer Council Victoria.

Table 2: (*Continued*)

16. Sharpley, C.F. and Rogers, H.J. (1984) Preliminary validation of the Abbreviated Spanier Dyadic Adjustment Scale: some psychometric data regarding a screening-test of marital adjustment, *Educational and Psychological Measurement,* 44(4): 1045–50.

17. Dadds, M.R. and Powell, M.B. (1991) The relationship of interparental conflict and global marital adjustment to aggression, anxiety, and immaturity in aggressive and nonclinic children, *Journal of Abnormal Child Psychology*, 19(5): 553–67.

18. Arnold, D.S., O'Leary, S.G., Wolff, L.S. and Acker, M.M. (1993) The Parenting Scale: a measure of dysfunctional parenting in discipline situations, *Psychological Assessment*, 5(2): 137–44.

19. Epstein, N.B., Baldwin, L.M. and Bishop, D.S. (1983) The McMaster Family Assessment Device, *Journal of Marital and Family Therapy*, 9(2): 171–80.

Data analyses

Exploratory statistical tests will be conducted to test the relationships between variables and differences by pattern of risk behaviours/states (for instance, t-tests, chi-squared tests, ANOVA tests). The role of contextual variables will be identified by examining the associations of these variables with outcomes and exposures. The confounding effects will be controlled by including contextual variables in the statistical models, and propensity score matching will be used where relevant when comparing by pattern of risk behaviours/states.

Clusters of behaviours within and across domains, and of outcomes, will be explored using a range of statistical techniques such as hierarchical agglomerative cluster analysis and latent cluster analysis to evaluate patterns of clustering after adjustment for confounding. Where risk behaviours/states were measured repeatedly over adolescence and young adulthood, latent class growth analysis and growth mixture models will be used to identify trajectory groups. Outcomes (or, where necessary, indices) will be related to individual risk behaviours, clusters of risk behaviours, and trajectories of risk behaviours in the context of other determinants. Supplementary analyses utilising recursive partitioning models (such as binary trees or regression trees) will be used to identify features of those individuals most at risk by investigating multivariable and non-linear associations between predictors and adverse outcomes. Multivariable logistic and log-binomial regression analysis will be used to estimate the risk magnitude for outcomes by identified risk groups obtained via clustering, recursive partitioning or latent class growth analysis. Analyses will be performed using statistical packages most suitable for each analytic model including Stata, R, SAS and Mplus.

Characteristics considered will include individual participant factors (mental health measures; cognitive measures; nutrition; growth), family factors (parental mental health measures; family structure; parental relationship quality measures; household income; parent education and employment; parent-reported parent time with child; parental substance use), school factors (attendance; achievement; teacher involvement; participant attitude towards school) and neighbourhood factors (parent and participant report of housing quality; neighbourhood liveability; facilities and building conditions; Socioeconomic Indexes for Areas). Standardised questionnaires and assessments used to collect data are listed in Table 2. The relative contribution of individual and family factors, versus higher-level factors, will be evaluated in multilevel analyses.

Power

Since we will link the whole cohort, up to 2,400 participants will be followed up for outcomes associated through linkage, about 50% of them female. We expect to see a broad range of frequencies of adverse outcomes. For example, based on existing WA infectious disease notification data of most common STI in young people, we expect 300 chlamydia infections by age 27 (male and female), with a cumulative incidence of 12%. In contrast, educational underachievement is considerably more common: in 2015, 28.5% of all Western Australians 25–29 years old had not completed Year 12. For categorical outcomes with frequencies as low as 5%, at a power of 80%, we will be able to detect relative risks of magnitude 2.3–2.6 in gender-specific logistic regression analysis, while adjusting for other candidate covariates of interest (conservative partial $r^2 = 0.20$). For more common outcomes, with incidence up to 30%, under the same conditions, we will be able to detect relative risks of 1.8–2.1. Similarly, multivariable Cox proportional hazard models for time until event will detect hazard ratios of 1.5–2.0 for candidate continuous covariates. All regression analyses using repeated measures will attain power close to 100% (PASS 2019 Power and Sample Size Program for Windows, Kaysville, Utah).

Health economic modelling

This project will develop statistical models to estimate the impact of early life, adolescent risk behaviours/states, and contextual variables associated with educational attainment (such as Tertiary Institutions Service Centre test scores and highest level of education completed), labour force outcomes (such as absenteeism and unemployment) and income outcomes in young adulthood, and will use microsimulation modelling (MSM) techniques to simulate how these outcomes may change with the implementation of an intervention. MSM is based on micro-data focusing on the characteristics, behaviours and decisions of individuals (Schofield et al, 2018). It allows the combination of data from multiple sources, modelling of population subgroups, and representation of individual heterogeneity to estimate distributions for projected consequences of various intervention choices (Krijkamp et al, 2018). The statistical and microsimulation modelling will use Raine Study data as the base population. The frequency of individual adverse outcomes, outcome indices and outcome clusters, and significantly associated clusters and trajectories of risk-taking behaviours will be estimated, and these frequencies will inform the microsimulation modelling. The MSM modelling and statistical analyses will be developed and undertaken in a suitable statistical package.

Organisation

Participant and public involvement

Raine Study participants have been involved in operations and research since they were young. Gen1 and Gen2 participants sit on all committees, including the Scientific Review Committee, which approved this project and will approve all related manuscripts and public communications.

Leadership of participant engagement is through the Community Advisory Committee. This committee meets quarterly with the Raine Study team and researchers, and helps guide the future direction of the Raine Study, providing input into issues raised by staff, researchers and research partners.

Selection of participants and consent

For child participation in the Raine Study, written informed consent was obtained from parents or guardians at each follow-up through to age 17 (Mountain et al, 2016). Assent was obtained from participants at ages 14 and 17. In 2008, when the cohort began to turn 18 years old, the University of Western Australia Human Research Ethics Committee (UWA HREC) approved contact of all living original participants (N = 2,829), including those whose parents had withdrawn, to acquire consent for the Raine Study to contact them for future assessments, use previously stored biological samples, and obtain additional information through data linkage to administrative records. Where the Raine Study was unable to contact the participant, a waiver of consent was granted by the UWA HREC. Five participants declined consent, 1,127 consented, and 1,692 (including 949 still actively participating in study follow-ups but unable to be contacted specifically for this consent form) were included in the waiver. The five participants who declined consent were not previously and will not be included in the linkage. Decedents (N = 39) and those covered by the waiver will be included. Creation of linkage keys to the Raine Study as part of the WADLS infrastructure was approved by the Government of Western Australia Department of Health (approval 2012/70).

Funding

The core management of the Raine Study is funded by The University of Western Australia, Curtin University, Women and Infants Research Foundation, Telethon Kids Institute, Edith Cowan University, Murdoch University, The University of Notre Dame Australia and the Raine Medical Research Foundation. In addition to funding the Raine Study Gen2 follow-ups, the Australian National Medical and Health Research Council (NHMRC) supports this work via Centre for Research Excellence APP1134894 and Project Grant APP1161445. Rebecca Ivers is supported by an NHMRC research fellowship APP1136430.

Ethics and other approvals

This work has been approved and is monitored by Human Research Ethics Committees in conformance with the Helsinki Declaration and the Australian National Statement on Ethical Conduct in Human Research (Australian Government et al, 2007 (updated 2018)). Although each Raine Study follow-up received its own ethics approval, in April 2020, the University of Western Australia Human Research Ethics Committee issued a single consolidated approval for use of research data and/or biosamples held in the Raine Study data collection (RA/4/20/5722). Use of this approval has been provided on the basis that the release of Raine Study data and/or biosamples for a specific project has been approved by the Raine Study. All projects must undergo the Raine Study governance and reviewing processes before a project can be submitted to

an HREC. The current linkage project has been approved by the Raine Study Scientific Committee (approvals PRIS2-2015, PRIS2amend and PRIS4-2017) and subsequently the Curtin University HREC (approval HRE2019-0774). Approvals for the current linkage project will also be obtained from the Government of Western Australia Department of Health HREC (WA DOH HREC) and WA Research Governance Office. Approval for use of justice data will be obtained from the Department of Justice Research Application and Advisory Committee. Each research manuscript generated using the linked data will be approved by the Raine Study Scientific Committee and the Government of Western Australia HREC. Any publications describing cause of death will be approved by the Australian Coordinating Registry.

Data access

The data sets to be generated for this study are not publicly available due to the terms of the ethics approval granted by WA DOH HREC and data disclosure policies of the Data Providers (including the Raine Study). The data sets may be available from the corresponding author upon request and subject to approval from the DOH HREC and relevant custodians, as well as the Raine Study. The data may be accessed by contacting the corresponding and senior authors (JLM and SRS, respectively) for approval to join the author group, then following the Raine Study Research Engagement Policy, available at the Raine Study website. The authors will facilitate further approval processes.

The original Raine Study data sets, including identifiers, will not be accessed by the researchers on this project. All data sets for the current study, including Raine Study data and data from Raine linked to WADLS data sets, will be de-identified prior to distribution to project researchers, and will be electronically stored at research centres of analysts (JLM, DJS, PLG, DAD, BL, SRS, RJT, SB, SL), at Melbourne University, Macquarie University, University of Western Australia, University of New South Wales, University of Sydney and Curtin University. Each site has a study-specific Data Management Plan, which has been registered with WADLS, and a study-specific Data Transfer Agreement, which will be approved by the WA Research Governance Office.

There will be no disruption to the storage of information should any members of the research team cease employment on this project. Data will be stored in electronic format on password-protected databases on secure servers at these universities, accessible only by the researchers with approved access. These servers are located behind digital firewalls and monitored continuously. Access to the servers, both physically and electronically, is highly restricted.

Following the required retention period of five years from the date of last publication, all data will be disposed of as outlined in the NHMRC guidelines. Records in electronic format will be destroyed/deleted by reformatting or rewriting to ensure that the data and any 'pointers' in the system are inaccessible, as required by Australian regulations. Hard copies will be shredded and disposed of using a secure confidential document destruction service.

Discussion

Comprehensive preventive child health programmes and policy are currently framed by our understanding of the importance of a healthy beginning to life.

The 'first 1,000 days from conception' is guiding investment in programmes in preventive and early intervention child health programmes in Australia as elsewhere (Moore et al, 2017). This is shifting, however, with increasing recognition of the need for sustaining intervention beyond age two years to school entry (first 2,000 days) (Black et al, 2017). Further, the World Bank's Disease Control Priorities calls for sustained intervention through middle childhood, early and late adolescence, to ensure that gains from investment in early life are maintained and that intervention addresses the unique developmental needs of these key life stages (Bundy et al, 2017). In Australia, health investment in adolescence has focused almost exclusively on the single health domain of mental health. This is insufficient to achieve the outcomes of improved mental health, let alone improvements in health and well-being in adolescence more broadly, and to improve outcomes across the life course.

Conclusions

Risk-taking behaviour is part of normal adolescent development, but much youth and young adult morbidity, mortality and social adversity such as poor educational attainment, reduced employment opportunities or criminal records, arises from risky behaviour in adolescence, notably co-occurring risk. No systemic way exists to distinguish adolescents with high probability of serious adverse outcomes, limiting our capacity to screen and intervene. Instead of taking the standard approach of measuring outcomes only in single health domains (such as sexual health or substance use), we will explore relationships across multiple domains, including different types of risk behaviours clustered in individuals, to single and clustered outcomes. To our knowledge, this is the first linkage study to attempt to take this holistic approach, and to integrate health economic modelling. In this way, we can capture both a more complete picture of the health and social impacts of risky behaviour in adolescence, and the potential benefits of investment in this age group. We will be poised to present recommendations to government and other agencies regarding costs and benefits of effective intervention in adolescence.

Conflict of interest

The authors declare that there is no conflict of interest.

References

AIHW (Australian Institute of Health and Welfare) (2019) *Australian Burden of Disease Study: Impact and Causes of Illness and Death in Australia 2015*, Australian Burden of Disease series no 19, cat no BOD 22, Canberra, ACT: AIHW.

Akasaki, M., Ploubidis, G.B., Dodgeon, B. and Bonell, C.P. (2019) The clustering of risk behaviours in adolescence and health consequences in middle age, *Journal of Adolescence*, 77: 188–97. doi: 10.1016/j.adolescence.2019.11.003

Almeida, O.P., Hankey, G.J., Yeap, B.B., Golledge, J. and Flicker, L. (2018) Substance use among older adults with bipolar disorder varies according to age at first treatment contact, *Journal of Affective Disorders*, 239: 269–73. doi: 10.1016/j.jad.2018.07.013

Almeida, O.P., Yeap, B.B., Alfonso, H., Hankey, G.J., Flicker, L. and Norman, P.E. (2012) Older men who use computers have lower risk of dementia, *PLoS ONE*, 7(8): 1-6, doi: 10.1371/journal.pone.0044239.

Australian Government, National Health and Medical Research Council, Australian Research Council and Australian Vice-Chancellors' Committee (2007 (updated 2018)) *National Statement on Ethical Conduct in Human Research*, Canberra, ACT: National Health and Medical Research Council.

Black, M.M., Walker, S.P., Fernald, L.C.H., Andersen, C.T., DiGirolamo, A.M., Lu, C., McCoy, D.C., Fink, G., Shawar, Y.R., Shiffman, J. et al. (2017) Early childhood development coming of age: science through the life course, *The Lancet*, 389(10064): 77–90. doi: 10.1016/S0140-6736(16)31389-7

Bundy, D.A.P., de Silva, N., Horton, S., Jamison, D.T. and Patton, G.C. (eds) (2017) *Child and Adolescent Health and Development*, Washington, DC: World Bank.

Castles, I. (1993a) *Australia's Aboriginal and Torres Strait Islander Population: 1991 Census*, Canberra, ACT: Australian Bureau of Statistics. https://www.abs.gov.au/AUSSTATS/free.nsf/log?openagent&27400_1991_10_Australias_Aboriginal_and_Torres_Strait_Islander_Population.pdf&2101.0&Publication&9E7C82E87FE099DFCA2574BF00164E9F&&1991&01.10.1993&Previous

Castles, I. (1993b) *Census Characteristics of Australia: 1991 Census of Population and Housing*, Canberra, ACT: Australian Bureau of Statistics. https://www.abs.gov.au/AUSSTATS/free.nsf/log?openagent&27100_1991_20_Census_Characteristics_of_Australia.pdf&2101.0&Publication&792BBD9457634FFECA2574BE00826627&&1991&01.10.1993&Previous

Castles, I. (1993c) *Census Characteristics of Western Australia: 1991 Census of Population and Housing*, Canberra, ACT: Australian Bureau of Statistics. https://www.abs.gov.au/AUSSTATS/free.nsf/log?openagent&27105_1991_100_Census_Characteristics_of_Western_Australia.pdf&2101.0&Publication&919B22599E18A744CA2574BE00836D27&&1991&01.10.1993&Previous

Dontje, M.L., Eastwood, P. and Straker, L. (2019) Western Australian pregnancy cohort (Raine) study: Generation 1, *BMJ Open*, 9(5): 1-11, doi: 10.1136/bmjopen-2018-026276.

Duell, N. and Steinberg, L. (2019) Positive risk taking in adolescence, *Child Development Perspectives*, 13(1): 48–52. doi: 10.1111/cdep.12310

Duell, N. and Steinberg, L. (2020) Differential correlates of positive and negative risk taking in adolescence, *Journal of Youth and Adolescence*, 49(6): 1162–78. doi: 10.1007/s10964-020-01237-7

Eitelhuber, T. (2016) *Data Linkage – Making the Right Connections*, Perth, WA: Government of Western Australia Department of Health, https://www.datalinkage-wa.org.au/wp-content/uploads/2019/02/Data-Linkage-Branch-Linkage-Quality.pdf.

Fergusson, D.M., Boden, J.M. and Horwood, L.J. (2013) Childhood Self-control and adult outcomes: results from a 30-year longitudinal study, *Journal of the American Academy of Child & Adolescent Psychiatry*, 52(7): 709–17. doi: 10.1016/j.jaac.2013.04.008

Gee, V. (1990) *Perinatal Statistics in Western Australia: Seventh Annual Report of the Western Australian Midwives' Notification System 1989*, Statistical Series, Perth, WA: Health Department of Western Australia.

Global Burden of Disease Collaborative Network (2017) *Global Burden of Disease Study 2016 (GBD 2016) Disability Weights*, Seattle, WA, US: Institute for Health Metrics and Evaluation (IHME), http://ghdx.healthdata.org/record/ihme-data/gbd-2016-disability-weights.

Hansen, M., Kurinczuk, J., Bower, C. and Webb, S. (2002) The risk of major birth defects after intracytoplasmic sperm injection and in vitro fertilization, *New England Journal of Medicine*, 346(10): 725–30. doi: 10.1056/NEJMoa010035

Hodges, S., Eitelhuber, T., Merchant, A. and Alan, J. (2019) Population data centre profile: the Western Australian data linkage branch, *International Journal of Population Data Science*, 4(2): 1-1. doi: 10.23889/ijpds.v4i2.1138.

Krijkamp, E.M., Alarid-Escudero, F., Enns, E.A., Jalal, H.J., Hunink, M.G.M. and Pechlivanoglou, P. (2018) Microsimulation modeling for health decision sciences using R: A tutorial, *Medical Decision Making*, 38(3): 400–22. doi: 10.1177/0272989X18754513

McKnight, C.M., Newnham, J.P., Stanley, F.J., Mountain, J.A., Landau, L.I., Beilin, L.J., Puddey, I.B., Pennell, C.E. and Mackey, D.A. (2012) Birth of a cohort: the first 20 years of the Raine study, *Medical Journal of Australia*, 197(11/12): 608–10. doi: 10.5694/mja12.10698

Mewton, L., Champion, K., Kay-Lambkin, F., Sunderland, M., Thornton, L. and Teesson, M. (2019) Lifestyle risk indices in adolescence and their relationships to adolescent disease burden: findings from an Australian national survey, *BMC Public Health*, 19(1): 1-10, doi: 10.1186/s12889-019-6396-y.

Moffitt, T.E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R.J., Harrington, H., Houts, R., Poulton, R., Roberts, B.W., Ross, S. et al. (2011) A gradient of childhood Self-control predicts health, wealth, and public safety, *PNAS*, 108(7): 2693–8. doi: 10.1073/pnas.1010076108

Mokdad, A.H., Forouzanfar, M.H. et al. (2016) Global burden of diseases, injuries, and risk factors for young people's health during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013, *The Lancet*, 387(10036): 2383–401. doi: 10.1016/S0140-6736(16)00648-6

Moore, T.G., Arefadib, N., Deery, A. and West, S. (2017) *The First Thousand Days: An Evidence Paper*, Parkville, VIC: Centre for Community Child Health, Murdoch Children's Research Institute.

Morris, L.J., D'Este, C., Sargent-Cox, K. and Anstey, K.J. (2016) Concurrent lifestyle risk factors: clusters and determinants in an Australian sample, *Preventive Medicine*, 84: 1–5. doi: 10.1016/j.ypmed.2015.12.009

Mountain, J.A., Nyaradi, A., Oddy, W.H., Glauert, R.A., de Klerk, N.H., Straker, L.M. and Stanley, F.J. (2016) Data linkage in an established longitudinal cohort: the Western Australian Pregnancy Cohort (Raine) Study, *Public Health Research & Practice*, 26(3): 1-5, doi: 10.17061/phrp2631636.

Newnham, J.P., Evans, S.F., Michael, C.A., Stanley, F.J. and Landau, L.I. (1993) Effects of frequent ultrasound during pregnancy: a randomised controlled trial, *The Lancet*, 342(8876): 887–91. doi: 10.1016/0140-6736(93)91944-H

Nguyen, B., Truong, L., Naidu, A., Tanvir, J., Choi, J. and Collin, P. (2021) *WH&Y Commission: 2021 Youth Health Research and Policy Priorities and Concerns*. Sydney, Australia: Wellbeing Health & Youth Centre for Research Excellence. https://apo.org.au/sites/default/files/resource-files/2021-09/apo-nid314877.pdf

Noble, N., Paul, C., Turon, H. and Oldmeadow, C. (2015) Which modifiable health risk behaviours are related? A systematic review of the clustering of Smoking, Nutrition, Alcohol and Physical activity ('SNAP') health risk factors, *Preventive Medicine*, 81: 16–41. doi: 10.1016/j.ypmed.2015.07.003

Nyaradi, A., Li, J.H., Hickling, S., Foster, J.K., Jacques, A., Ambrosini, G.L. and Oddy, W.H. (2015) A Western dietary pattern is associated with poor academic performance in Australian adolescents, *Nutrients*, 7(4): 2961–82. doi: 10.3390/nu7042961

Patton, G.C., Sawyer, S.M. et al. (2016) Our future: a *Lancet* commission on adolescent health and wellbeing, *The Lancet*, 387(10036): 2423–78. doi: 10.1016/S0140-6736(16)00579-1

Schofield, D.J., Zeppel, M.J.B., Tan, O., Lymer, S., Cunich, M.M. and Shrestha, R.N. (2018) A brief, global history of microsimulation models in health: past applications, lessons learned and future directions, *International Journal of Microsimulation*, 11(1): 97–142. doi: 10.34196/ijm.00175

Stanley, F.J., Croft, M.L., Gibbins, J. and Read, A.W. (1994) A population database for maternal and child health research in Western Australia using record linkage, *Paediatric and Perinatal Epidemiology*, 8(4): 433–47. doi: 10.1111/j.1365-3016.1994.tb00482.x

Straker, L., Hall, G.L., Mountain, J., Howie, E.K., White, E., McArdle, N., Eastwood, P.R. and Raine Study 22 Year Follow-Up Investigators (2015) Rationale, design and methods for the 22 year follow-up of the Western Australian Pregnancy Cohort (Raine) Study, *BMC Public Health*, 15: 1-16, doi: 10.1186/s12889-015-1944-6.

Straker, L., Mountain, J., Jacques, A., White, S., Smith, A., Landau, L., Stanley, F., Newnham, J., Pennell, C. and Eastwood, P. (2017) Cohort profile: the Western Australian pregnancy Cohort (Raine) Study – Generation 2, *International Journal of Epidemiology*, 46(5): 1384–85j.

United Nations Secretary-General (2015) *The Global Strategy for Women's, Children's, and Adolescents' Health (2016–2030)*, Every Woman Every Child, New York: United Nations.

White, S.W., Eastwood, P.R., Straker, L.M., Adams, L.A., Newnham, J.P., Lye, S.J. and Pennell, C.E. (2017) The Raine study had no evidence of significant perinatal selection bias after two decades of follow up: a longitudinal pregnancy cohort study, *BMC Pregnancy Childbirth*, 17: 1-10, doi: 10.1186/s12884-017-1391-8.