Kirk Wilson

Acupuncture as an Adjunct Therapy in the Treatment of Depression

Doctor of Philosophy

2014

Certificate of Original Authorship

I certify that the work in this thesis has not previously been submitted 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.

Signature of Student:

Date:

Acknowledgement

First and foremost I would like to thank Peter Meier. Even after all these years I still appreciate your offer to supervise my study. We took on a big project and you are to be commended for doing so mostly on your own. Many thanks also for your hard work, patience and encouragement along the way. I look forward to continuing working together to realise the full benefit of our efforts.

During the course of this study we lost two staff members who sat on the panel. The late Carole Rogers, who was generous with her time and assisted with the design of the study. The late Narrelle Smith who made herself available and offered friendly and helpful advice to oversee my statistical analysis. Thank you to you both, I wish you were both here to see the end result and to share my sense of achievement. Thank you to those at the university who worked behind the scenes. Rochelle Seneviratne thanks for your kindness when I needed administrative help. Not to be forgotten is the University of Technology, Sydney for giving me the opportunity to conduct the research and for the privilege of being a PhD candidate.

Thanks children Joy, Phon and Ava. I have spent an incalculable amount of hours being absent. I appreciate your understanding of the importance of my studies and your encouragement along the way. Thank you to my Mum and Dad, my extended family and friends. Your frequent and often well timed encouragement and interest in my progress assisted me to keep on track and see it through to the end.

Lastly, thank you to those who participated in the study. Without you all this study would not have been possible. I especially appreciate those who have returned to visit me after the completion of the study to express their gratitude. Your continued improvement makes the effort worthwhile.

Contents

List of Tables

List of Figures

Abstract

OBJECTIVE: The main analysis in this study assessed the effectiveness of using acupuncture as an adjunct therapy to Selective Serotonin Reuptake Inhibiting (SSRI's) antidepressants in the treatment of Major Depressive Disorder (MDD). A secondary analysis explored if the intervention led to any differences between subjects based on sex.

DESIGN: This study used a randomised, single blind, repeated measures design. A standardised acupuncture intervention was administered as an adjunct therapy to SSRI. The researchers proposed a best-fit synthesis model which upheld the integrity of the scientific method whilst maintaining the integrity of the Chinese Medicine (CM) model. Inclusion in the study required subjects to both satisfy the criteria for MDD and to present with liver qi stagnation (a CM diagnostic category). The Beck Depression Inventory and the Hamilton Rating Scale for Depressive Illness were the primary measures of depression.

RESULTS: The outcomes of the study showed that those who received the acupuncture intervention experienced a statistically significant improvement in their depression scores compared to those who participated in a wait list control group who experienced no change. Analysis based on diagnostic status (DSM-IV-TR) indicated an 87.3% remission rate. An eight week follow up analysis indicated subjects were able to maintain their improvement and remain significantly less depressed than they were before receiving the intervention. The data were stratified according to sex and suggested there were few differences between females and males. Further analysis was conducted to include an anxiety scale (STAI) and a general mental health scale (SCL-90). As with the depression analysis, the subjects showed statistically significant improvement in anxiety and mental health dimension scores. This was similarly true for the female and the male subjects alike.

CONCLUSION: Acupuncture may be an effective adjunct therapy to SSRIs for both females and males in treatment of MDD. In addition to this, the outcomes from this study have interesting implications within the wider context of the CM model. It would appear that in addition to the link between liver qi stagnation and depression, there is also a link to a broader spectrum of mental health dimensions.

Chapter 1: Introduction

1.1 The Burden of Depression

The burden of a disease is a measure of the relative impact it has on the population by considering health loss due to the disease, the injury which remains after the disease as well as the effort taken in rehabilitation and prevention (AIHW, 2014). Out of all diseases, depression has the third highest burden and is the greatest cause of non-fatal disability in Australia (AIHW, 2007). In 1997, it was estimated that one in five Australians will be affected by a mental health problem at some stage in their life (AIHW, 1997). Of all the mental health problems in Australia, depression is the most commonly reported (ABS, 1997). In 2008 it was estimated that over one million were suffering from a current case of depression. This equated to one in seven Australians experiencing depression at some stage during their life (ABS, 2009). This reflected an increase in the prevalence of depression from 6.8% to 10.3% over the preceding ten years (ABS, 2008). These data also indicate the incidence of depression differs between the sexes with Females (6.8%) having a higher incidence of depression than males (3.4%) (ABS,1998). The World Health Organisation estimates that by 2030, depression will be the greatest health concern worldwide in both developed and developing nations (WHO, 2008).With such a high prevalence, and no single modality which universally treats depression, it is clear there is a need for research which tests the effectiveness of various treatment modalities or combinations of modalities to determine the best patient outcomes for the treatment of depression.

1.2 Diagnosis of Depression

The diagnosis of depression is initially identified by the presence of criteria for a Major Depressive Episode (MDE). The most overriding symptoms of MDE are either a depressed mood for most of the day or a loss of interest/pleasure in nearly all activities. Either one of these symptoms is required to be present for at least two weeks preceding diagnosis. There are

a further seven inclusion criteria which identify subsequent symptoms of an MDE. Each of these signs and symptoms contributes to a deeper understanding of depression. These include less specific symptoms, such as sleep disturbances and changes in appetite, as well as specific symptoms such a suicidal ideation. Major depressive disorder (MDD) is characterized by one or more episodes of MDE. (APA, 2000)

The standard method for diagnosing depression is a structured interview. The Diagnostic and Statistical Manual for Mental Disorders is the gold standard in diagnosis of both categories of depression. At the time when this trial was conducted, the fourth edition with the text revision (DSM-IV TR) was the most up to date version. It outlines the diagnosis of depression via a recommended structured interview which is designed to standardise the process for mental health professionals. (APA, 2000)

The literature also notes that depression has a high comorbidity with anxiety (Dunlop, 2008). Despite this, depression and anxiety have their own unique diagnostic criteria. There are multiple diagnostic classifications of anxiety. For example Social Anxiety, Panic Attacks, Post Traumatic Stress Disorder and Obsessive Compulsive Disorder to name a few. Generalised Anxiety Disorder is a common classification of anxiety which has a highly frequent comorbidity with mood disorders such as MDD. Hence this is of particular interest to this study. Classification of this type of anxiety is characterized by at least six months of persistent or excessive anxiety or worry which occurs on more days than not. An important distinction is that the worry needs to be difficult to control. (APA, 2000)

1.3   Measurement of Depression

There are a number of scales which can be used to measure depression. The most commonly used are the Beck Depression Inventory and the Hamilton Rating Scale for Depressive Illness. Each depression scale has its own unique understanding of depression. Consequently, each will have varying sensitivities to the various symptoms of depression. It is therefore

important to include multiple depression measures when conducting a clinical trial which involves the treatment of depression. This will ensure a more informative understanding of the benefits from the treatment. Given that depression has a high comorbidity with anxiety, the inclusion of an anxiety scale would offer additional data which would be of interest. Separate anxiety scales would be required since the commonly used measures of depression have very few or no items which address anxiety. The State-Trait Anxiety Inventory for Adults is one of the most commonly used scales in clinical trials. It is structured as a global scale with two subscales. One subscale measures state anxiety or how the subjects feel at a point in time, and the other measures trait anxiety or how they feel generally.

1.4 Treatment of Depression

It is common for sufferers of depression to enter into a combination of treatments. These may include conventional or alternative therapies. Current conventional treatments can be categorised as either psychologically based or medically based. Psychological treatments focus on improving negative patterns of thinking as well as the introduction of coping strategies. There are a range of techniques which can be used to assist recovery and to improve future relapse.

Cognitive Behaviour Therapy (CBT) is considered an effective treatment for depression for some cases. It works off the premise that an individual's cognition and action influence the way they feel. Its main mechanism is the identification of the thought patterns which attribute to the depressed state. The emphasis of the treatment is to change these thoughts via the introduction of rational thinking, the removal of negative thought patterns and introducing a positive outlook based on problem solving.

Behaviour therapy is not just a component of CBT. It can be a standalone treatment for depression. The focus of this treatment is on the level of activity and pleasure in life. This involves the introduction of strategies such as

positive reinforcement, to move away from behaviours which may facilitate depression, and to consolidate healthy behaviours.

Another alternate to CBT is mindfulness based cognitive therapy (MBCT). This technique is based on mindfulness meditation and it requires keeping the mind in the present moment. The objective is to help sufferers of depression to be unaffected by external stimulus as well as from thoughts and feelings on the past and future. This technique is designed to block negative thoughts and physical sensations which are associated with depression.

These forms of therapy focus on the interpersonal relationships surrounding a sufferer of depression. They identify relationship patterns which may make someone more vulnerable to experiencing depression. The strategy of the therapist is to introduce or improve coping strategies to assist healthy social functioning.

The main conventional treatment for depression is antidepressant medication. There are a variety of classes of antidepressant drugs prescribed to those who are experiencing a moderate to severe episode of depression or if psychological treatments have been ineffective. The length of time a depression sufferer needs to take antidepressants will depend on the course of the illness.

Each type of antidepressants has different pharmacological actions. Currently, the most common type of antidepressants prescribed in Australia are Selective Serotonin Reuptake Inhibitors. These are considered by the medical community as the first choice for most depression sufferers. SSRI's are a new form of medication and they are believed to have less side effects than their predecessors, Tricyclic Antidepressants (TcAs) which are broader spectrum drugs. Side effects can still be experienced with SSRI's. Common side effects include nausea, headaches, anxiety, sweating, dizziness, agitation, weight gain, dry mouth and sexual difficulties (such as becoming and staying aroused). SSRI's block the reabsorption (or the "reuptake") of

the neurotransmitter serotonin in the brain causing the prevention of the natural reabsorption back into nerve cells. This allows the neurotransmitter to temporarily remain in the synapses (the gap between nerve cells). The increased serotonin in turn assists the brain cells to more effectively send and receive chemical messages which causes an improvement in mood.

Since there is no one treatment which can reliably treat all forms of depression there remains a need to explore further treatment options. Given it is common for depression sufferers to require a combination of treatments, it stands to reason that research into alternate or new treatments for depression would best focus on those which have the capacity to be used as adjunct to the existing treatments. This is the case for acupuncture.

1.5   Chinese Medicine and Depression

Chinese Medicine (CM) is a system of healthcare which can be used to treat a wide range of conditions. Acupuncture is a modality of treatment which falls under the banner of CM. In general terms, acupuncture involves the insertion of fine needless into specific locations on the body to elicit a therapeutic benefit. The model assumes there are channels or meridians of energy which flow through the body to function as a closed loop circuit. Most of these meridians are located superficially under the skin. Situated on specific locations along the meridians are acupoints in which acupuncture needles are inserted. Acupuncture is well suited to being used as an adjunct to other treatments. In particular, it can be used in conjunction with medication since it is not pharmacologically active by nature.

The practice of CM is open to interpretation and this has lead to the development of many schools of thought. Despite this, there is a general acceptance of protocol which combines diagnostically based theory and the practice of treatment. Acupuncture is founded on the assumption that disease is a byproduct of an underlying energetic imbalance. There are a number of possible energetic imbalances. These are classified via groupings of signs and symptoms (patterns of disharmony) which are generally labeled with names which are descriptions of the cause of the respective imbalance.

During a CM consultation the practitioner takes note of the signs and symptom experienced by the patient. These are then matched to one or more patterns of disharmony. The identification of the pattern of disharmony will in turn determine the approach or principle of treatment. From this an acupoint prescription is constructed to uphold the principle of treatment.

CM is highly idiosyncratic and there are a number of different ways it can account for depression. Schnyer (2001) identified twelve main basic diagnostic patterns and pattern combinations some of which included phlegm confounding the orifices of the heart; phlegm damp; heart/spleen deficiency and kidney deficiency. The most frequently presented in the list was liver qi stagnation. There were also three other patterns which were based on liver qi stagnation. These were liver depression transforming into heat, liver qi stagnation affecting the heart and the lung and liver stagnation with blood stasis. Schnyer (2001) did explicitly stipulate that liver qi stagnation accounted for a significant component of most cases of MDE. In response to this, the researchers in this study chose to test the hypothesis that liver qi stagnation is the pattern of disharmony which is most synonymous with the current understanding of depression.

Liver qi stagnation is an energetic imbalance which is characterised by a slowing or blocking of the energy as it moves through the meridians. Chinese Medicine theory suggests multiple possible aetiologies for stagnant liver qi, in particular emotional upset. Due to the lack of a commonly agreed standardised definition of liver qi stagnation, the researchers in this study undertook a manualisation process to obtain a consensus on which signs and symptoms best account for liver qi stagnation. The full process has been described in the method chapter, but the most common symptoms were hypochondriac stuffiness/distention, chest stuffiness/distention, sighing, irritability/frustration, plumb stone throat and a wiry pulse (note appendix B for details). Some of these are commonly identified as comorbidities by individuals suffering depression. A similar process of manualisation was undertaken to determine the acupoint prescription.

1.6   Randomised Clinical Trials

A randomised clinical trial is a method of scientific testing which is designed to add new or clarify existing knowledge. It involves the formation of a set of standardised and repeatable procedures which are designed to create an effective enquiry. RCTs are a suitable method of testing the effectiveness of a treatment in healthcare. This is most commonly done using a design of "repeated measures". Designs of this nature involve assessments both before and after an intervention to determine the therapeutic benefit of a treatment. In addition a follow up assessment may also be added. This is done some time after the discontinuation of the treatment to ascertain it's long-term benefit. Best practice in RCTs randomly assigns subjects into groups. Generally speaking one of these groups receives the treatment, and is called the "treatment group", and the other does not and is called the "control group". By comparing both groups it allows greater clarity in determining that any change is directly attributable to the treatment. This will be further discussed below. In some studies the data from each of the groups can be further divided into sub-groups, hence facilitating smaller groups within the groups to be compared to one another. This allows for a better understanding of the sub-populations.

1.7   Issues Related to Designing an RCT

There are essential considerations which need to be addressed when designing an RCT using acupuncture. The demands of the scientific model are not well aligned with the idiosyncratic nature of CM. Consequently a pragmatic approach is required which employs a synthesis model combining the best practices of the scientific method whilst maintaining the integrity of the CM model.

A key element of a suitable design is to define subjects from both disciplines via the implementation of suitable inclusion criteria. Hence subjects will need to satisfy the diagnosis of MDD from a psychological/psychiatric perceptive as well as satisfy the diagnosis of having liver qi stagnation according to CM.

As stated above, the CM model is open to interpretation. This carries inherent problems since standardisation is a necessary component of adequate design. A process of manualisation is an effective method which can be used to satisfy the need for standardisation since it is well suited to determining clearly defined principles. In essence, it is a process which arrives at a general consensus by surveying relevant sources. The researchers in this study used a manualisation process to achieve a consensus on the signs and symptoms of liver qi stagnation. This in turn served as inclusion criteria.

Another important component of design is finding a suitable method of control. These are included in a RCT to help ascertain the cause of any change over the intervention phase. It enables the researchers to determine if the benefits of an intervention are the direct cause of the treatment. A therapeutic intervention has an inherent positive bias. This is commonly known as the "placebo effect". When subjects receive what they believe to be an active treatment, there is an implicit expectation of improvement. A method of control can account for the placebo effect because it provides a base line measure of a perceived treatment. This is directly comparable to the active treatment. If the active treatment can outperform the inactive treatment (the control), then the extra improvement can be attributable to the intervention.

Most of the studies in the existing literature used non-specific acupuncture as a method of control. This either involves the needling of the skin in an area which does not contain known acupoints, or the needling of a known acupoint which is not believed to treat the condition in question. From a theoretical perspective, both these ideas are inherently problematic. The acupuncture meridians are a closed loop system which makes it difficult to exclude a therapeutic benefit. Further to this, the conception of what constitutes an acupoint is at best vague. It is therefore not possible to fully understand the potential therapeutic benefit of needling any area of the skin (despite not being located on a known acupuncture meridian). An alternative is "sham" acupuncture. This involves the use of retractable needles which do

8

not pierce the skin. This fools subject to believe they have received acupuncture when in fact they have not. It is difficult to maintain the naivety of subjects by successfully fooling them over a series of treatments. Moreover, this method raises concerns since it requires either the recruitment of those who have never received acupuncture or it will lead to other problems associated with randomisation.

The use of a wait list control group avoids the inherent problems associated with non-specific and sham acupuncture. This involves a group of subjects who do not receive the active intervention for a period of time equal to those who do. This provides a base line which can reflect on the natural rate of improvement or what can otherwise be deemed as spontaneous healing. This method of control is suitable when testing of an adjunct treatment since it answers the question if it is better to treat or not to treat. This method of control does however have its own shortcomings. It is not able to fully account for the placebo effect since the testing of acupoint specificity is not addressed i.e. testing if the therapeutic benefit is the result of the action of the acupoint.

Randomisation is another important element of design. This involves the assignment of subjects into groups based on chance. If done correctly, the law of probability will ensure the subject profiles will be the same if not similar in the treatment group as they are in the control group. This in effect assists creating homogeneity between the two groups. In turn this helps facilitate the inference that any change is attributable to the intervention.

Sound design also incorporates the use of blinding. This involves maintaining the naivety of certain aspects of the study for those who are involved. This creates a situation less conducive to biases. In its most basic form, it ensures study participants remain unaware as to if an active treatment is being received. Blinding can also involve withholding information from the participants as is the case of a wait list control, where subjects will not be informed that they are on a wait list. There are varying degrees of blinding. The least stringent form of blinding is called "single blinding". This

involves the blinding of the subjects. Further to this is "Double blinding" where both the subjects and the researchers administering the intervention remain naïve. Then there is "triple blinding" which extends the blinding to include the researchers performing the assessments.

1.8   Hypothesis

The hypothesis of the study is that acupuncture, administered as an adjunct to SSRI antidepressants to subjects who present with liver qi stagnation, will demonstrate a greater benefit in the treatment of depression compared to when SSRI's are taken alone.

The null hypothesis of this study is that there will be no added benefit when acupuncture is used as an adjunct treatment to SSRI's when treating depression in subjects who present with liver qi stagnation.

The aims of this study are to:
- develop and test a more rigorous acupuncture study design which will better determine if acupuncture has a significant therapeutic effect in the treatment of depression;
- improve on the level of understanding of how acupuncture can be shown to be an effective adjunct to SSRI's in the treatment of depression;
- demonstrate that acupuncture can be an effective adjunct to drug therapy which may help reduce drug dependency and thereby reduce the risk of side effects;
- explore possible differences between how each of the sexes respond to an acupuncture treatment for depression.

1.9   Conclusion

This study was conducted off campus in an acupuncture clinic located in the central business district of Sydney. This is a primary care, outpatient clinic which is appropriately equipped to conduct a clinical trial.

In summary, this thesis reports on a randomised, single blind, repeated measures design clinical trial which used acupuncture as an adjunct therapy to SSRI antidepressants to treat depression. The researchers devised a synthesis model to create a best-fit approach to satisfy the demands of scientific method whist maintaining the integrity of the model in Chinese Medicine. Subjects were defined as suffering from Major Depressive Disorder (MDD) and as experiencing liver qi stagnation. Those in the treatment group received the acupuncture treatment combined with SSRI's and were compared to those in the control group who only received the SSRI's. A follow up assessment was done eight weeks after the conclusion of the treatment. Analysis was based on depression inventories as well as on diagnostic status. Further analysis considered anxiety and general mental health. Data was stratified based on sex to allow females to be compared with males.

This thesis is divided into a number chapters. The first chapter is the introduction and provides the setting for the study.

A literature review is presented in chapter 2. This places this study within historical context. Consideration is made towards the difficulties associated with designing a clinical trial of this nature. The literature is reviewed in respect to the key elements of design for an acupuncture clinical trial.

Chapter 3 addresses the method used in this clinical trial. It outlines how the researchers created a synthesis model of best fit to suitably combine scientific method and the CM model. Each of the processes and procedures which were implemented to ensure standardisation are reported in turn.

Chapter 4 reports on the results. The results for the depression analysis are addressed first. This is done from the perspective of the change in symptom severity based on the depression inventories then in regards to changes in diagnostic status. Then the results from the anxiety analysis are reported followed by those from the general mental health analysis. Within each of

these areas of investigation subjects are considered as a homogenous group as well as in respect to their sex.

Chapter 5 is the discussion. This considers the outcomes from this study and discusses them in reference to the findings in the existing literature. The shortcomings of this clinical trial are also addressed. This is followed by a discussion on the implications of the findings on the understanding of liver qi stagnation as it relates to depression. The conclusions from this study are also addressed and compared to contemporaneous developments in the literature.

Chapter 6 is the conclusion. This highlights the main findings from the study and makes recommendations for further areas of research.

Chapter 2: Literature Review

Database searchers were made using Pubmed for Complementary and Alternate Medicine as well as on Google Scholar and Medline to identify the existing literature which addressed the use of acupuncture in the treatment of depression. The search parameter used on each of the databases was "acupuncture depression". All the articles which were published in English were considered and articles written in a language other than English were excluded. The search in the initial stages of this study (in the year 2007) included 457 relevant articles. In the later part of the study an additional search was carried out to include the most recent studies in the area using the same inclusion and exclusion criteria. This search in 2013 included 648 relevant studies. This was reduced down to a final number of 40 studies.

2.1 Systematic Reviews

At the time of design of this study, there were four systematic reviews which focused on studies using acupuncture to treat depression. All four considered the effectiveness of acupuncture to treat depression and offered analysis on the clinical trial designs. Mukaino (2005) and Leo (2006) applied the Jadad criteria to assess the quality and scientific rigour of the relevant literature. Leo (2006) was most critical of the existing studies claiming 56% of the 21 studies examined were of low quality. He identified multiple design issues which included small sample sizes; imprecise enrolment criteria; shortcomings with randomisation; absence of appropriate blinding; brief duration of treatment programs; and a lack of follow up assessments. Smith (2004) voiced concern over the poor application of scientific method and identified a high frequency of studies characterised by small sample sizes. Later Halbreich (2008) also expressed concern over a clear lack of quality studies and identified the need for further research using acupuncture to treat depression, but with a focus on more stringent scientific rigour.

In reporting on the effectiveness of acupuncture to treat depression, there was a consensus among the systematic reviews of insufficient conclusive evidence to suggest acupuncture is effective in the treatment of depression. Mukaino (2005) and Leo (2006) both attempted to circumvent the issue of poor scientific rigour by reviewing studies of "better quality". Despite this, both authors still found the effectiveness of using acupuncture in the treatment of depression to be inconclusive. This was mostly due to poor methods of control within the studies under review. Whilst the studies under review employed methods of control they deemed as neutral in effect, such as sham or nonspecific acupuncture, according to the CM model, these methods of control can still elicit a therapeutic benefit. This may have resulted in a comparative underestimation of the effectiveness of the acupuncture interventions. The suitability of the methods of control for acupuncture studies will be addressed later in this chapter.

Smith (2004) approached the assessment of effectiveness from a different perspective, applying selection criteria in a different manner. She focused on studies which compared active acupuncture treatments to non-specific acupuncture, wait lists, medication alone and acupuncture as an adjunct to medication. Despite the more thorough analysis, Smith was unable to determine whether medication was better at treating depression than acupuncture. She cited the poor application of scientific rigour as the main impediment and stressed there was insufficient evidence to draw accurate conclusions.

In contrast Halbreich (2008) questioned the systematic natures of the reviews by Mukaino (2005), Leo (2006) and Smith (2004). He claimed they placed too much focus on studies written in English. This resulted in the omission of multiple clinical trials published in Chinese. He argued that the articles written in Chinese, demonstrated a clear consensus suggesting acupuncture is highly effective in treating depression. Halbreich's (2008) also added that like the studies published in English, the Chinese clinical trials also suffered similar shortcomings in their application of scientific method and poor reporting

standards. Nevertheless, despite these shortcomings, Halbreich (2008) believed the Chinese clinical trials need to be noted. There are however no English translations of these articles available and this chapter will not independently assess the articles written in Chinese.

Based on the systematic reviews it is evident there remains a need for further research into the use of acupuncture to treat depression. In order to fill the gaps in the existing literature, future studies require the application of sound scientific method and need to utilise more detailed reporting standards. This will allow the effectiveness of using acupuncture to treat depression to be assessed more accurately and in a more clinically meaningful manner.

2.2   Designing Acupuncture Clinical Trials

When designing a clinical trial which uses acupuncture to treat depression there are two primary considerations. Firstly, there is the need to account for depression from a Western psychological/psychiatric perspective and to align that interpretation to the Chinese Medicine (CM) model. Secondly the inherent difficulties when designing any acupuncture clinical trial must be addressed. These difficulties are associated with the incompatibility between the CM model and the demands of the scientific method.

There is a clear consensus as to how the psychological/psychiatric communities have defined depression with well-established psychological diagnostic inclusion and exclusion criteria. This is however not as clear when depression is considered from a CM perspective. There is a greater scope for interpretation within the CM model. It focuses on the underlying energetic imbalances which can lead to depression. Whilst both models have a similar understanding of depression and its accompanying signs and symptoms, in the CM model there are differing weights of importance placed the various symptoms. This is what

enables the CM model to have the sensitivity to identify the various energetic imbalances which can lead to depression and hence offer differential diagnoses.

The studies in the literature have not addressed CM's scope for differential diagnosis very well. This remains an area for improvement when designing acupuncture RCTs for depression. Few studies in the literature made an explicit attempt to provide an account for depression (via a differential diagnosis) according to the CM model. Instead, the majority of studies treated depression sufferers as a heterogeneous group. Even though many of these studies have made attempts to uphold the idiosyncratic nature of the CM model, there remains a lack of clarity for this reason. Consequently, much of the theory within how the CM model accounts for depression is yet to be scientifically tested. It is yet to be determined which differential diagnosis can best account for depression. In turn, a consensus is yet to be achieved as to how to best treat depression using acupuncture. As a result the studies in the literature lack clinical relevance. These issues were further exacerbated by the implementation of poor reporting standards.

In any acupuncture clinical trial there are difficulties due to the incompatibility of the CM model and the demands of scientific method. Two of the most pressing problems surround finding a suitable method of control and implementing blinding. The studies in the literature used a range of controls. These included non-specific acupuncture, antidepressant medication and wait list groups. These models will be examined to determine which method of control best upholds the integrity of both models for the purpose of this study.

This chapter will further explore the relevant literature to address each component of the study design. Inclusion/exclusion criteria from a psychological perspective will be examined. Subject demographics will be outlined followed by the use of psychometric measures. The implementation of the Chinese Medicine model will also be addressed with reference to the use of diagnosis according to

patterns differentiation, treatment programs and acupoint prescriptions. Finally the use of blinding suitable methods of control will be discussed. The findings from the literature will be used to inform the design of this study.

2.3 Psychological Diagnosis / Inclusion Criteria

Identification of a target population is an important requirement of any study. This is especially the case when a study involves the recruitment of a clinical population. It is imperative to employ valid inclusion criteria. Given the line of enquiry, the studies in the literature were concerned with recruiting subjects who were clinically depressed. There was a clear consensus as to the preferred method to determine an adequate level of depression in order to meet inclusion criteria. The primary requirement was for subjects to satisfy the structured interview for the diagnosis of either a Major Depressive Episode (MDE) or Major Depressive Disorder (MDD) as outlined in the Diagnostic and Statistical Manual for Mental Disorders (DSM). There have been a number of different editions of the DSM. Each study in the literature used the most up to date version for the DSM at the time of recruitment. Since most studies were recent, DSM IV was the most commonly used edition (Allen, 1998; Allen, 2000; Gallagher, 2001; Blitzer, 2004; MacPherson, 2004; Manber, 2004; Allen, 2006;). Two studies used the previous edition, the DSM III (Luo,1985 and Roschke, 2000).

Within the literature there is also an identifiable trend of using scores from depression scales as an inclusion criteria. The use of depression scales requires subjects to score over a certain level in order to meet inclusion criteria. The Hamilton Depression Inventory was the most frequently used scale in this fashion (Luo (1985), Luo (1990), Yang (1994), Luo (1998), Roschke (2000), Manber (2004), Han (2004) and Allen (2006). There were less frequently employed scales. Luo (1985) used the National Survey and Coordination Group of Psychiatric Epidemiology. Luo (1990) used the Present State Examination (PSE) and the Standardized Assessment of Patients with Depressive Disorders

(SADD). Luo (1998) used the Depression criteria from the Huangshan symposium in China (1981). The Beck Depression Inventory was used as inclusion criteria by Quah-Smith (2005).

There was clear consensus in the literature that the DSM is the most widely accepted inclusion criteria. The researchers in this study will adhere to this standard and will use the DSM –IV-TR to diagnose MDD (single episode) as inclusion criteria. This will allow for direct comparison to the majority of studies in the literature.

The diagnosis of MDD, in a simplified form, requires subjects to be either:
- feeling depressed for most the day (this can be a subjective account or observed by others); or
- have markedly diminished interest or pleasure in activities for most of the day nearly every day.

In addition to the above, a combination of the following symptoms needed to be met so that a total of at least five symptoms were present.
- Significant loss of weight outside of dieting;
- Insomnia or hypersomnia nearly every day;
- Psychomotor retardation or agitation nearly every day;
- Fatigue or loss of energy nearly every day;
- Feelings of worthlessness or excessive or inappropriate guilt nearly every day;
- Diminished ability to think/concentrate or indecisiveness nearly every day; and
- Recurrent thoughts of death or suicide ideation, or suicide attempt.

In order to meet each of the criteria, the symptoms need to cause clinically significant distress or impairment in social, occupational, or other important areas of functioning. (APA, 2005)

2.4 Psychological Exclusion Criteria

In line with the use of the DSM as inclusion criteria, the same studies in turn used the DSM as exclusion criteria. (Luo, 1985; Allen, 1998; Allen, 2000; Roschke, 2000; Gallagher, 2001; Blitzer, 2004; MacPherson, 2004; Manber, 2004; and Allen, 2006)

In summary, exclusion from the diagnosis of MDD occurs if:

- there has never been a manic episode or the criteria of a mixed episode;
- the symptoms can be attributed to underlying medical condition or substance abuse;
- the symptoms can be attributed to bereavement; and
- the MDD is not better accounted for by schizoaffective disorder or it is not superimposed on schizophrenia, schizophreniform disorder or delusional disorder. (APA, 2005)

Other exclusion criteria were used in the literature. Pregnant women were excluded from the Gallagher (2001), Allen (2006) and Quah-Smith (2005) studies. Depression lasting for more than two years was also excluded by Quah-Smith (2000) and Allen (2000). Subjects with wound healing disorders or bleeding disorders were excluded by Roschke (2000) and MacPherson (2004). Those taking additional herbal supplements were excluded from Quah-Smith (2000). Those taking psychotropic medication were excluded from Blitzer (2004).

Despite these variations, most commonly the exclusion criteria as outlined in the DSM accounts for the greater proportion of criteria. This creates a consensus with the literature that best practice uses the criteria in the relevant edition of the DSM. This stands to reason since the DSM clearly defines what does and does not constitute a diagnosis of MDD. In order to uphold best practice and to maximise comparative relevance to the existing literature, the researchers in this study will use DSM-IV-TR's exclusion criteria.

2.5 Subjects

As identified in the systematic reviews, most studies had poor reporting standards. Often the population from which the sample of subjects was drawn was not reported. Then once selected, vital and important demographic information such as age, sex and method of recruitment were often not reported.

2.5.1 Number of Subjects

All studies reported how many subjects were involved in each of their trials. There were however instances when the number of subjects which were assigned to the respective treatment groups were omitted. Nevertheless, the most noteworthy aspect was the small sample sizes in many of the studies as shown in Table 1.

Table 1 highlights the concerns reported by Smith (2004) and Leo (2006). The only studies in the literature which clearly included suitable sample sizes were Allen (2006), Luo (1990) and Luo (1998) Phase 2. The studies with 33 or less subjects carry a question over adequate statistical power. Once the subjects get divided into two groups, the size of each group is small. This is especially of concern if the results are not characterised by a statistically large effect size. Small group sizes require large effect sizes to ensure valid and meaningful outcomes. The studies with over 60 subjects have a more suitable sample size. Though it must be noted that the design in the Manber (2004) and the Roschke (2000) studies required subjects to be divided into three groups. This again meant fewer subjects were allocated to each of the groups. This again brings into question the adequate statistical power. The Blitzer (2004) and MacPherson (2004) studies should be taken within the context that they are pilot studies and ten subjects may be a suitable number..

Table 1: Number of subjects in the studies in the literature

Blitzer (2004)	10
MacPherson (2004)	10
Luo (1998) Phase 1	29
Han (2004)	30
Quah-Smith (2005)	30
Allen (1998)	33
Allen (2000)	33
Gallagher (2001)	33
Yang (1994)	41
Luo (1985)	47
Manber (2004)	61
Roschke (2000)	70
Allen (2006)	151
Luo (1990)	241
Luo (1998) Phase 2	241

2.5.2 Longevity of Depression Suffered

Few of the studies reported the length of time the subjects involved in their clinical trails had been suffering from depression. This can however be implied by the choice of inclusion criteria. For example, the diagnosis of MDE would suggest the depression would be of a more short-term nature compared to the diagnosis of MDD. Though this is not a very robust way to account for the longevity of the depression.

Some studies in the literature did report how long subjects were depressed before entering the study. Luo (1985) reported the course of depression for subjects to be between one month to two years. Luo (1990) reported the mean depressive episode of the subjects in that study was 5.4 months. Han's (2004)

study involved subjects who had been diagnosed with depression from between two months and 15 years before entry into the study, with current episodes of depression reported to be between two months to three years in duration. These large time ranges as reported by Han (2004) and Luo (1990) serves as examples of the problematic nature when only means are reported. In such cases, the mean can be misleading and/or uninformative.

In order to fully understand the sample of the depressed population it is imperative that descriptive statistics are reported on the longevity of the depression experienced by the subjects as well as any statistical means. This provides a valuable insight into the nature of the depression experienced by the subjects. It may also provide clinical insights which may suggest there are subgroups of the depressed population who respond more favourably to acupuncture treatment.

2.5.3   Other Classifications and Severity of Depression

Not all the studies in the literature used the DSM as the focus of their diagnosis of depression. Lou (1985), Lou (1990) and Han (2004) for example, recruited a sample of depressed subjects who were inconsistent to those used in the other studies. Out of the 30 subjects who were recruited by Lou (1985) 14 were reported as depressive, 11 were involutional depressive and four had reactive depression. Thirty of the subjects in the Han (2004) study and 193 subjects in Lou (1990) were reported to have bi polar disorder, which according to the DSM, is a different diagnosis of depression to that of MDD. Roschke (2000) reported all subjects were severely depressed, and despite using DSM as inclusion criteria, he reported 18 subjects as having bi polar disorder, ten having involutional depression and 13 having depressive neurosis. The inclusion criteria used in Roschke (2000) however, was extended beyond the inclusion of only those with an MDD.

As with the above demographical information it is equally important the exact classification of depression experienced by the patients is reported. Again, this provides valuable information on potential subgroups of the depressed population who may respond more favourably to acupuncture treatment. The lack of differentiation of subgroups within a study population confounds any potential clinical conclusions which may be drawn. To improve the clinical application of research, more effort will need to be made in future studies to address a more homogenised group of depressed patients.

2.6   Differentiation Based on Sex

No studies in the literature made comparisons based on sex. There were no studies that examined the use of acupuncture for depression in male populations. There were a group of studies which did however use acupuncture to treat depression in females. Allen (1998), Allen (2000) both compared an acupuncture intervention to non-specific acupuncture and a wait list group. Gallagher (2001) considered the relapse rates in females who had previously undergone acupuncture treatment. In contrast, Manber (2004) more specifically focused on treating pregnant women.

Allen (1998) found acupuncture to be effective in reducing depression amongst females. At the conclusion of the study he reported that 64% of all the subjects who took part, experienced a full remission according to DSM-IV criteria. This included subjects who received active acupuncture, non-specific acupuncture and those who were in a wait list group. According to the Hamilton Depression Scale (Ham-D) and the Beck Depression Inventory (BDI), symptom reduction was found to be greater amongst the females who received the active acupuncture intervention compared to females who received a non-specific acupuncture intervention ($p < 0.05$). Though, less encouraging were the results when those who received the active acupuncture intervention were compared to those in the wait list group. Ham-D analysis on this comparison indicated no

significant difference (p < 0.12). No BDI analysis was offered to support or refute this outcome. From the statistics reported, it appears the Allen (2000) used the same data as Allen (1998). Both studies reported the same Ham-D mean and standard deviation of Ham-D scores. So it follows that Allen (2000) also found that an active acupuncture treatment was more effective at treating depression in females than a non-specific acupuncture treatment or participating in a wait list control group.

Manber (2004) recruited depressed female subjects who were also pregnant. She found a response rate of 68.8% amongst those who received an active acupuncture treatment. Comparisons were made which tested the relative effectiveness of the acupuncture intervention to a non-specific acupuncture intervention and those who received massage therapy. Results on the Hamilton Rating Scale for Depressive Illness (Ham-D) indicated no significant differences between the active and non-specific acupuncture groups (p = 0.18). This was supported by the Beck Depression Inventory (BDI) analysis which indicated no significant difference in both the first month of treatment (p = 0.092) and in the second month of treatment (p = 0.115). Active acupuncture was found to be significantly more effective than massage therapy using the Ham-D (p = 0.031). This was supported by the BDI analysis in the first (p = 0.047) only. In the second month the BDI found no difference between the active acupuncture and the massage therapy (p = 0.83). Manber (2004) concluded that acupuncture was effective in treating depression. Though she was not able to attribute this effect to the CM model since the active acupuncture treatments were found to be no different to non-specific acupuncture. Moreover, the active acupuncture barely outperformed a massage treatment.

Results which have focused on depression in females have yielded mixed outcomes with the results of Allen (1998) and Manber (2004) conflicting. Allen (1998) did offer some encouraging outcomes, but this was only limited to the comparison with the non-specific acupuncture. As no exact P scores were

reported, the degree of the statistical significance cannot be determined. Couple this with an insignificant finding of Manber (2004) it is evident that there is insufficient evidence to determine the effectiveness of acupuncture to treat depression in females. More importantly, the review of the literature highlights the clear lack of consideration given to sex differences in acupuncture studies for depression.

This study will seek to add to the knowledge on the effectiveness of acupuncture in the treatment of females with depression. In addition, it will address the gap in the literature where males are yet to be tested separate to females. Through stratifying data based on sex, this study will be the first to test and report the effectiveness of acupuncture treatment within independent groups of depressed female and male subjects.

2.7 Measures of Depression

The most commonly used depression measure in the literature was the Hamilton Rating Scale for Depression Illness (Ham-D) (Hamilton, 1967). This scale is both a reliable and valid measure of depression. It is an older measure which may account for its more frequent use amongst the earlier studies. The Beck Depression Inventory (BDI) (Beck, 1987) is the most frequently used measure of depression in more recent studies. This is considered to be a more contemporaneous scale which better reflects current views on depression. Table 2 below shows the frequency of use of various scales in the literature.

Some of the studies conducted in China used an alternate method to quantify their results. The Chinese Grading System was used by Luo (1990) and Luo (1998). This system rates changes in subjects according to the categories of improved markedly, improved and failed. Consequently this scale is not able to be used for statistical analysis.

Table 2: Depression measures used in the literature.

	Ham-D	BDI	ASBERG	CGI	HADS	GAS	BRMS	RDS
Lou (1985)	X			X				
Lou (1990)	X		X					
Yang (1994)	X							
Allen (1998)	X	X						
Luo (1998)	X		X	X				
Allen (2000)	X							
Roschke (2000)				X		X	X	
Gallagher (2001)	X							
Blitzer (2004)		X						X
Han (2004)	X		X					
MacPherson (2004)		X			X			
Manber (2004)	X	X						
Quah-Smith (2005)		X			X			
Allen (2006)	X	X						

Ham-D: Hamilton Rating Scale for Depression Illness

BDI: Beck Depression Inventory

Asberg: The Self-Rating Depression Scale

CGI: The Clinical Global Impressions Scale

HADS: The Hospital Anxiety and Depression Scale

GAS: Global Assessment Scale

BRMS: Bech-Rafaelsen Melancholia Scale

RDS: Reynolds Depression Survey Inventory

Interestingly, there is a trend in the literature towards using multiple scales as evident from Table 2. Only three studies, Yang (1994), Allen (2000) and Gallagher (2001) used a single measure, the Ham-D. The remainder of the

studies used two or more measures. This may have been done to gain a broader understanding of depression and its sensitivity to acupuncture treatment. Not all the scales used in the literature were measures of depression. Other scales were included to add further information which is not available in the depression scales. For example, Luo (1985) used an Antidepressant Side Effect Rating Chart to gauge the side effects of the antidepressants in the control group.

The researchers in this study will use the BDI as the primary scale, as it is widely considered to be the most relevant measure of depression in a clinical setting. Compared to the Ham-D, it is the most up to date scale and better reflects the recent expert understanding of depression. It also has the advantage of being an inventory which implicitly means it is a questionnaire which is completed by the subjects. This removes the need to have an expert assessor. In order to maintain relevance to other studies in the literature, this study will additionally administer the Ham-D. This approach is consistent with Allen (1998), Manber (2004) and Allen (2006) who all used both the BDI and the Ham-D. The use of both measures will increase the comparability of this study to that of the greater majority of the studies in the literature.

In addition to these two main depression measures, this study will also use the State Trait Anxiety Index (STAI) (Speilberger, 1983) and the Systematic Checklist 90 (SCL 90) (Derogatis, 1977). The STAI is a measure of anxiety. It is commonly understood that anxiety and depression have a high incidence of co-morbidity (Dunlop, 2008). The SCL 90 is a general mental health inventory which will provide insight into symptomatology and psychological distress. The STAI and the SCL 90 are being included to allow for a more thorough reporting of psychological symptoms and an analysis of how they might be associated with liver qi stagnation.

2.8   Chinese Medicine Diagnosis

In the paradigm of CM, therapeutic processes are always relative. An acupoint prescription is only meaningful relative to the principle of treatment. This in turn is only meaningful relative to a CM diagnosis of pattern differentiation or energetic imbalance within the body. Despite this being an integral aspect of CM practice, few studies in the literature adequately reported which differential patterns of disharmony were presenting in their subjects. Surprisingly, some studies made no attempt at all.

The studies which made no mention of CM differential diagnosis were the studies of Luo (1985), Luo (1990), Luo (1998), Roschke (2000), Gallagher (2001) and Allen (2006). The Allen (1998) and (2000) studies along with the Manber (2004) and MacPerson (2004) did claim to have made a diagnosis based on CM pattern differentiation. Though, they did not specifically report on the diagnoses which were included in the study. Consequently, no conclusions could be drawn as to which CM patterns of disharmony are more likely to be associated with depression nor which patterns could best be treated to reduce depression.

Other studies did make some attempt to report on CM differential diagnosis. Blitzer (2004) based treatments on the extraordinary vessels which involved a CM diagnosis accordingly based on the extraordinary vessels. Despite reporting so, Blitzer (2004) did not report any of the specific details used within the method of using the extraordinary vessels diagnostically. For example, the range of differential diagnoses were not reported nor tallied. Quah-Smith (2005) did use a CM diagnosis but did not report on the method in which the diagnosis was determined. It was only reported that most of the included subjects were treated for either liver qi stagnation or liver qi deficiency. No stratification of data was offered which considered the results for each respective pattern of disharmony. This may be of interest since both of these patterns, although

associated with the liver, are different by nature. One is primarily characterised as an excess condition and the other a deficient condition. This would have provided an interesting insight for how the CM model accounts for depression.

The most thorough reporting of CM diagnosis was done by Yang (1994) and Han (2004). They reported the incidence of each of the respective diagnoses. Han (2004) reported 19 cases of liver and spleen qi deficiency; 27 cases of heart and spleen deficiency and 15 cases of liver and kidney yin deficiency. Yang (1994) reported the diagnoses of liver qi stagnation with spleen deficiency; liver blood stasis; heart and spleen deficiency and spleen and kidney yang deficiency. Even in these studies however, they did not report on systems to ensure standardisation of the diagnosis. Like Quah-Smith (2005), both Yang (1994) and Han (2004) did not stratify their results to determine if any of the patterns of disharmony could be more effectively treated to reduce depression.

When taking a critical stance, the above studies have an apparent absence of an acupoint prescription being based on a principle of treatment to rectify a pattern of disharmony. This renders any reporting of outcomes clinically uninformative. Hence, improved adherence to detail within the CM model is needed in future research. To add knowledge in a meaningful manner, there is the requirement to first determine which of the CM patterns of disharmony are more closely associated with depression. There should be a reported, standardised method to determine the particular pattern. Once this is achieved it is necessary that each pattern of disharmony be independently treated with a corresponding principle of treatment. This will provide insight as to which pattern of disharmony can be best treated to reduce depression. In turn, this will more thoroughly and more correctly (in terms of the CM model) determine which subgroups of the depressed population can be effectively treated with acupuncture.

Given the patterns of disharmony which were reported by Yang (1994), Han (2004) and Quah-Smith (2005) it appears that liver qi stagnation is the pattern of disharmony which is most frequently associated with depression. This is further supported in Schnyer's (2000) theoretical CM account for depression. Following this, the researchers in this study will focus on depression sufferers who satisfy the diagnostic criteria of liver qi stagnation and related sub-patterns. The acupoint prescription can then be correctly based on a core and relevant principle of treatment. This will facilitate the implementation of best practice according to the CM model which in turn provides greater clinical relevance to the outcomes.

2.9   Acupuncture Treatment Programs

The literature provided a general consensus on the most suitable treatment program for acupuncture to treat depression. The duration of treatment programs was generally for an eight week period. It constituted two treatments per week for the first four weeks, then one treatment per week in the remaining four weeks. This protocol was utilised by Allen (1998) Blitzer (2004), Manber (2004), Quah-Smith (2005) and Allen (2006). An eight week treatment program was also used by Allen (2000) however the timing of the treatments was not reported. Though as previously suggested it appears Allen (1998) and Allen (2000) were based on the same data given they share the same descriptive statistics and statistical analysis. Gallagher (2001) focused on follow up assessments but the treatment program which preceded the follow up period, was consistent with an eight week period.

Similarly, the studies which were conducted in China had a general consensus to a treatment program but used a different approach. This approach was exemplified by the electro-acupuncture studies of Luo (1990), Luo (1998) and Han (2004). Their treatment program consisted of a six week treatment program which included six treatments per week. During each treatment the needles

remained in situ for 45 minutes. Luo (1985) used a similar protocol but differed in that the needles remained in situ for one hour and the program lasted for only five weeks. The Yang (1994) study also used the five week treatment program but did not report how long the needles remained in situ.

Some studies used unique approaches. MacPherson (2004) adopted a pragmatic trial approach and left the treatment program to the discretion of the acupuncturist who administered up to ten treatments either once or twice a week. Roschke (2000) administered three treatments a week for four weeks with the needles remaining in situ for 30 minutes. These studies appeared to be outliers in the literature.

Due to the poor standards of reporting, there is insufficient information on treatment protocols. The STRICTA recommendations highlight the integral importance in upholding the standards of scientific reporting. In order to duplicate a treatment program all relevant information needs to be reported. This includes reporting the rationale for the acupuncture intervention, details of the needling used, treatment regimen and the background of the practitioner (MacPherson, 2001).

Complete replication of the treatment programs used in the literature is not possible. Each of the studies fails to report one or more of the basic STRICTA requirements. Although the number of treatments may have been reported, the necessary detail such as the timing between treatments, and the number of treatments as well as the time the needles remain in situ was omitted. This poor standard of reporting further highlights the concerns as discussed in the systematic reviews (Smith 2004; Makaino 2005; Leo, 2006 and Halbreich, 2008).

Consistent with the general consensus of the literature, this study will implement an eight week treatment program. Based on the clinical expertise of the

researchers, two treatments a week for the first four weeks, and then one treatment a week for the remaining four weeks will be administered. Not all studies in the literature used this treatment program but it was used with the greatest frequency compared to the other reported treatment programs. The study will aim to improve the reporting standards and implement the STRICTA recommendation. Full disclosure will be upheld. This will include the reporting of the timing of the treatment sessions and the time the needles remained in situ, as well as reporting the needling technique and the underlying acupuncture rationale. This will enable more accurate future replications of the treatment protocol which will be used in this study.

2.10 Acupuncture Prescriptions

The reporting of acupoint prescriptions in the literature also did not satisfactorily adhere to reporting standards. Some studies failed to report any detail on acupoint prescriptions. Others did report which acupoints were used, but did not give an account as to the reasoning behind their selection. Then there were studies which provided the required supportive details by either providing a differential diagnosis or principle of treatment to justify the acupoint prescription.

Allen (1998) and Gallagher (2001) reported the least information on acupoints. Both studies did not report which acupoints were used, nor did they report a CM diagnosis. Allen (2000), MacPherson (2004), Manber (2004) and Allen (2006) reportedly based their acupoint prescriptions on differential diagnosis. They however did not explicitly report the acupoints nor did they state on which differential diagnoses the acupoints were based. Blitzer (2004) also omitted to report which specific acupoints were used, but did offer more information by reporting the prescriptions were based on the extraordinary channels which limits the number of potential acupoints to a total of eight.

Other studies in the literature did report on acupoints but did not put the prescription into context by reporting on which CM principle of treatment they were based. The Luo studies of (1985), (1990) and (1998) used electro-acupuncture between Governing Vessel 20 (Baihui) and Yin Tang (extra) and he provided no CM justification for this prescription. Given the acupoints, a sophisticated reader could infer the principle of treatment was to calm *shen.* This assumption, although highly probable, may not be correct. Luo did however provide more detail on the technique used for his treatment protocols. He reported the needling was done obliquely towards each other and the electric current was at 2Hz with between 3.0 and 5.0 mA which was strong enough to allow for slight twitching of the skin. Roschke (2000) administered an acupoint prescription which included Urinary Bladder 15 (Xinshu), Urinary Bladder 17 (Geshu), Heart 7 (Shenmen), Pericardium 6 (Neiguan), Stomach 40 (Fenglong), Spleen 6 (Sanyinjiao), Spleen 5 (Shangqiu) and Lung 1 (Zhongfu). Like Luo, Roschke (2000) did not justify this prescription by providing the CM diagnosis. Although these studies did report on the acupoint prescriptions, their failure to report on CM diagnosis is still problematic as it is unclear, from a CM perspective, which subset of the depressed population was treated.

There were studies in the literature which did adhere to better reporting standards. Along with the acupoint prescription they also made some attempt to put into context their selection by either reporting the CM diagnosis or the principle of treatment on which the acupoints were based. Yang (2006) used the extra channels due to their relationship with the heart and the brain. All subjects received Governing Vessel 24 (Shenting), Governing Vessel 20 (Baihui) , Governing Vessel 14 (Dazhui), Governing Vessel 12 (Shenzhu), Conception Vessel 17 (Shanzhong), Conception Vessel 14 (Juque), Gall Bladder 20 (Fengqchi) & Pericardium 6 (Neiguan) as a core treatment. Electro-acupuncture was used on Governing Vessel 24 (Shenting), Governing Vessel 20 (Baihui) and Gall Bladder 20 (Fengchi). Then subjects who presented with liver qi stasis & spleen deficiency received Stomach 36 (Zusanli), Spleen 6 (Sanyinjiao), Liver 3

(Taichong) in addition. Those with liver blood stasis also received Large Intestine 4 (Hegu), Liver 3 (Taichong), Spleen 10 (Xuehai). Those with Heart and spleen deficiency also received Heart 7 (Shenmen), Pericardium 7 (Daling), Spleen 6 (Sanyinjiao) and Stomach 36 (Zusanli). Those with spleen and kidney yang deficiency also received Kidney 3 (Taixi), Spleen 3 (Taibai), Spleen 6 (Sanyinjiao), Stomach 36 (Zusanli), and Conception Vessel 4 (Guanyaun).

The idea of a core set of acupoints with the additions of variation to match a principle of treatment was also used by Han (2004). In that study, as a core prescription, all subjects received Governing Vessel 20 (Baihui) and Yin Tang (extra) with an electric current being applied between the two points. Then additional points were administered based on differential diagnosis. Those with liver qi and spleen deficiency received Gall Bladder 34 (Yanglingquan) and Spleen 6 (Sanyingjiao). Those with heart and spleen deficiency received Pericardium 6 (Neiguan) and Spleen 6 (Sanyinjiao). Those with liver and kidney yin deficiency received Kidney 3 (Taixi) and Spleen 6 (Sanjinjiao).

Quah-Smith (2005) also administered a core set of acupoints to all subjects. These six primary points were Liver 14 (Qimne) on the right hand side only as well as Conception Vessel 15 (Jiuwei), Conception Vessel 14 (Juque) , Heart 7 (Daling), Liver 8 (Ququan) (left). Kidney 10 (Yingu) was additionally given if there were signs of anxiety. An additional set of acupoints which were based on differential diagnosis were also administered. These acupoints were however not reported.

The poor reporting standards employed in the literature in regards to acupoint selection does have consequences. It has not allowed for identification of a consensus on which acupoints (relative to the respective patterns of disharmony) can best be used to treat depression. Consequently this study cannot effectively replicate acupoint prescriptions identified in the literature. This limits comparisons to other studies. There are however good elements of design

within the literature. The more recent studies of Han (2004), Quah-Smith (2005) and Yang (2006) introduced the notion of a core set of points with the inclusion of additional points which were based on the principle of treatment matching the differential diagnosis. This approach better maintains the integrity of the CM model and is more closely aligned with clinical practice. This study will adopt this approach and attempt to construct and test an acupuncture prescription which may better account for the subtle variations of liver qi stagnation. As a result, this study will be the first to administer an acupoint prescription based on a single principle of treatment with a protocol that can be replicated by future studies. This may serve as a starting point for a future consensus.

2.11 Methods of Blinding

Best practice in scientific method requires the implementation of blinding within a clinical trial. The use of blinding reduces the likelihood of bias. In essence it maintains the naivety of those involved in the study in such a manner as to facilitate the integrity of the clinical trial. The blinding of subjects is considered the most basic necessity and is referred to as single blinding. This entails procedures aimed at preventing subjects from differentiating between inactive and active treatment. This allows for effective comparison between treatment (active) and control (inactive) groups to determine the effectiveness of an intervention. Blinding can be also applied to the experimenters (i.e. to those who administer the treatment as well as on those who perform assessments) and it is referred to as double blinding since it is done in addition to single blinding. This again involves withholding the knowledge as to which group of subjects are receiving active versus inactive intervention. It may also be applied so that experimenters remain unaware of the aim of the study and of what is being measured. This will reduce the likelihood of the experimenters from interacting with the subjects in a manner which may affect the outcomes.

Despite being an integral aspect of reporting standards, only about half of the studies in the literature reported on what efforts were made to implement suitable blinding. These were the studies by Luo (1998), Allen (2000), Roschke (2000), Gallagher (2001), Manber (2004), Quah-Smith (2005), and Allen (2006). With the exception of Roschke (2000), all the other studies employed standards of double blinding. This involved the implementation of varying techniques to blind the subjects, treating practitioners and the assessors. Evidently, the subjects, assessors and treating acupuncturists were unaware of which groups the subjects were allocated. Only the head researcher, who in each instance did not participate in the study, knew of the allocation of subjects. In addition to this, those who performed the assessments were kept separate to those who administered the acupuncture. Roschke (2000) used a similar design. The only difference was that the treating acupuncturists were aware of which group the subjects were assigned. Such practices are questionable since it is not clearly established that patients can be effectively blinded to an acupuncture treatment, or that practitioners with defined knowledge could administer a treatment without attempting to determine its effectiveness or validity.

The design of this study will uphold the requirement for single blinding. All though all involved will receive an active treatment, the subjects in the wait-list group still taking their standard medications will not be informed that they are being used as a control group. This will allow for a valid comparison of acupuncture as an adjunct therapy. The requirement of double blinding will not be maintained. Double blinding is inherently problematic for acupuncture clinical trials. It is impossible to blind the acupuncturist who is administering the treatments. The intended principle of treatment can be easily revealed by the choice of acupoints. In turn, the experimenters can only be blinded from the allocation of subjects into groups. This study is being carried out for the purposes of a Doctor of Philosophy. Consequently the student will be the researcher as well as the treating acupuncturist and will be aware as to which group each of the subjects will be assigned. This will be noted as a limitation of

this study. To an extent this bias will be minimised via the use of self-reporting inventories such as the BDI, STAI and SCL90 which are completed by the subjects.

2.12 Acupuncture Interventions

Five of the studies in the literature offered analysis in which the effectiveness of the acupuncture intervention was assessed independent to a control group (within group analysis). This means these studies consisted of one treatment group which received the acupuncture intervention. Two of these were the pilot studies by Blitzer (2004) and MacPherson (2004) and another two were the electro-acupuncture studies of Luo (1985) and Han (2004). Yang (1994) used the extra channels to independently test the effectiveness of acupuncture to treat depression. In addition to these five studies was Quah-Smith (2005) who used laser as a form of active acupuncture intervention. The other studies in the literature did not offer an independent measure the effectiveness of the acupuncture. They instead only provided statistical analysis of the acupuncture intervention relative to a control group (between group analysis).

Blitzer (2004) and MacPherson (2004) offered encouraging results which suggested acupuncture is an effective treatment for depression sufferers. Both studies had clear significant results comparing pre and post intervention scores using the BDI. The outcome was more significant for Blitzer (2004) ($p = 0.001$) than for MacPherson (04) ($p = 0.03$). Further to this, Blitzer (2004) found all six participants who completed the study experienced a full remission as per DSM IV criteria. Blitzer (2004) also administered the Reynolds Depression Survey Inventory, which confirmed acupuncture's effectiveness in treating depression ($p = 0.003$). MacPherson (04) additionally administered the HADS which again indicated significant improvement in depression scores on its depression subscale ($p = 0.04$).

The electro-acupuncture studies by Lou (1985) and Han (2004), as well as Yang's (1994) extra channel study, all supported the findings of Blitzer (2004) and MacPherson (2004). Analysis using the Ham-D across all three studies indicated significant reductions in depression scores as the result of an acupuncture intervention ($p < 0.01$). Luo (1985) also used the CGI and found the same outcome when considering the severity of the disease ($p < 0.01$) and the general progress of the treatment ($p < 0.01$). These outcomes are consistent with Quah-Smith (2005) who's analysis using the BDI indicated that a laser acupuncture intervention was also effective in reducing depression scores ($p = 0.006$).

Despite the shortcomings with the studies in the literature, there are grounds to suggest there is a rough consensus which indicates that acupuncture may be an effective treatment for depression. This is encouraging from a clinical perspective. Moreover, it suggests further research to better understand how acupuncture can treat depression is warranted. Sound scientific method however demands the use of suitable controls. The inclusion of a method of control will clarify if the effectiveness of the treatment is attributable to the use of acupuncture as used within the CM model. A consensus over suitable methods of control for an acupuncture clinical trial is however yet to be achieved. Conventional methods of control are problematic for acupuncture clinical trials. This challenge remains when constructing a clinical trial in which acupuncture is used to treat depression.

2.13 Methods of Control

Throughout the literature, studies which have used acupuncture to treat depression have utilised a range of methods of control. The effectiveness of active acupuncture treatments have been compared to the effectiveness of non-specific acupuncture, antidepressant medication and no treatment as in the case for wait list control groups. The use of a wait list as a method of control has in

turn facilitated the use of acupuncture as an adjunct treatment to antidepressant medication.

2.13.1 Non-Specific Acupuncture

Non-specific acupuncture has been used in the literature as a method of control in acupuncture RCTs which treat depression. Its alternate name is sham acupuncture. This process involves the administration of acupuncture needles outside of the CM model in a manner to ensure there is no therapeutic benefit. It is assumed that since the application of the needling does not maintain the integrity of the CM model, any benefit of the needling cannot be attributed to the CM model and/or acupoint specificity. Since it closely replicates a legitimate acupuncture treatment, subjects are fooled into believing they have received a valid acupuncture intervention when in fact they have not. This method of control has been used in the literature to facilitate between group comparisons by comparing those who received an active acupuncture intervention to those who receive non-specific acupuncture. This comparison is believed to offer insight into any involvement of the placebo effect. If an active acupuncture treatment had greater rate of effectiveness than a non-specific treatment, then the added benefit of the active acupuncture can be attributed to acupoint specificity within the CM model.

Yeung (2006) highlighted the problematic nature of using non-specific acupuncture as a means of control. It was explained that differing schools of CM thought may have opposing views on the principle of treatment and indications associated with an acupoint. Further to this, it was suggested that under the current understanding of the CM model, the specificity of an acupuncture point cannot be fully accounted. In turn Yeung (2006) concluded that a true control using non-specific acupuncture is not possible.

Non-specific acupuncture was used as a method of control by five studies in the literature (Allen, 1998; Allen, 2000; Roschke, 2000; Manbar, 2004 and Allen, 2006). The most common form of non-specific acupuncture involves needling known acupoints which are unrelated to the targeted principle of treatment (or are not indicated for the condition in question). In the depression studies in the literature, Allen (1998) and Allen (2006) both used this method. In each case Allen used acupoints indicated to treat symptoms which were unrelated to depression. Between group analysis using the Ham-D indicated the active acupuncture used in Allen (1998) was significantly more effective at reducing depression scores than the non-specific acupuncture ($p < 0.05$). Extra analysis using the BDI confirmed this outcome ($p < 0.05$). In contrast Allen (2006) found no difference in their between group analysis. The active acupuncture was reported to be no more effective than the non-specific acupuncture according to the Ham-D ($p > 0.2$) and the BDI ($p > 0.17$). Though, it is important to note the within group analysis in Allen (2006) indicated the active acupuncture intervention was effective on the Ham-D ($p < 0.001$) and the BDI ($p < 0.001$). This would indicate that although the active acupuncture was effective, it was not able to be significantly more effective than the non-specific acupuncture.

An alternate method of non-specific acupuncture involves needling areas of the skin which are not known locations of acupoints. This technique is designed to prevent the stimulation of any acupoint or meridian. In the literature, Roschke (2000) used this method and needled subjects in areas of the skin which were proximal to the same acupoints which were used to treat the subjects who received the active treatment. To further reduce the effect of the stimulation, the needles were reported to merely pierce the skin. The results of Roschke (2000) yielded mostly insignificant results when the active acupuncture was compared to the non-specific acupuncture. Interaction analysis considered group and time and found no significant difference on all but one of the scales. The scales which indicated no significant difference on group time interaction were in terms of the functioning of subjects according to the GAS ($p = .0.052$), in terms of

melancholia according to the BRMS (p = 0.226), and in terms of the severity of the illness according to the CGI Item 1 (p = 0.092). The significant outcome (p = 0.048) was found on the CGI Item 2. This is a measurement of global improvement.

It can be argued that these results from Roschke (2000) may offer limited insight into the use of non-specific acupuncture. The analysis was not strictly between active and non-specific acupuncture. Instead, the results were compared between three groups at once (active acupuncture, non-specific acupuncture and a wait list control group). When data for the active and non-specific acupuncture was combined and compared to the waitlist group there was a greater statistical significance (p = 0.026). This suggests the significant difference when all three groups were considered is more attributable to differing outcomes for those in the wait list group and there may be little difference between the active acupuncture and non-specific acupuncture groups.

There were two studies in the literature which did not report their method of non-specific acupuncture. These studies were Manber (2004) and Allen (2000) as a consequence, the outcomes from these two studies are less informative. Allen (2000) used the Ham-D and found active acupuncture to be significantly more effective at treating depression than non-specific acupuncture (p <0.05). In contrast, Manber (2004) indicated no significant differences between Ham-D scores and the specific and non-specific acupuncture groups (*p* = 0.18). BDI scores in the Manber (2004) study supported the Ham-D with no significant difference in the first month of treatment (p = 0.092) and the second month of treatment (p = 0.115).

It is important to note that the difference in outcomes in the above studies may well be due to how the non-specific was administered. This may well be the case for Roschke (2000) since the needling used in that study was close to the acupoints used in the active treatment. This may account for the insignificant

differences in results between active and non-specific acupuncture. The non-specific acupuncture may have simulated an active acupuncture too closely. In turn, the non-specific acupuncture may have resulted in an unintended therapeutic benefit. Allen (1998) an Allen (2006) used the same method of non-specific acupuncture and yet still these studies produced contradicting results. It is unclear if the conflicting results are due to the difficulties surrounding non-specific acupuncture or if they are due to other variables. Nevertheless, it is evident from the literature the comparison between active acupuncture and non-specific acupuncture is a source of ambiguity and does not provide an adequate control.

The researchers involved in this study consider the points made by Yeung (2006) are valid. This study will not use non-specific acupuncture as a control. CM theory is complex and it still is yet to be determined if non-specific acupuncture functions as a true placebo. There are two reasons for this. Firstly, criteria to determine a valid acupoint is at best vague since there is little understanding on the mechanism of acupuncture. Hence there is no criterion to facilitate the identification of an area of the body which is devoid of acupoint specificity. Secondly, *Five phase theory* or even the acupuncture channel system *(jing luo)* are theoretically recognised as closed loop systems. Treatment at any one point within the system will ultimately affect the balance of the entire system. This *balancing/regulating* mechanism is one of the principle mechanisms behind acupuncture. Consequently, the use of so-called unrelated acupuncture points as controls does not account for the non-specific effects of acupuncture.

2.13.2 Antidepressant Medication

The use of drug therapy as a means of control, was employed in three studies conducted in China. In each instance the effectiveness of an active acupuncture intervention was compared to a control group whose participants were

administered an antidepressant. Luo (1985) and Han (2004) used electro-acupuncture whilst Yang (1994) used an acupuncture prescription based on the extra channels. Luo (1985) and Yang (1994) are older studies and hence used amitriptyline. Amitripyline is a generic name and belongs to the tricyclic family of antidepressants. Han (2004) on the other hand administered maprotiline, which is also a generic name, and belongs to the tetracyclic family of antidepressants.

All three studies, which compared an active acupuncture treatment to an antidepressant drug, drew the same conclusions. Luo (1985), Yang (1994) and Han (2004) used the Ham-D and found both acupuncture and the respective antidepressant they used as a means of control significantly reduced depression scores ($p < 0.01$). There were however no significant difference between the effectiveness of the acupuncture and the antidepressant ($p > 0.05$). Luo (1985) also used the CGI and found the same outcomes when considering the severity of the disease and the general progress of the treatment. Both groups showed significant improvement in scores ($p < 0.01$), but there was no difference between the groups ($p > 0.05$).

This approach produces meaningful outcomes as it allows for direct comparison to be made between acupuncture and conventional depression treatments. For this reason, the researchers in this study appreciate the inclusion of medication within the design of depression studies. It additionally improves the relevance to the clinical setting and facilitates the satisfaction of ethical considerations. Consequently, the researchers have decided to account for the use of medication in the design of this study. The role of medication in this study will however need to take a different role from the studies in the literature. It is highly unlikely the researchers in this study will be able to replicate the design used by Luo (1985), Yang (1994) and Han (2004). The methodological designs used in these studies carry inherent problems associated with recruitment and ethical considerations. The antidepressant will also need to be changed for another which is more reflective of current treatment of depression.

The main issue of concern when incorporating the use of medication with the design of a study is around the recruitment of subjects. A method of recruitment to duplicate the design used within the studies in the literature is problematic. It is not possible to facilitate the situation in which to form a group of subjects who are taking the one type of medication and another group who are taking no medication. To independently recruit separate subjects for each group would contravene randomisation practices. It would implicitly require those taking the medication to be automatically assigned to the control group and those taking no medication to be automatically assigned to the treatment group.

Another issue is finding enough subjects who are taking the one generic medication. Ethics approval would not allow the researchers to require subjects to change or cease taking the medication which they have been prescribed by their treating physician. This is of additional concern as it would jeopardise the simulation of a clinical setting as subjects would not be taking the antidepressant they would ordinarily be prescribed. Hence the method of control used by Luo (1985), Yang (1994) and Han (2004) not only contradicts the practice of sound scientific method but it raises ethical concerns.

Both Tricyclics and Tetracyclics used in the studies within the literature are no longer frequently prescribed. These are older types of antidepressant medications and there has since been further refinement in the medicating of depression sufferers. Consequently, comparatively few sufferers of depression continue to take these drugs to treat their depression. Additionally, these studies also used a generic name of an antidepressant as a parameter within the inclusion criteria. If this were to be used in this study it may create difficulty in satisfying inclusion criteria. The tighter parameters of inclusion will increase the difficulty of recruitment. It may prevent obtaining suitable subject numbers. The researchers in this study have decided to extend the inclusion criteria for medication from the generic name to the broader parameter of the family name

and then monitor the use of the generic named drugs. This will assist in recruitment whilst maintaining meaningful outcomes and clinical relevance.

This study will limit the inclusion of subjects to those taking Selective Serotonin Re-Uptake Inhibitors (SSRI's). This is currently the most commonly prescribed antidepressant medication (Parker, 2006).

In response to the problems associated with including the use of antidepressant medication into the design of a study, the researchers have decided not to design this study so as to compare an active acupuncture treatment to an antidepressant medication. An alternate method needs to be considered. Other studies in the literature have used acupuncture as an adjunct treatment to medication. Designs of this nature are less fraught with difficulties.

2.14 Acupuncture as an Adjunct to Medication

In the literature, the use of acupuncture as an adjunct to antidepressant medication as an element of design is not common. The design of such studies requires all subjects who are participating in the study to be taking a certain medication throughout their involvement in the study. An acupuncture intervention is therefore administered to those in the treatment group in addition to the medication. Valuable inferences can however be made from studies of this nature. Since they closely simulate a clinical setting, the outcomes provide more clinically relevant and valuable implications. This type of trial design also satisfies ethical considerations when the target population is likely to be medicated. Additionally, studies designed in this manner help circumvent other issues such as randomisation problems and recruitment challenges which are present when medicated subjects are involved. Examples of studies in the literature which use acupuncture as an adjunct treatment within a trial design are Luo (1990), Luo (1998) and Roschke (2000).

Roschke (2000) medicated all subjects in his study with an antidepressant called mianserin. Comparisons were made between a treatment group who received active acupuncture, a non-specific acupuncture group and a control group (who only received the mianserin). This study yielded mostly insignificant results on interaction analysis on group and time. The only significant ($p = 0.048$) factor was global improvement as measured on the CGI Item 2. Both the active and non-specific acupuncture group showed greater global improvement compared to those in the wait group. The other measures indicated no significant difference on group time interaction in terms of the functioning of subjects according to the GAS ($p = .0.052$), in terms of melancholia according to the BRMS ($p = 0.226$), and in terms of the severity of the illness according to the CGI Item 1 ($p = 0.092$). The outcomes in Roschke (2000) provided little evidence to suggest acupuncture increased the effectiveness of a treatment when combined with mianserin, compared to when the mianserin was administered alone.

Phase 1 of Luo (1998) partially employed a design using acupuncture as an adjunct to antidepressant medication. The medication used in this study was amitriptyline. An active electro-acupuncture treatment was administered in conjunction with medication and was compared to those who only received the medication. According to the Ham-D, both groups experienced significant improvement in depression scores ($p < 0.001$) but there was no significant difference between the groups ($p > 0.05$). Like Roschke (2000) this outcome suggests the addition of acupuncture did not improve the effectiveness of treating depression with an antidepressant.

In Luo (1990) and both phases of the Luo (1998) study, acupuncture was administered as an adjunct treatment to a placebo drug. In all three instances significant reductions in Ham-D scores were found ($p < 0.01$). The acupuncture group was then compared to a control group which received the antidepressant medication of amitriptyline alone. Those who received the medication alone

experienced significant improvements in Ham-D scores across all three studies (p < 0.01). Between group analysis compared those who received the acupuncture and the placebo drug to those who only received the amitriptyline. No statistical significance was found (p > 0.05). These results indicate that acupuncture may add to the effectiveness of a placebo and together this treatment is equal to that of an active antidepressant.

There is an alternate way to consider the design used in phase 1 of Luo (1998). This would involve the comparison of those who received the acupuncture and the medication to those who received the acupuncture along with the placebo medication. This in effect would mean all subjects involved in the comparison received an electro-acupuncture treatment. The point of difference between the two groups is that one received an active antidepressant medication (amitriptyline), and the other received a placebo (non active drug). This is in effect testing if the drug is more effective when it is administered with acupuncture compared to when it is not. Analysis using the Ham-D found significant reductions in the depression score for both groups (p <0.001), however there was no difference between the groups (p <0.05). This means the acupuncture was equally effective when it was administered with a placebo medication than when it was administered with an active drug. This raises an interesting point of consideration. Conventional medical opinion would more likely attribute a placebo effect to an acupuncture treatment over antidepressant medication. This study however suggested there is also a placebo effect associated with taking amitriptyline since the acupuncture intervention is constant across both groups.

Interestingly, the results from both of the phase 1 of Luo (1998) and Roschke (2000) studies were mostly insignificant. These outcomes appear to be at odds with conventional clinical opinion and practice which would assert patients would be less depressed, if they received acupuncture in addition to their medication.

Further research is required to provide an understanding of why the outcomes in the literature contradict accepted clinical understandings.

This study will implement the approach of using acupuncture as an adjunct treatment to antidepressant medication as done by Roschke (2000) and in phase 1 of Luo (1998). That is, all subjects will be prescribed and taking antidepressant medication and those who receive acupuncture will be compared to those who do not. This is the most informative design and is most suitable to test the effectiveness of using acupuncture as an adjunct treatment to the antidepressant medication. This then introduces a wait list control group as a component of trial design.

2.15 Wait List Control Groups

The use of a wait list is an interesting method of control. This essentially entails no treatment is administered to those who are assigned to a wait group. In effect a wait list group facilitates the proposition of the question as to if it is better to receive treatment or none at all. This method of control was employed by Allen (1998), Allen (2000) and Allen (2006). They used a simple comparison between an active acupuncture intervention and a wait list group. Allen (2006) made the same comparison but also included another by comparing between the wait list group to those who received non-specific acupuncture.

According to both Allen (1998) and Allen (2000), the Ham-D analysis indicated no significant differences when those who received the active acupuncture were compared to those in the wait group [($p < 0.012$) and ($p < 0.12$) respectively. These outcomes are different to the findings of Allen (2006). Ham-D analysis in this study found subjects in both the active and the non-specific acupuncture significantly out-performed those in the wait group ($p < 0.001$). This outcome was confirmed by the analysis using the BDI ($p < 0.001$).

The inconsistent findings between the Allen (1998), Allen (2000) and the Allen (2006) studies indicates further research is required to add clarification. It is yet to be conclusively determined if those who receive an acupuncture intervention are less depressed compared to those who take part in a wait group. This is a very valuable point of differentiation, which needs to be answered. It carries important clinical ramification since it determines if it is better to treat or to not treat those who suffer from depression with acupuncture. There will be no value in administering an acupuncture treatment if it is not able to facilitate greater improvement in depression than when no treatment is administered.

2.16 The Combined Use of a Wait List with Acupuncture as an Adjunct to Medication

Compared to the Allen (1998, 2000 and 2006) studies, Roschke's (2000) employed a method of design which is clinically more interesting and informative. This study compared an active acupuncture treatment to a wait list control group whilst all subjects in both groups were taking an antidepressant. In effect this meant those in the treatment group took medication and received acupuncture as an adjunct therapy. In contrast, those in the wait list group just continued to take the medication. This design combines the use of an adjunct therapy and a wait list control group. The results from this study indicated those who received the acupuncture interventions showed significantly more improvement than those who received medication alone. This was the case in terms of global improvement of symptoms according to the CGI Item 2 ($p = 0.026$) and in the functioning of subjects according to the GAS ($p = .0.049$). No difference was however found between the two groups in terms of melancholia according to the BRMS ($p = 0.384$) and in terms of the severity of the illness according to the CGI Item 1 ($p = 0.065$).

It is difficult to ascertain exactly why there are inconsistencies in the Roschke (2000) findings. Though it does need to be acknowledged that each of the

scales in Roschke (2000) were measuring different dimensions associated with depression. For example the GAS measured the functioning of subjects, which is different to melancholia as measured by the BRMS. This may account for the conflicting outcomes. The same can also be said between all four of the measures used in the study.

Luo (1998) partially used the same design as Roschke (2000) in phase 1 of his study. An adjunct therapy was used in conjunction with a wait list group. Subjects who were taking antidepressant medication (amitriptyline) were also administered electro-acupuncture. Comparisons were then made to another group of subjects who just received the amitriptyline. According to the Ham-D, both groups experienced significant improvement in depression scores ($p < 0.001$) but there was however no significant difference between the groups ($p > 0.05$).

Despite the ambiguity surrounding the results of Roschke (2000) and Luo (1998), the design of these two studies is of interest. The combination of a wait list as a method of control, in conjunction with using acupuncture as an adjunct treatment to medication, is valuable methodological design. It poses the question as to whether the depression sufferer who is taking medication is better off if they also receive acupuncture treatment, as opposed to taking the medication alone. Based on their clinical experience, the researchers in this study are of the opinion this combination of these design elements has the greatest relevance to the clinical setting. In response, this study will incorporate a wait list group and combine it with using acupuncture as an adjunct treatment to antidepressant medication.

2.17 Other Control Methods

Two alternate methods of control were noted in the literature. The Quah-Smith's (2005) study focused on the use of laser acupuncture. This involved the

construction of a laser with two settings, which enabled the administration of an active laser to the treatment group and an inactive laser to the control group. This left the treating acupuncturist unaware as to which group was receiving the active treatment. BDI analysis indicated the active laser was significantly more effective at treating depression ($p > 0.001$).

Manber (2004) focused her study on pregnant women and made comparisons between those who received an acupuncture intervention to those who received massage therapy. Analysis using the Ham-D indicated those in the acupuncture group experienced significantly greater improvement in their depression scores than those who received the massage ($p = 0.031$). Subsequent analysis using the DBI supported this finding after the first month of treatment ($p = 0.047$) but found no difference between the groups after the second month of treatment ($p = 0.083$).

Each of these methods of control has their respective shortcomings. The use of laser acupuncture is an effective method of control as it facilitates a double blind study design. Both the treating acupuncturist and the subject are unaware as to who is receiving the active treatment. Despite this, there is conjecture surrounding the effectiveness of laser acupuncture and whether "laser acupuncture" constitutes acupuncture at all. It is yet to be determined if laser acupuncture has the capacity to effectively stimulate an acupoint. Given this study will be focusing on the specific principle of dispersing liver qi stagnation, it will be more suitable to use a method of acupuncture which ensures the necessary acupoint stimulation. This would provide a more suitable test of the CM model. The researchers in this study consider the use of massage as a method of control would be outside of the scope of this study. This study is designed to determine if acupuncture is an effective adjunct to antidepressants. A comparison to massage therapy would not offer further the understanding on this line of enquiry.

2.18 Follow Up Analysis

Follow up analysis was included in five of the studies in the literature for using acupuncture to treat depression. Each of these studies approached their respective follow up assessment in their own way.

2.18.1 Within Group Comparisons

Gallagher (2001) and Manber (2004) were the only studies which made within group comparisons to test if the benefits of an acupuncture intervention were able to be maintained at the conclusion of a wait period. The focus of the Gallagher (2001) study was specifically on results from a follow up assessment. It considered how well females were able to maintain changes in their depression scores from the treatment they received whilst pregnant. The follow up assessment was carried out six months after the conclusion of the treatment. Statistically, remission status at the conclusion of the treatment predicted clinical status at the end of the six months ($p = 0.04$). The results on follow up indicated a relapse rate of 24% according to DSM IV. Using the Ham-D, symptom severity at the end of treatment also remained unchanged ($p = 0.033$ – one-tailed) over the wait period. This means female subjects were able to maintain their improved depression scores for a period of up to six months after the acupuncture intervention.

Manber (2004) also focused on follow up assessments on women who received acupuncture whilst pregnant. In contrast to Gallagher (2001), the follow up in the study was done 10 weeks from the conclusion of the intervention. Ham-D analysis indicated significant reductions in symptoms after delivery of the baby ($p < 0.0001$) compared to the scores before the acupuncture intervention was administered. This implies the subjects were able to improve their depression score as a result of the intervention and this improvement was maintained over a wait period. Despite both of these studies being focused specifically on

pregnant females, both studies offered encouraging evidence on the longevity of improved depression scores from an acupuncture intervention.

2.18.2 Between Group Comparisons

Between group comparisons were made by Luo (1990), Gallagher (2001), Manber (2004) and Quah-Smith (2005). Manbar (2004) showed there were no significant differences in between group interactions when the pregnant women received acupuncture compared to either non-specific acupuncture or massage according to the Ham-D (p = 0.12) and the BDI (p = 0.21). Luo (1990) conducted a follow up assessment between two to four years after the subjects received either acupuncture as an adjunct to a placebo medication or an antidepressant drug alone. No significant difference was found between the two groups (p > 0.05). It must be noted that the length of this follow up period is not suitable for a study of this nature. A four year follow up period for a condition such as MDD may not be clinically relevant.

The between group follow up analysis of Quah-Smith (2005) however showed an interesting pattern of results. This study compared active laser acupuncture treatment to a sham laser. There was no significant difference in BDI scores at 4 weeks post treatment (p = 0.088), but at 12 weeks post treatment, those who received the active laser were significantly less depressed than those who received the sham laser (p = 0.007).

The encouraging follow up results from Gallagher (2001) and Manber (2004) were not evident when between group comparisons were considered. When the follow up results for those who received an active acupuncture intervention were compared to those from a control group, the between group analysis suggested the acupuncture was no more beneficial. These outcomes indicate there is a need for further research to better understand the long-term benefits of using acupuncture to treat depression.

To date, the majority of studies in the literature focused on follow up analysis on women, during and after pregnancy. With the exclusion of Luo (1990) and Quah-Smith (2005), it remains to be adequately tested if men or non-pregnant women will maintain improved depression scores as well in these instances. In addition, all the follow up analysis in the literature was mostly confined to comparisons to post treatment scores. One study in the literature, Manber (1994), compared follow scores to base scores before the acupuncture intervention. This is a more meaningful comparison as it offers a broader scope on changes in depression. Additionally, further research is required which will incorporate a design which is more suitable for an acupuncture clinic trial. As discussed earlier, the studies in the literature which consider between group comparisons have inherent shortcomings since consensus on an adequate method of control is yet to be determined. This in turn has an implication when considering follow up scores for the studies concerned.

The researchers in this study will include a follow up analysis. The inclusion of both men and women will allow a homogeneous assessment to be made amongst depression sufferers regardless of sex. In addition results with be stratified to enable the comparison of follow up scores to be made between the sexes. As done in Manber (2004), two follow up score comparisons will be also carried out. The follow up scores will be compared to both the base scores before the intervention, as well as post intervention scores. The researchers can then see if subjects can maintain their depression scores over the wait period, and test the subjects' scores relative to their pre-treatment scores. Analysis of this nature is of greater clinical significance. Furthermore, this study will also serve to offer new knowledge. It will be the first study to report a follow up analysis when acupuncture is used as an adjunct treatment to antidepressants.

2.19 Conclusion

This review of the literature on studies which have used acupuncture to treat depression has revealed numerous shortcomings. Firstly, the studies in the existing literature are characterised by poor reporting standards. This has engendered ambiguity and it has jeopardised the clinical implications which could have been drawn from the outcomes. Secondly, the design of many of these studies needs to be reassessed. There is an existing tendency to utilise methods of control which are not compatible with the CM model, such as the case for non-specific acupuncture. Thirdly, the use of antidepressant medication which was not only irrelevant to current developments but was incorporated in a manner which raised concerns over randomisation.

The researchers involved in this study will endeavour to implement improvements to rectify these shortcomings. Greater efforts will be made to adhere to improved standards of reporting. A method of control will be employed which best maintains the integrity of the CM model. The use of antidepressant medication will be updated and a suitable design will be used to ensure randomisation requirements are maintained. The follow up analysis in this study will also incorporate greater detail. Follow up scores will be compared to those before and after the acupuncture intervention. Finally this study will be the first study to make provisions for stratification based on sex. This will allow for comparisons to be made between females and males across all components of the study.

Chapter 3: Method

3.1   Elements of Good Design

There are known difficulties in designing randomised clinical trials (RCT) for acupuncture. The scientific method demands standardisation of experimental variables, yet acupuncture treatments require flexibility, often based on idiosyncratic prescriptions and individualised CM diagnostic patterns. This causes randomised clinical trials (RCT) for acupuncture to have a number of problems including difficulty with standardisation of diagnostic criteria, treatment intervention/programming, determining suitable controls and the implementation of appropriate measures.

In designing an acupuncture RCT, it is important the method be scientifically rigorous. The implementation of good design combines the best practices of the scientific method whilst maintaining the integrity of the CM model. Compromise is needed from both paradigms in order to produce a meaningful clinical outcome. If a design does not maintain the integral components from both paradigms, then the outcomes from the study will be problematic and limited. In turn the validity of the study will be questionable and the clinical implications of the outcomes will be uninformative.

In the context of this study, best practice also required clear definitions of depression. Subject inclusion/exclusion criteria were implemented from both CM and psychiatric/psychological perspectives. The CM perspective required the definition of the relevant diagnostic patterns of disharmony and appropriate inclusion criterion, as well as the justification of relevant acupoint prescription. This was done via a manualisation process of the relevant diagnostic patterns and ensuing acupoint prescriptions. The psychological perspective required the utilisation of relevant diagnostic criteria along with the selection of relevant and

reliable psychometric measures. This was done to achieve a practical synthesis of both models whilst maximising clinical relevance.

The intention of this study was to design a RCT with outcomes which had the greatest clinical impact in the treatment of depression. This study therefore defined the most common representation of depression sufferers according to each paradigm. Since this study best reflected clinical reality it maximised the relevance of the results for acupuncture practitioners, mental health professionals and patients.

3.2 Design

This study tested the effectiveness of using acupuncture as an adjunct treatment to the SSRI class of antidepressant medication. This was done by using a repeated measures design which included a follow up analysis.

Comparisons were made between subjects in a treatment group and a control group. These comparisons involved both sexes as a homogenous group. Further provisions were made to allow for stratification of subjects based on sex. This facilitated comparisons to be made between the sexes within the groups as well as comparison for each of the respective sexes across the two groups.

The researchers in this study proposed a synthesis model which was designed to best test the effectiveness of using acupuncture to treat depression. This is a "best fit" model which will appropriately combine the requirements of scientific method with the idiosyncratic nature of the CM model. This chapter will define and outline each of the components of this proposed model. The first component of the model involved the inclusion of suitable methods of control. This involved a wait list control group coupled and the use of acupuncture as an adjunct treatment to antidepressant medication. As previously highlighted, there is a requirement for depression to be defined from both psychological and CM

perspectives. This in turn served to assist the formulation of suitable inclusion/exclusion criteria. Given this, valid and reliable psychometric measures were determined. Moreover, the clear definition of subjects assisted in the construction of a suitable acupuncture intervention. Consistent with the CM model, this was based on a principle of treatment in accordance with a pattern of disharmony. Lastly, the details of the procedure will be reported to provide clarity and to assist future replication of the design.

3.3   Method of Control

As noted in the literature review, designing adequate controls for acupuncture RCTs is fraught with difficulties and controversy. Numerous trials in the literature have used unrelated acupuncture points as controls. Neither enterprising patients nor experienced acupuncturists can be effectively blinded to the acupoint functions and locations. Additionally, the CM model is complex and the use of so-called unrelated acupuncture points as controls does not account for the non-specific effects of acupuncture. Equally problematic is the use of "sham" points. Needling in the area of an acupuncture point, anywhere along the same dermatome as an "active" point, may stimulate the same non-specific acupuncture effect. "Sham" needles are also a problematic alternative. It is questionable if the necessary deception can be perpetrated over a series of 12 treatments. In addition, "sham" is characterized by recruitment and randomisation difficulties since it requires acupuncture naive individuals be recruited for all groups.

This clinical trial employed a wait list control. This method has shortcomings, but on consideration, it is the best alternative. Despite not fully accounting for the placebo effect, wait list controls are suitable for pragmatic acupuncture RCTs where the question being asked is if the acupuncture intervention is better than standard care or no intervention at all. Wait lists provide adequate control for

disease remission, and from an ethical perspective, all subjects have the opportunity to receive the active treatment.

3.4 Defining Depression: DSM-IV-TR

The Diagnostic and Statistical Manual for Mental Disorders Fourth Edition Text Revision was used to define subjects from a psychological/psychiatric perspective. The DSM-IV TR is published by the American Psychiatric Association and was designed to be used by a diverse array of mental health professionals in a variety of settings. The diagnostic criterion for a Major Depressive Disorder (MDD) was formulated using a task force comprised of a wide range of health experts. This panel of experts considered published literature, pertinent depression related issues, and participated in field-trials and data-reanalysis projects to construct the diagnostic criteria for a MDD. The diagnosis of a MDD is determined via a recommended Structured Clinical Interview (SCID). Within the psychiatric community DSM-IV TR is considered to be the gold standard and most up to date in the diagnosis of a major depressive episode. (APA, 2000)

The Structured Clinical Interview for MDD was formatted by the researchers in this study into a tick box arrangement (Appendix A). This led to the creation of a form which included both the DSM-IV TR inclusion and exclusion criteria for MDD. This form was administered during each of the assessments throughout the study to ensure standardisation. In addition, training was administered by a psychologist who was expert in the diagnosis of depression to ensure correct application of the structured interview.

3.5 Serotonin Selective Re-Uptake Inhibitors

To more accurately define the sample population from a psychological/psychiatric perspective, this study classified subjects according to

their prescribed medication. At the time when the researchers designed this study, Parker (2006) recommended Serotonin Re-Uptake Inhibitors (SSRI's) as the "first port of call" medications in the treatment of depression. This recommendation in turn has led SSRI's to be the most frequently and most commonly prescribed medication to treat major depressive disorder in Australia at that time. In order to allow for the greatest clinical implication SSRI's were chosen as inclusion criteria and subjects taking other classes of antidepressant medication were excluded.

Noradrenaline-serotonin reuptake inhibitors (NSRIs) are a similar antidepressant medication to SSRIs. These two medication can often be classified together. In order to more accurately define the subject population, the researchers in this study however decided not to include those taking noradrenaline-serotonin reuptake inhibitor (NSRIs).

3.6   The Manualisation Process

As discussed, it is an element of good design to implement clearly defined principles within a study. The manualisation process is a suitable method to fulfill this requirement. It is essentially a method of enquiry used to identify general consensus. It involves the surveying of the various relevant sources to find commonalities via the tallying up of the most frequently offered understandings and interpretations.

The CM model is open to interpretation. There are differing schools of thought and each has created a different approach in applying the CM model. Consequently, when acupuncture is used within both an RCT and in clinical practice there are a number of possible variables which need to be considered. The variations between the various schools of thought are mostly centered on diagnosis and treatment. In both theory and practice, the CM model is founded on identifying a diagnosis via the identification of a pattern of disharmony. Then

subsequent to the identification of a pattern of disharmony a matching and relevant acupoint prescription can be constructed. The identification of the signs and symptoms for a pattern of disharmony as well as its matching acupoint prescriptions are open for debate. In turn this study required two manualisation processes. The first process identified the pattern of disharmony which most commonly occurred in sufferers of depression. This involved the identification of the most common symptoms representing the pattern in question. The second process was concerned with ascertaining a consensus of an acupoint prescription which best applied the principle of treatment that matched the pattern of disharmony.

3.6.1 Liver Qi Manualisation

According to Schnyer (2000) there are twelve Chinese medicine patterns of disharmony which can account for depression. Four of these patterns either have as their root cause or are characterised by varying degrees of liver qi stagnation. In addition to this, epidemiological data from the UTS Acupuncture Clinic Database indicated liver qi stagnation as the second most frequently presenting pattern of disharmony in the UTS clinic. As noted in table 3 below, it accounts for 10.2% of all diagnoses. When these data are broken down into the symptom subcategory of "feeling depressed" 24.8% of all diagnostic categories associated with this symptom are identified as liver qi stagnation. If this is expanded to include all related liver patterns then "feeling depressed" is associated with 34.3% of the cases. Similarly for the symptom "depressive disorder", liver qi stagnation accounts for 22.6% of all diagnoses associated with this symptom and all liver related diagnoses account for 32% of cases. Since the UTS clinic and the location of this study were co-located and both drew on the same population it was reasonable to assume that liver qi stagnation would be the most frequently presenting pattern of disharmony in the prospective subjects. The diagnosis of liver qi stagnation was therefore the most suitable pattern of disharmony to be addressed in this study. To increase the clinical

relevance and sample size, the researchers included the diagnoses of *liver qi stagnation* combined with *wood attacking earth* and *liver qi stagnation* with *heat*. The manualisation process indicated these two extra patterns are closely connected to liver qi stagnation. This is consistent with the theoretical understanding of liver qi stagnation and the CM model. When a pattern of disharmony is experienced long term the said imbalance will become more systemic. *Liver qi stagnation* combined with *wood attacking earth* and *liver qi stagnation* with *heat* are the most common extensions of liver qi stagnation as it becomes more systemic.

Table 3: UTS clinic patient database

UTS Patient Database		
Total diagnoses (all categories)	127903	
Diagnosis	**Count of Diagnosis**	**Percentage of total**
Spleen Qi Deficiency	13889	10.9%
Liver Qi Stagnation	13038	10.2%
Total Number of Diagnoses per ICPC* category P03 "feeling depressed"		
Total Diagnoses (all categories)	2323	
Diagnosis: Liver Qi Stagnation	575	24.8%
Diagnosis: Wood (liver) invades Earth (spleen)	61	2.6%
Diagnosis: other liver related e.g. liver wind, liver heat etc.	160	6.9%
Total Number of Diagnoses per ICPC* category P76 "depressive disorder"		
Total Diagnoses (all categories)	1386	
Diagnosis: Liver Qi Stagnation	313	22.6%
Diagnosis: Wood (liver) invades Earth (spleen)	31	2.2%
Diagnosis: other liver related e.g. liver wind, liver heat etc.	100	7.2%
* ICPC: International Classification of Primary Care		

A manualisation process was carried out to determine exactly what signs and symptoms constitute each of the variations of liver qi stagnation. The sources included in the process were relevant literature, consultation with an expert panel of experienced acupuncture researchers, and the epidemiological data for the UTS clinic. Refer to Appendix B for the results of the manualisation process.

The findings from the manualisation process were used to create a diagnostic template to assist with subject selection. This was done by employing a tick box format to create a form which was administered during the initial interview (refer to Appendix C). The form first assessed the presence of liver qi stagnation. This was done by dividing the symptoms of liver qi stagnation into two groups, the key symptoms and the more secondary symptoms. The manualisation process identified 12 key symptoms of liver qi stagnation. The corresponding tick boxes for each of these key symptoms were placed on the form above a line. The secondary symptoms were placed below the line. Based on the findings from the manualisation process and the experience of the researchers, it was determined that a positive diagnosis of liver qi stagnation and inclusion into the study would require the presence of at least four of the key symptoms.

The same process was undertaken to gauge if the liver qi stagnation had become more systemic and satisfied either the diagnosis of *liver qi stagnation* combined with *wood attacking earth* or *liver qi stagnation* with *heat*. This was used to determine which principle of treatment and acupoint prescription would best treat each subject. The manualisation process identified an additional three key symptoms of *wood attacking earth* and an additional six key symptoms of *heat.* Again each of these key symptoms was placed above a line on the form under their respective pattern of disharmony. To satisfy the diagnosis of *wood attacking earth,* one of the key symptoms were required. The diagnosis of *heat* required two symptoms to be present. Refer to Appendix C to view the form as used by the researchers.

3.7 Inclusion/Exclusion Criteria

In accordance with the line of investigation and in the context of the literature review, the researchers in this study decided on the best parameters for inclusion/exclusion criteria.

Inclusion Criteria required subjects to be:

1. prescribed and currently taking selective serotonin re-uptake inhibitors;

2. experiencing Major Depressive Disorder as outlined in the Diagnostic and Statistical Manual of Mental Disorders Fourth Edition Text Revision (DSM-IV-TR);

3. diagnosed as experiencing *liver qi stagnation, liver qi stagnation* with *heat* or *liver qi stagnation* with *wood attacking earth* according to Chinese medical patterns of disharmony.

Subjects were excluded if they:

1. were prescribed and taking an antidepressant other than an SSRI. Subjects who were taking a noradrenaline-serotonin reuptake inhibitor were excluded form the study. Those who were also taking any form of medication in addition to their SSRI were excluded from the study.

2. satisfied the exclusion criteria for Major Depressive Disorder according to DSM IV-TR. These criteria include depression due to a side effect of a drug, a general medical condition or bereavement. The DSM-IV also excludes the alternate diagnoses of manic episodes, hypomanic episodes, schizophrenia, schizoaffective disorder, schizophreniform disorder, delusional disorder and psychotic disorders not otherwise specified.

3. did not satisfy the diagnosis of *liver qi stagnation, liver qi stagnation* with *heat* or *liver qi stagnation* with *wood attacking earth.*

4. under the age of 18.

5. presented with a level of suicidal ideation indicating they were at risk of causing self harm.

3.8 Ethics Approval

This study was granted approval by the Human Ethics Research Committee at the University of Technology, Sydney (approval number 2007-28A). The approval included guidelines for advertising, procedures and duty of care to subjects. The approval was also granted for the content of information sheets, consent forms and questionnaires used in the study.

3.9 Outcome Measures

As noted in the literature review, the most common measure of depression identified from the relevant studies was the Ham-D. The BDI however, was beginning to be used more frequently amongst the more recent studies and was judged as the more contemporaneous scale better reflecting current views on depression. The researchers in this study decided to administer the BDI as the primary measure of depression. The Ham-D was administered as a secondary measure. The combination of depression measures allowed for valid and reliable assessments whilst facilitating comparison with the majority of the existing literature. The comparison of multiple measures also allowed for potential identification of a measure that was the most sensitive for determining change due to acupuncture. Given that no outcome scale will cover all measurable outcomes, the use of multiple scales would reveal additional or otherwise missed results if only a single scale were used. In addition to the depression measures, the Systematic Check-List 90 and the State-Trait Anxiety Inventory for Adults was administered. Both scales were included to give further insight and understanding into conditions of co-morbidity. In turn they were also expected to offer insight as to how the dispersal of liver qi may assist in the reduction of a broader spectrum of reported psychological and somatic symptoms in sufferers of depression.

3.9.1 Beck Depression Inventory – Second Edition

The primary measure and focus in the study was the Beck Depression Inventory (BDI-II). This measures severity of depression in those who are over the age of 13 years. It consists of a 21-item self-report instrument developed in accordance with most of the diagnostic criteria for major depressive disorder as set out in DSM-IV. The BDI-II is one of the most accepted measures of depression over the last 35 years. Reviews of the DBI-II have reported high reliability in psychometric properties and it's applications for clinical and non-clinical populations.

This measure was chosen since it best matched the requirements of the study. Its implicit method of self-reporting allowed for the questionnaire to be administered in the absence of either an expert in psychology or psychiatry. The time parameters of the BDI also suited the design of this study since it measured the severity of depression symptoms over the preceding two weeks. (Beck, 1987)

3.9.2 Hamilton Rating Scale for Depressive Illness

The Hamilton Rating Scale for Depressive Illness (Ham-D) was chosen as a secondary measure of depression. It provided a simplistic assessment of the severity of depression. It can be used to assess longitudinal changes in the severity of depression within patients. The scale is completed using practitioner ratings based on a structured interview using a varying number of items (up to 21) according to updated versions. Only 17 of these items are used to calculate the scores, which are reported to have high internal reliability. Questioning is directed towards how the patient has been feeling the last few days preceding the assessment.

The shortcoming of this measure rests in the requirement of an expert assessor. The value of each rating is dependent on the skill and experience of the

assessor and the information made available. The head researcher in this study received training from a psychologist who is a specialist in depression and is also an expert in using the Ham-D. This was done to ensure correct application of the structured interview. There still remained, however the risk of problematic outcomes stemming from an incorrect application by a non-expert assessor. Hence, the purpose of utilising the Ham-D in this study was to serve as a secondary measure and to provide a direct point of comparison to studies in the literature. (Hamilton, 1967)

3.9.3 State-Trait Anxiety Inventory for Adults

Depression and anxiety have a high incidence of co-morbidity and SSRI's are commonly prescribed to treat both depression and anxiety (Dunlop, 2008). It was deemed appropriate to include a measure to provide a better understanding of how depression and anxiety interacted within the study sample group. It would also serve to provide insight as to how the dispersal of liver qi stagnation could treat a broader spectrum of mental health symptoms.

The STAI was used to assess changes in anxiety scores. This scale consists of two subscales those being for trait based anxiety and state based anxiety. These scales are otherwise referred to as Y1 (state) and Y2 (trait). The state anxiety scale best describes intensity of anxiety related feelings, while the trait based scale indicates the frequency of anxiety related feelings. Each scale consists of twenty self-reported statements based on items which are measured on a four point likert scale. The STAI has been used extensively in both research and clinical practice and is considered a reliable and valid measurement of anxiety. (Speilberger, 1983)

3.9.4 Systematic Checklist 90

The Systematic Checklist 90 (SCL 90) measures current point in time symptoms. It indicates psychological status of mental health patients, medical patients and individuals in the community who are not clinical patients. It consists of 90 self-reported items which are scored using a five point likert scale from zero to four. Scores are interpreted in terms of nine primary dimensions. These dimensions are: somatisation, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation and psychoticism. In addition there are three further global indices which provide further information on symptomatology and psychological distress. The SCL 90 was deemed relevant to this study as it provided an alternate depression measurement. In addition it was included to facilitate investigation into liver qi stagnation and it's co-morbidity with other psychological primary dimensions. (Derogatis, 1977)

3.10 Acupoint Selection – Manualisation

Contrary to the flexible, personally tailored acupuncture treatments seen in clinical practise, the scientific method requires standardisation of acupoint prescriptions to control variables. Hence to achieve a consensus, a manualisation process for acupoint selection was implemented. According to the CM model, the dispersal of liver qi is the principle of treatment which is most effectively used to treat liver qi stagnation. The manualisation process focused on finding the most suitable acupoints which disperse liver qi. This process considered information from the relevant literature and consultation with an expert panel of acupuncture researchers. The raw data from this manualisation process is reported in
Appendix D.

The results from the manualisation process indicated the acupoints most frequently prescribed to disperse *liver qi* were Liver 3 (Taichong) and Liver 14 (Qimen). Liver 3 (Taichong) is said to act as the great passage way for the flow of qi in the liver channel (Deadman, 1998). It is located on the dorsum of the foot, in the hollow distal to the junction of the first and second metatarsal bone (Rogers, 2006). Liver 14 (Qimen) is effective in dispersing liver qi as it is the "front mu" point of the liver (Rogers, 2006). "Front mu" points are where the qi of the respective organs gather on the surface of the body (Deadman, 1998). It is located on the mammillary line, in the sixth intercostal space 4 cun lateral from the midline (Deadman, 1998).

Based on advice of the panel of experienced practitioners, the researchers included Pericardium 6 (Neiguan) as a support point. Pericardium 6 (Neiguan) is an acupoint which is indicated for and commonly administered to *calm shen* (Rogers, 2006). The researchers considered Pericardium 6 (Neiguan) to be a suitable support point since the pericardium and the liver channels have a Jue Yin connection. Pericardium 6 (Neiguan) is located on the flexor aspect of the arm, 2 cun proximal from the wrist crease, between the tendon of palmaris longis and flexor carpi radialis (Deadman, 1998).

The three acupoints of Liver 3 (Taichong), Liver 14 (Qimen) and Pericardium 6 (Neiguan) were administered to each subject as a core prescription. The method of basing the treatments on a core set of acupoints was consistent with Han (2004) and Quah-Smith (2005). This better maintained the integrity of the CM model since their selection was based on the principle of treatment matching the diagnosis. This method also better simulates clinical practice.

Using the same method of manualisation, the acupoints which were administered in addition to the core prescription for each of the three variations of liver qi stagnation were identified (Refer to appendix D). The results of this process were, those with *liver qi* symptoms alone received Gall Bladder 34

(Yanglingquan) in addition to the core points, those with symptoms consistent with *wood attacking earth* received Spleen 6 (Sanyinjiao), and those presenting with *heat* symptoms received Liver 8 (Ququan).

As identified earlier, the core acupoints to disperse liver qi were Liver 3 (Taichong) and Liver 14 (Qimen). These are both located on the liver channel. The inclusion of a third liver qi dispersing point would be best located on a channel other than the liver channel. This will allow for consistency with both the CM model and clinical practice. Gall Bladder 34 (Yanglingquan) was the most frequently occurring acupoint in the manualisation process which was not located on the liver channel (refer to Appendix D). It is also suitable to accompany the core acupoints since it can be administered when the subject is positioned in the supine position. Gall Bladder 34 (Yanglingquan) soothes the qi of the liver and the gall bladder (Rogers, 2006). It is located on the lateral aspect of the knee approximately one cun inferior and anterior from the head of the fibula (Deadman, 1998).

The manualisation process identified Spleen 6 (Sanyinjiao) as being the most suitable to be administered as the additional acupoint for those who satisfy the pattern of wood attacking earth (refer to Appendix D). It is located on an intersection of the liver and spleen channels and it has the function of strengthening the spleen whilst harmonizing the liver (Rogers, 2006). It is located on the medial side of the lower leg, in the depression posterior to the tibia, 3 cun superior to the prominence of the medial malleolus (Deadman, 1998).

Maciocia (2004) was the only source in the relevant literature which listed acupoints for liver qi stagnation with heat. In the absence of a consensus in the relevant literature, a decision was based on the clinical expertise of the panel of practitioners. Amongst Maciocia's (2004) suggestions was Liver 8 (Ququan). The researchers decided and agreed to utilise it as it is the water point of the

liver channel and it is said to cool the liver. It is located when the knee is flexed and is at the medial end of the popliteal crease, in the depression anterior to the tendons semimembranosis and semitendonosis, posterior to the head of the medial condyle of the tibia (Deadman, 1998). Consequently, each patient received four acupuncture points for each treatment as noted in table 4.

Table 4 Acupoint prescriptions.

CORE ACUPOINTS	Liver 3 (Taichong)
	Liver 14 (Qimen)
	Pericardium 6 (Neiguan)
ADDITIONAL ACUPOINTS	
Liver qi stagnation	Gall Bladder 34 (Yanglingquan)
Wood attacking earth	Spleen 6 (Sanyinjiao)
Liver qi stagnation with heat	Liver 8 (Ququan)

3.11 Materials

Needles

The acupuncture needles used for the treatments were Viva 0.25mm X 30mm single use needles manufactured by Helio Medical Supplies Inc.

CM case history template

A template for a full CM case history was developed using experience from a panel of expert acupuncture researchers. It was based on the case history questionnaire from the UTS student clinic. The template was used to assist the assessing diagnostician in conducting a structured and standardised CM case history. Refer to Appendix E to view the template.

3.11.1 Recruitment

Subjects were recruited via a range of mediums including:

- newspaper and magazine articles;
- newspaper advertisements;
- radio interview
- distribution of flyers in pharmacies; and
- the UTS website.

A total of 81 subjects were accepted into the study. A total of 66 subjects completed the study. Of those who completed the study, females and males were each represented by 33 subjects.

3.12 Randomisation

The method of randomisation was chosen based upon the constraints of the study in terms of both design and implementation. Subjects for this study were self selecting from the wider population. Upon screening and acceptance to the trial they became members of the sample population. This study used a wait list control whereby all subjects would, at some point in time receive the active treatment. As such, it was determined that stratified simple random sampling with replacement was sufficient to account for bias, giving each subject an equal chance of being chosen for either the intervention or control group.

The random allocation of subjects into groups was done in advance of each subjects' initial assessment. This allowed the researcher to accurately inform each subject of their commitment without subsequent amendments. Only the researcher was aware of which group the subjects allocated.

3.13 Statistical Analysis

The statistical analysis in this study was done with the SPSS statistical program. P scores were calculated using matched pairs repeated measures. This analysis measured the change in scores when the treatment group received the acupuncture intervention and the changes in scores for the control over the wait period. The change in scores when the control group received the acupuncture intervention was also analysed. In addition, follow up analysis included comparisons of scores both before and after the acupuncture intervention. Since this study stratified subjects based on sex, the above analysis was repeated to assess the females and males separately. This was done over all the phases of the study, i.e. the interventions, the wait period and the follow up period. P scores were calculated to test any interaction which may have occurred between the sexes.

The study also included an intention to treat (ITT) analysis. In effect this type of analysis measures the effectiveness of the treatment whilst taking into consideration the amount of attrition experienced by subjects who could not complete the treatment program. This type of analysis better reflects the effectiveness of the treatment since it considered all subjects who attempted treatment. The ITT analysis involved only the depression measures, i.e. the Ham-D, BDI and the SCL 90 depression subscale. It was included across all the phases of the study and it was done with consideration of the stratification based on sex. Consideration was made for those who opt out of the study by assuming their level of depression would remain unchanged for the subsequent assessments in which they did not take part. This was done irrespective of which phase of the study the subject was in when they decided to withdraw from the study.

3.14 Intervention Procedure

After meeting inclusion criteria, subjects were randomly assigned to either the treatment group or the control group. Those assigned to the treatment group received the acupuncture intervention over an eight week period. Those assigned to the control group entered a wait period of eight weeks (equaling the length of time used in the acupuncture intervention). At the conclusion of the wait period, the subjects in the control group received the same acupuncture intervention as those in the treatment group. At the conclusion of the intervention, subjects in both groups entered a follow up phase that lasted for another eight weeks (refer to Figure 1).

Figure 1: Chart of Procedures
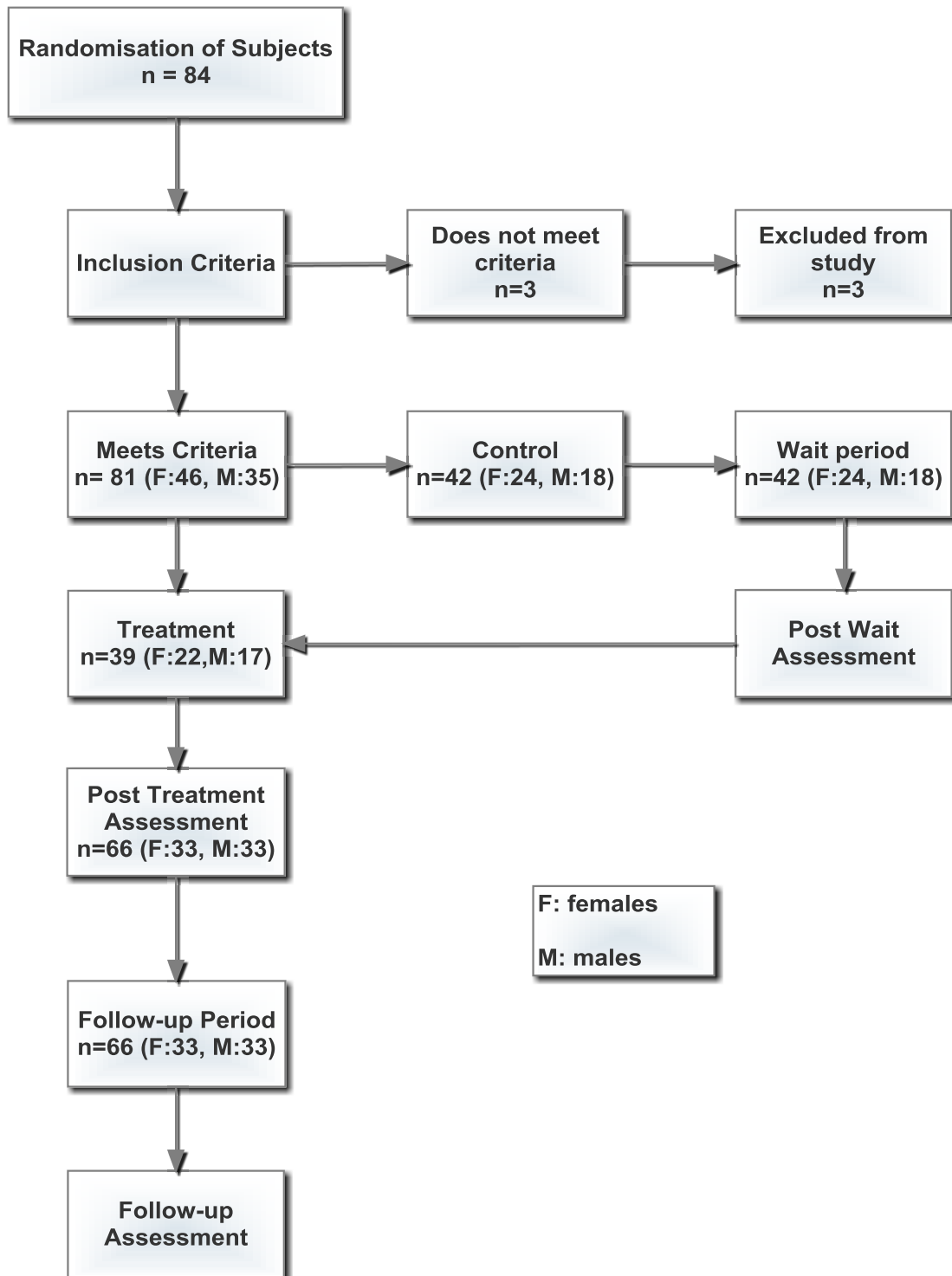
3.15 Assessments

3.15.1 Initial Assessment

Potential subjects were initially screened over the phone to check if they met basic inclusion/exclusion criteria. Those who met the requirements were randomly allocated into either the treatment or control groups and invited to attend an initial assessment. The initial assessment consisted of the liver qi stagnation checklist and the DSM-IV-TR structured interview for MDD. These two assessments served to confirm inclusion in the study. Subjects were required to read and sign an information sheet and consent form. Once consent to take part in the study was obtained, the initial assessment was carried out. This involved a full CM case history as well as the administration of the psychometric measures. The student administered all the initial assessments (refer to section 5.16.5: Practitioner Expertise – Assessments located in the section which addresses the limitations of this study). The Ham-D was administered first, followed by the BDI, then the STAI and finally the SCL 90.

3.15.2 Subsequent Assessments

Subsequent assessments were performed using the same measures. In each instance they were administered in the same order as the initial assessment. Those in the treatment group received subsequent assessments after the eight week treatment program and again at the conclusion of the eight week follow up period. Those in the wait list group received a subsequent assessment after the eight week wait period. This in effect is their initial assessment in respect to the acupuncture intervention. Subsequent assessments were administered after the eight week treatment program and again at the conclusion of the eight week follow up period. As was the case for the initial assessments, the student also administered all the subsequent assessments.

3.16 Treatment Program

The acupuncture was administered twice per week for the first four weeks, then once a week in the remaining four weeks. This is consistent with the consensus established by the literature which suggested this treatment program was best practice for an acupuncture intervention to treat depression. This treatment program will also facilitate ease of comparison of results across studies. It is reasonable to suggest it also reflects clinical practice, as it consists of ample treatments and enough time to produce an effect. All the treatments were administered by the researcher (the student).

All subjects were administered the three common or core acupoints, these being Liver 14 (Qimen), PC 6 (Neiguan) and LR (Taichong). In addition to the core acupoints, subjects also received an additional acupoint. This was determined according to which liver qi stagnation subgroup the subject best matched. The subjects who presented with simply liver qi stagnation, received the additional acupoint of Gall Bladder 34 (Yanglingquan). Those who showed symptoms of wood attacking earth received Spleen 6 (Sanyinjiao), and those who presented with liver qi stagnation and heat received Liver 8 (Ququan). Once a subject was assigned a prescription, that same prescription was administered for the entirety of the acupuncture intervention.

3.16.1 Standardisation of Point Location

The acupoints were located using the location descriptions offered by Deadman (1998). Standardisation using proportional measurement was used to locate those acupoints not directly located adjacent to anatomical structure or landmark. Fabricated expandable elastic straps were designed to reflect the number of equal units of space (*cun*) between respective anatomical locations used to locate the relevant acupoints. The units of space were marked out on the elastic straps with a marker. When the elastic straps were stretched between

two anatomical landmarks, the *cun* measurement expanded proportionally. This ensured consistency between subjects with differing body size and shape.

3.16.2 Treatment Procedure

This study employed a protocol consistent with standard practice in Australia. The needles were inserted and then "de qi" was achieved. The needles were left in situ for twenty minutes. This was considered to be ample time for a therapeutic effect to occur. There was no needle manipulation performed. The needles were inserted into and removed from the acupoints in order from the anatomically superior to inferior points. The depth of needling was standardised as noted in table 5.

Table 5: The depths of needling.

ACUPOINTS	NEEDLING DEPTH (CUN)
Liver 14 (Qimen)	0.5
Pericardium 6 (Neiguan)	1.0
Gall Bladder 34 (Yanglingquan)	1.0
Spleen 6 (Sanyinjiao)	1.0
Liver 8 (Ququan)	1.0
Liver 3 (Taichong)	0.5

3.17 Conclusion

This study implemented the elements of design which best suit an acupuncture RCT designed to treat depression. This was done by creating a synthesis model which satisfied the key requirements of scientific method whilst also maintaining the integrity of the CM model. Inclusion/exclusion criteria were defined from both psychological/psychiatric and CM perspectives. This led to the decision to use acupuncture as an adjunct therapy to SSRI antidepressant medication. Liver qi

stagnation was determined to be the main CM inclusion criteria (with the addition of liver qi stagnation with wood attacking earth and liver qi stagnation with heat). The isolation of these patterns of disharmony enabled close adherence to the CM model via the construction acupoint prescriptions which were specifically designed to treat each of the respective patterns of disharmony. Manualisation processes were used to identify a general consensus approach to assist the standardisation of both the CM diagnosis and acupoint prescriptions. A wait list control group was used as a method of control since it circumvents the problematic nature implicit in the other alternatives. It also enabled the clinical valuable question to be asked of whether is better to treat of not to treat.

Chapter 4: Results

To provide a full account of the subjects who were involved in the study, demographical information was collected and collated. This information included the age of the subjects; their highest level of education; the length of their depressive illness and their current depressive episode; as well as which SSRI they had been prescribed. Statistical analysis was also undertaken to ensure a homogeneity between subjects at the entry point of the study. The absence of significant differences improves the validity of the inferences which can be drawn from the statistical analysis. Base scores between the treatment group and the control group were analysed. Consideration was also made for the stratification of subjects based on sex. Base scores between the sexes were compared for the treatment group and the control group.

4.1   Subject Demographics

4.1.1   Age of Subjects

Table 6: The age of subjects.

		MAIN ANALYSIS	FEMALES	MALES
AGE (years)	M	37.13	32.42	41.55
	SD	11.64	9.17	11.96
AGE RANGE	20-29	16	14	2
	30-39	28	15	13
	40-49	14	2	12
	50+	8	2	6

Table 6 reports that the average age of the subjects in this study was 37.13 (SD = 11.64). This is consistent with the studies in the literature, particularly those

noted in the studies by Luo (1985) 39 years, Lou (1998) 36.9 years (SD = 2.3) and Han (2004) 35.1 years (SD = 12.63).

The most frequently represented age bracket was between 20-29 years of age. Two thirds of the subjects were under the age of 40 years and there were few subjects over the age of 50 years. Females had a much higher representation in the age bracket between 20-29 years and males were on average nine years older. The males had a higher representation in both the 40-49 years and the over 50 years age brackets. This representation of both sexes is consistent with data released from the Black Dog Institute. They reported that females most commonly experience a mood disorder between the ages of 25- 34 years and males between the ages of 35-44 years. (Black Dog Institute, 2012)

4.1.2 Highest Education of Subjects

Table 7: The subjects' highest level of education.

		MAIN ANALYSIS	FEMALES	MALES
EDUCATION	School Cert	14	8	6
	HSC	8	4	4
	Certificate	10	7	3
	Diploma	5	3	2
	Bachelor	23	10	13
	Masters	4	1	3
	PhD	2	0	2

As seen in table 7, almost half of the subjects had achieved a university education. The most frequently occurring highest level of education was at bachelor level. One third of the subjects however did not hold a qualification beyond school level. In general terms, the males had received more education than the females. A higher proportion of males had achieved a qualification of

bachelor level or above. They had a greater representation than females on all three qualifications of bachelor, masters and PhD. In contrast, more females had School Certificate and a Tertiary Certificate as their highest level of education.

4.1.3 Depression Profiles

Table 8: Subject depression profiles.

		MAIN ANALYSIS	FEMALES	MALES
ILLNESS	M	12.24	12.00	12.45
(years)	SD	8.22	7.07	9.14
EPISODE	M	3.18	4.50	1.98
(years)	SD	4.22	5.50	1.91

The subjects in this study had been suffering from a depressive illness for an average of 12.24 years (SD = 8.22). The average length of the current depressive episode at the time of the initial interview was 3.18 years (SD = 4.22). Refer to table 8. In contrast, the studies in the literature focused on subjects who were experiencing depression of a shorter duration. Only Han (1994) reported on the length of depressive illness since first diagnosed. The average length of depressive illness in his study was 40.94 months (SD = 50.24). This is much shorter in time than the depressive illness experienced by the subjects in this study. Three studies in the literature reported the average length of depressive episode. Those who received the electro acupuncture intervention on Luo (1985) had experienced an average episode of 6.3 months. This is similar to average depressive episodes of 7.9 months in Luo (1998) and 9.47 months (SD = 10.06) in Han (2004). These three averages are significantly shorter in duration than experienced in this study.

There was only one study in the literature which reported the length of depressive episode in relation to sex. The female subjects in Allen (1998) were

reported to experience an average depressive episode of 9.2 months (SD = 6.9). This again is much shorter in duration than experienced in this study where the average current episode for females was 4.50 years (SD = 5.50). No studies in the literature reported on the length of male depression (neither in terms of depressive illness or episode). The sample populations used in this study demonstrated an important difference between females and males. Despite both sexes experiencing similar lengths of depressive illness, their length of depressive episodes were very different. Males experienced an average length of depressive episode of 1.98 years (SD = 1.91).

There were studies in the literature which used exclusion criteria to omit subjects who were suffering from a long-term depressive episode. Manber (2004), Quah-Smith (2005) and Allen (2006) excluded potential subjects if they were experiencing an episode of two years or more. In effect, this may have created a positive bias with these studies. The over inclusion of sufferers of acute depression means there is a greater likelihood of improvement in depression scores and/or a greater likelihood of remission.

It is interesting to note that in studies by Manber (2004), Quah-Smith (2005) and Allen (2006), the exclusion clause in each case was coupled with the exclusion of dysthymia. This may have implied a similarity or confusion between a long term MDE (Major Depressive Episode) and dysthymia. It must be noted these are very different conditions. Dysthymia is characterised by chronic feelings of a depressed mood of a lesser severity. In contrast, an MDE requires symptoms to be the source of significant distress and to cause impairment in the ability to function. An MDE can be long term and not be dysthymia.

The researchers in this study chose to include all sufferers of MDD irrespective of the length of their depressive episode. To exclude subjects based on the length of depressive episode was considered by the researchers to be contrary to ethical standards. Moreover, the researchers wanted to ensure an RCT

design that better reflected clinical practice, where patients are generally not excluded from treatment on the basis of length of depressive episode. In addition, the exclusion of long-term depression sufferers was not considered to best reflect a clinical setting. In contrast to the studies in the literature, the inclusion of long-term MDD sufferers may have created a negative bias within this study as such subjects would be expected to be more difficult to treat.

4.1.4 The SSRI's Prescribed to Subjects

Table 9: Frequency of SSRI's prescribed to subjects.

		MAIN ANALYSIS	FEMALES	MALES
MEDICATION	Citrilopram	12	8	4
	Escitalopram	19	11	8
	Fluoxetine	8	4	4
	Fluvoxamine	2	0	2
	Paroxetine	4	2	2
	Sertraline	21	8	13

As noted in table 9 above, only three generic drug brand names had significant representation in the study. Escitalopram and Sertraline were clearly the most frequently prescribed. Citrilpram had the third greatest frequency. In contrast, Fluoxetine, Fluvoxamine and Paroxetine each had a small representation. The distribution of the prescribed generic names for the female subjects followed a similar trend as the main analysis. In contrast, the male subjects had a more even distribution across all the generic names which were represented in the study. They did however have a greater frequency of Escitalopram compared to the other generic names.

4.2 Base Statistics

4.2.1 Depression Measures

T-tests were carried out to detect differences in base scores between and within groups at the entry point of the study (refer to table 10). Similarities between groups at entry point are used to validate the statistical analysis. When groups are shown to be similar before an intervention, it authenticates the assumptions that the change in scores may be attributable to the intervention. This analysis was done across all the depression measures i.e. the BDI, Ham-D and the SCL 90 depression subscale. Comparisons were made between the treatment group and the control group as well as between the sexes within the treatment group and between the sexes within the control group.

Statistically significant differences were found between those in the treatment group to those in the control group according to both the Ham-D (α = 0.05, t = 0.024) and the Ham-D ITT (α = 0.05, t = 0.029). Subjects in the treatment group had significantly higher base scores than those in the control group. This may indicate the Ham-D has a different sensitivity in measuring the levels of depression of the subjects in each group. This may influence the analysis of the outcomes involving the Ham-D when the two groups are compared.

Statistically significant differences were also found between females and males in the treatment group. The females had significantly higher base scores than did the males on both the Ham-D (α = 0.05, p = 0.004) and the Ham-D ITT (α = 0.05, p = 0.007). The females in the treatment group were also found to have significantly higher scores than the males in the treatment group on the BDI (α = 0.05, t = 0.011). This may indicate the BDI and the Ham-D has different sensitivity in measuring the levels of depression between the sexes in the treatment group. In contrast no differences were found in base scores between the sexes in the control group according to the BDI and Ham-D.

Table 10: Base statistics on depression measures.

TEST		MAIN ANALYSIS (n=66)			SEX COMPARISONS					
					TREATMENT GROUP			CONTROL GROUP		
		TREAT. (n=33)	CONT. (n=33)	t TEST	FEMALES (n=16)	MALES (n =17)	t TEST	FEMALES (n=17)	MALES (n=16)	t TEST
BECK	M	27.39	28.30	0.705	32.38	22.71	**0.011**	30.12	26.38	0.177
	SD	11.23	7.89		11.67	8.75		7.25	8.31	
BECK ITT	M	27.95	29.07	0.601	32.00	31.17	0.762	22.71	26.28	0.211
	SD	11.02	7.81		11.05	7.28		8.75	7.81	
HAM-D	M	22.55	19.97	**0.024**	25.19	20.53	**0.004**	20.06	19.38	0.645
	SD	5.42	3.36		4.98	3.04		4.68	3.67	
HAM-D ITT	M	22.67	20.45	**0.029**	24.68	21.33	**0.007**	20.06	19.28	0.577
	SD	5.16	3.53		4.66	3.38		4.68	3.46	
SCL DEP	M	26.70	29.18	0.313	30.13	30.88	0.831	23.47	27.38	0.236
	SD	10.79	8.97		10.99	9.21		9.84	8.62	
SCL DEP ITT	M	27.05	30.62	0.118	29.82	33.00	0.306	23.47	27.44	0.201
	SD	11.07	9.21		11.38	9.42		9.84	8.13	

4.2.2 Secondary Measures

Base statistics were also done using T-Test on the STAI and the SCL 90 (refer to table 11). This analysis considered the STAI global scale and the Y1 and Y2 subscales as well as the SCL 90 global scale and all it's subscales. As done with the depression measures, comparisons were made between the treatment group and the control group as well as between the sexes within the treatment group and between the sexes within the control group.

The only statistically significant difference in base scores was found in the control group. Males had higher base scores than females according to the SCL interpersonal sensitivity subscale ($\alpha = 0.05$, p = 0.038). This indicates the subscales may have different sensitivity in measuring their interpersonal sensitivity between the sexes. This may need to be considered when comparisons are made between the sexes on the interpersonal sensitivity subscale.

Table 11: Base statistics on STAI and SCL 90.

TEST		MAIN ANALYSIS (n=66)			SEX COMPARISONS					
					TREATMENT GROUP			CONTROL GROUP		
		TREAT.	CONT.		FEMALES	MALES		FEMALES	MALES	
		(n=33)	(n=33)	t TEST	(n=16)	(n=17)	t TEST	(n=17)	(n=16)	t TEST
STAI	M	112.79	113.24	0.919	119.06	113.06	0.349	106.88	113.44	0.294
	SD	18.63	17.62		17.02	19.11		18.60	16.51	
STAI - Y1	M	52.76	52.82	0.982	54.63	52.29	0.547	51.00	53.38	0.523
	SD	10.69	10.68		11.02	10.95		10.39	10.73	
STAI - Y2	M	59.94	60.15	0.930	64.44	60.76	0.221	55.71	59.50	0.291
	SD	10.24	9.17		6.91	9.66		11.20	8.88	
SCL - ANX	M	13.12	14.36	0.547	14.94	13.18	0.532	11.41	15.63	0.171
	SD	7.11	9.41		6.89	8.93		7.07	10.04	
SCL - GLOB	M	128.09	140.85	0.354	150.94	143.71	0.716	106.59	137.81	0.090
	SD	53.83	57.12		53.69	59.23		45.62	56.56	
SCL - SOM	M	13.58	14.58	0.625	15.56	15.71	0.963	11.71	13.38	0.538
	SD	7.95	8.57		8.28	9.15		7.37	8.02	
SCL - OC	M	18.03	20.42	0.185	20.75	22.00	0.650	15.47	18.75	0.131
	SD	7.76	6.69		8.36	7.33		6.39	5.70	
SCL - IS	M	15.30	16.91	0.418	19.50	17.24	0.431	11.35	16.56	**0.038**
	SD	8.31	7.68		7.83	8.44		6.80	7.04	
SCL - P.ANX	M	5.12	6.24	0.446	6.25	5.65	0.794	4.06	6.88	0.135
	SD	5.10	6.67		5.52	7.44		4.58	5.91	
SCL - HOST	M	7.36	7.27	0.944	5.35	7.56	0.224	9.50	7.00	0.168
	SD	5.61	4.84		4.68	5.55		5.85	4.23	
SCL - PID	M	8.00	8.42	0.740	9.94	7.71	0.248	6.18	9.19	0.077
	SD	4.68	5.62		5.20	5.67		3.34	5.66	
SCL - PSYCH	M	9.03	11.03	0.200	10.69	11.53	0.703	7.47	10.50	0.173
	SD	6.32	6.23		5.91	6.60		6.47	5.99	

4.2.3 Dropouts

A total of 81 subjects were took part in this study. This was divided into 39 in the treatment group and 42 in the control group. This was lower than the rate of withdrawal which was expected by the researchers at the inception of the study. Considering the subjects in the study were experiencing significant levels of depression, anxiety and a range of other psychological symptoms, it was thought the withdrawal rate would be higher. A total of 15 subjects chose to withdraw their participation in the study. Six of these were from the treatment group, and nine from the control group. It was expected the control group would have a greater amount of withdrawals. The design of the study required them to undertake an extra eight weeks of involvement due to their added requirement to take part in the wait period.

The majority of those who chose to end their involvement in the study were female. A total of 13 females withdrew compared to only two males. Six of these females were in the treatment group and 7 were in the control group. The two males who withdrew from the study were both in the control group. This means a total of 46 females were involved in the ITT analysis. 22 females being in the treatment group, and 24 were in the control group. A total of 35 males were involved in ITT analysis. 18 of these were in the control group whilst the treatment group remained to the standard analysis with 17 male subjects.

4.2.4 Adverse Events

There we no adverse events experienced during the study

4.3 Reporting of the Results

Given that depression is the focus of this study, this chapter first addresses the results from the depression scales; followed by analysis of the diagnostic criteria for depression; anxiety and finally general mental health.

The analysis was undertaken using the SPSS statistical program. P scores were calculated using matched pairs repeated measure t-tests. This was done for all the groups across the relevant phases of the study. Analysis of variance (ANOVA) based on repeated measures was also used to test interactions between the sexes.

The statistical outcomes for each section of the study (being depression, anxiety and general mental health measures) are displayed using a consistent structure of three tables for each section presented as follows:.

- tables labeled with **"treatment group - intervention"** in the title refer to the trial phase which reports the results from the treatment group that received acupuncture intervention and SSRI medication.
- tables labeled with **"control group – wait period"** in the title refer to the trial phase which reports the results when the control group received no acupuncture and only continued taking their SSRI's.
- tables labeled with **"control group – intervention"** in the title refer to the trial phase which reports the results from when the control group were later administered acupuncture along with their SSRI's.

4.3.1 Stratification

Data for the males and females within the study were stratified by sex. In effect four groups were created: a female treatment group; a male treatment group; a female control group; and a male control group. Analysis taking into consideration stratification based on sex was applied to all measures in each section of the study (i.e. depression, anxiety and general mental health). The same reporting structure using the same table format has been applied to reporting the stratified results within each of the sections. The results for females are reported first, then the results for males and then the analysis of sex interaction which compares the rate of the change in scores between the sexes.

The inclusion criteria in this study noted three variations of liver qi stagnation. These being: liver qi stagnation; liver qi stagnation with wood attacking earth; and liver qi stagnation with heat. It may be proposed to stratify subjects for analysis based on these three sub diagnoses. This in turn would pose the question as to which sub diagnoses of liver qi stagnation could be best treated to reduce depression, anxiety and improve general mental health. The researchers decided this investigation would be of little value. Firstly, the treatment prescriptions applied to the three diagnostic categories shared a core set of acupoints and would not offer a sufficient point of difference both statistically and clinically. Secondly, the sample size in each sub pattern would not likely produce statistically significant differences. Thirdly, given the diagnoses of wood invading earth and liver qi stagnation with heat were undertaken in the context of sub-patterns of liver qi stagnation for purposes of inclusion, analysis of these subsets of data would be of little statistical or clinical interest.

4.4   Depression

4.4.1 Main Analysis

The following tables report the statistical outcomes from the depression analysis. The subjects involved in each of the three phases of the study were administered the depression scales before and after each phase. The depression measures used in this study were the Beck Depression Inventory (BDI), the Hamilton Rating Scale for Depressive Illness (Ham-D) and the Systematic Checklist 90 depression subscale. An intention to treat analysis (ITT) was conducted using each scale in all three phases of the study and is reported immediately below the results of each measure.

In each table, the mean scores (M) and their corresponding standard deviations (SD) from both before (PRE) and after (POST) the respective phase of the study

are reported alongside each measure. Also listed adjacently are the P scores (P). These give the statistical significance in the change in the corresponding mean scores for that measure during the respective phase of the study.

4.4.2 Treatment Group

Table 12: Depression - Main Analysis – Treatment Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	33	27.39	11.233	12.03	11.240	0.000
BECK ITT	39	27.95	11.019	14.95	12.957	0.000
HAM-D	33	22.55	5.420	10.82	6.121	0.000
HAM-D ITT	39	22.67	5.162	12.74	7.369	0.000
SCL DEP	33	26.70	10.789	12.70	11.137	0.000
SCL DEP ITT	39	27.05	11.067	15.21	12.794	0.000

- Pre and post intervention comparisons for the acupuncture treatment group showed significant improvement in the degree of depression across all depression scales (BDI, Ham-D and SCL 90 depression subscale).
- The intention to treat analysis for the acupuncture treatment group also showed significant reductions in the degree of depression across all depression measures.

4.4.3 Control Group

Table 13: Depression - Main Analysis - Control Gp – Wait Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	33	28.30	7.892	27.15	9.365	0.371
BECK ITT	42	29.07	7.813	27.45	8.744	0.137
HAM-D	33	19.97	3.359	19.27	4.118	0.313
HAM-D ITT	42	20.45	3.528	19.74	4.423	0.221
SCL DEP	33	29.18	8.967	27.21	10.343	0.210
SCL DEP ITT	42	30.62	9.213	28.19	9.993	0.075

- The control group experienced no significant change in their depression over pre and post comparisons during the wait period on the BDI, Ham-D and the SCL depression subscale.
- The intention to treat analysis also showed no significant change in the degree of depression experienced over the wait period on all depression measures.

Table 14: Depression - Main Analysis - Control Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	33	27.15	9.365	12.12	8.268	0.000
BECK ITT	42	27.45	8.744	15.64	10.373	0.000
HAM-D	33	19.27	4.118	9.61	4.123	0.000
HAM-D ITT	42	19.74	4.423	12.14	6.554	0.000
SCL DEP	33	27.21	10.343	12.03	8.099	0.000
SCL DEP ITT	42	28.19	9.993	16.26	11.455	0.000

- When the control group were administered the acupuncture intervention they showed significant improvement in the degree of their depression on the BDI, Ham-D and the SCL 90 depression subscale
- The intention to treat analysis also indicated the control group experienced a significant reduction in their depression across all depression measures.

4.4.4 Follow Up

The reporting of the results from the follow up period differs in how it is presented since it involves two tables. The first table will report the results over the follow up period. This will involve comparisons between scores at the **end** of the intervention to those at the **end** of the follow up period. The second table will report the comparisons between scores **before** the intervention to those at the **end** of the follow up.

Table 15: Depression - Main Analysis – Follow Up Period

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	66	12.08	9.791	12.58	11.504	0.618
BECK ITT	81	15.31	11.620	15.72	12.729	0.618
HAM-D	66	10.21	5.214	11.20	8.644	0.267
HAM-D ITT	81	12.43	6.921	13.23	9.114	0.266
SCL DEP	66	12.36	9.668	12.05	10.940	0.748
SCL DEP ITT	81	15.75	12.054	15.49	12.976	0.748

- There was no significant change in the rates of depression on the BDI, Ham-D and the SCL 90 depression subscale over the eight week follow up period.

94

- The intention to treat analysis also indicated no significant change in the depression experienced over the follow up period across all three measures.

Table 16: Depression - Main Analysis – Pre Intervention to Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	66	27.27	12.580	12.58	11.504	0.000
BECK ITT	81	27.69	9.845	15.72	12.729	0.000
HAM-D	66	20.91	5.053	11.20	8.644	0.000
HAM-D ITT	81	21.15	4.985	13.23	9.114	0.000
SCL DEP	66	26.95	10.490	12.05	10.940	0.000
SCL DEP ITT	81	27.64	10.473	15.49	12.976	0.000

- When comparing the pre intervention scores to the follow up scores, there was a significant improvement in the rates of depression on all three measures. This was also the case for the intention to treat analysis.
- These results indicate subjects were able to maintain their post treatment improvement in their depression scores (refer to table 12) and remain significantly less depressed than before the intervention phase of the study.

4.5   Stratification Based on Sex

The following data shows the results stratified according to sex. The results for females in their respective treatment and control groups as well as their follow up analysis are reported first. The analysis for the men will follow. The stratified results are reported using the same tables, structure, labeling and abbreviations as the main analysis. The sex interaction analysis follows the results from the male groups. It compares the relative improvement of females and males during

the intervention phase and over the follow up period on all three depression scales.

4.6 Female Subjects

4.6.1 Treatment Group

Table 17: Depression - Females – Treatment Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	16	32.38	11.673	17.63	12.847	0.001
BECK ITT	22	32.00	11.045	21.27	13.392	0.002
HAM-D	16	25.19	4.983	14.13	6.228	0.000
HAM-D ITT	22	24.68	4.664	16.64	6.973	0.000
SCL DEP	16	30.13	10.990	17.81	12.210	0.001
SCL DEP ITT	22	29.82	11.379	20.86	13.246	0.002

- Pre and post intervention comparisons for females in the acupuncture treatment group showed significant improvement in the degree of depression across all depression scales (DBI, Ham-D and SCL 90 depression subscale).
- The intention to treat analysis for females in the acupuncture treatment group also showed significant reductions in the degree of depression across all depression measures.

4.6.2 Control Group

Table 18: Depression - Females – Control Gp – Wait Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	17	30.12	7.253	28.76	8.035	0.396
BECK ITT	24	31.17	7.281	29.13	7.368	0.117
HAM-D	17	20.53	3.044	20.00	3.841	0.589
HAM-D ITT	24	21.33	3.384	20.63	4.271	0.336
SCL DEP	17	30.88	9.212	27.06	11.729	0.096
SCL DEP ITT	24	33.00	9.418	28.83	11.115	0.027

- Females in the control group experienced no significant change in their level of depression during the wait period on the BDI, Ham-D and the SCL 90 depression subscale.
- The intention to treat analysis indicated female subjects did not change in their degree of depression over the wait period according to the BDI and the Ham-D.
- Females in the intention to treat analysis did demonstrate significant improvement in the depression on the SCL 90 subscale during the wait period.

Table 19: Depression - Females – Control Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	17	28.76	8.035	12.53	9.247	0.000
BECK ITT	24	29.13	7.368	17.63	11.590	0.000
HAM-D	17	20.00	3.841	9.41	4.556	0.000
HAM-D ITT	24	20.63	4.271	13.13	7.508	0.000
SCL DEP	17	27.06	11.729	9.88	6.918	0.000
SCL DEP ITT	24	28.83	11.115	16.67	13.031	0.000

- When females in the control group received the acupuncture intervention, they demonstrated a statistically significant reduction in their degree of depression on all measures – Beck, Ham-D and SCL.
- The intention to treat analysis also showed that when females in the control group received the acupuncture intervention they significantly reduced their depression.

4.6.3 Follow Up

Table 20: Depression - Females – Follow Up Period

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	33	15.00	11.261	16.21	12.626	0.433
BECK ITT	46	19.37	12.480	20.24	13.095	0.430
HAM-D	33	11.70	5.855	13.45	6.874	0.061
HAM-D ITT	46	14.80	7.393	16.07	7.517	0.062
SCL DEP	33	13.73	10.489	15.48	12.125	0.261
SCL DEP ITT	46	18.67	13.159	19.93	13.688	0.259

- Female subjects showed no significant change in their rates of depression over the follow up period on all three measures. This was also the case for the intention to treat analysis.

Table 21: Depression - Females – Pre Intervention to Follow Up Period

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	33	30.52	9.975	16.21	12.626	0.000
BECK ITT	46	30.50	9.316	20.24	13.095	0.000
HAM-D	33	22.52	5.094	13.45	6.874	0.000
HAM-D ITT	46	22.57	4.866	16.07	7.517	0.000
SCL DEP	33	28.55	11.306	15.48	12.125	0.000
SCL DEP ITT	46	29.30	11.127	19.93	13.688	0.000

- When comparing the pre intervention scores with the follow up scores for female subjects, there was a significant improvement in the rates of depression on all three measures. This was also the case for the intention to treat analysis.
- These results indicate female subjects were able to maintain their post treatment improvement in their depression scores (refer to table 17) and remain significantly less depressed than before they received the intervention.

4.7   Male Subjects

4.7.1   Treatment Group

Table 22: Depression - Males – Treatment Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	17	22.71	8.745	6.76	6.098	0.000
BECK ITT	17	22.71	8.745	6.76	6.098	0.000
HAM-D	17	20.06	4.683	7.71	4.165	0.000
HAM-D ITT	17	20.06	4.683	7.71	4.165	0.000
SCL DEP	17	23.47	9.837	7.88	7.582	0.000
SCL DEP ITT	17	23.47	9.837	7.88	7.582	0.000

- Male subjects in the treatment group demonstrated a statistically significant reduction in their degree of depression between pre and post intervention comparisons on all measures – Beck, Ham-D and SCL 90 depression subscale.
- The intention to treat analysis also indicated males in the treatment group significantly reduce their level of depression due to acupuncture intervention.

4.7.2 Control Group

Table 23: Depression - Males – Control Gp – Wait Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	16	26.38	8.310	25.44	10.589	0.660
BECK ITT	18	26.28	7.813	25.22	10.085	0.581
HAM-D	16	19.38	3.667	18.50	4.382	0.393
HAM-D ITT	18	19.28	3.461	18.56	4.462	0.462
SCL DEP	16	27.38	8.617	27.38	9.025	1.000
SCL DEP ITT	18	27.44	8.126	27.33	8.506	0.955

- Male subjects in the control group showed no significant change in their depression scores over the wait period on all three measures.
- The intention to treat also indicated the males in the control group showed no significant change in their levels of depression over the wait period.

Table 24: Depression - Males – Control Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	16	25.44	10.589	11.69	7.364	0.000
BECK ITT	18	25.22	10.085	13.00	8.051	0.000
HAM-D	16	18.50	4.382	9.81	3.746	0.000
HAM-D ITT	18	18.56	4.462	10.83	4.914	0.000
SCL DEP	16	27.38	9.025	14.31	8.837	0.000
SCL DEP ITT	18	27.33	8.506	15.72	9.285	0.000

- Male subjects in the control group demonstrated a statistically significant reduction in their degree of depression when they were administered the acupuncture intervention on all three measures.

- The intention to treat analysis also indicated males showed statistically significant improvement in their levels of depression on pre and post acupuncture intervention analysis.

4.7.3 Follow Up

Table 25: Depression - Males – Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	33	9.15	7.089	8.94	9.062	0.872
BECK ITT	35	9.97	7.740	9.77	9.499	0.871
HAM-D	33	8.73	4.049	8.94	9.695	0.889
HAM-D ITT	35	9.31	4.770	9.51	9.775	0.889
SCL DEP	33	11.00	8.718	8.61	8.467	0.046
SCL DEP ITT	35	11.91	9.272	9.66	9.299	0.047

- After the follow up period male subjects showed no significant change in their rates of depression compared to their post intervention scores on the Beck and Ham-D. These outcomes were repeated in the intention to treat analysis.

- Male subjects did demonstrate significant improvement in their depression according to the SCL 90 depression subscale over the follow up period. This was again repeated in the intention to treat analysis. This indicates males continued to significantly improve after the acupuncture intervention had concluded.

Table 26: Depression - Males – Pre Intervention to Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
BECK	33	24.03	9.629	8.94	9.062	0.000
BECK ITT	35	24.00	9.406	9.77	9.499	0.000
HAM-D	33	19.30	4.538	8.94	9.695	0.000
HAM-D ITT	35	19.29	4.567	9.51	9.775	0.000
SCL DEP	33	25.36	9.513	8.61	8.467	0.000
SCL DEP ITT	35	25.46	9.249	9.66	9.299	0.000

- Males were still significantly less depressed after the follow up period compared to their pre intervention scores on all three measures. These outcomes were again repeated in the intention to treat analysis.

- The results indicate the male subjects were able to maintain or improve their post treatment improvement in their depression scores (refer to table 22) and remain significantly less depressed than when they entered the intervention phase of the study.

4.8 Sex Interactions

4.8.1 Intervention Phase

Table 27: Depression - Sex Interactions - Intervention.

TEST	n	ANOVA
BECK	66	0.738
BECK ITT	81	0.153
HAM-D	66	0.691
HAM-D ITT	81	0.068
SCL DEP	66	0.141
SCL DEP ITT	81	0.014

- Over the intervention phase of the study, there was no significant differences in the rate of improvement between females and males on the BCI, Ham-D and SCL depression subscale.

- The intention to treat analysis also showed no significant difference in the rate of improvement between females and males for the BDI and Ham-D. In contrast, the intention to treat analysis using the SCL 90 depression subscale indicated males had significantly greater improvement in their scores than females.

4.8.2 Follow Up Phase

Table 28: Depression – Sex Interactions – Follow Up Period.

TEST	n	ANOVA PRE - FU	ANOVA POST - FU
BECK	66	0.768	0.480
BECK ITT	81	0.116	0.581
HAM-D	66	0.530	0.383
HAM-D ITT	81	0.083	0.467
SCL DEP	66	0.172	0.035
SCL DEP ITT	81	0.011	0.029

- According to the follow up scores on the BDI and Ham-D, female and male subjects demonstrated no significant difference in their rates of improvement in their depression. Comparisons between the follow up scores to both the pre and post intervention scores showed no significant differences. This was again the case for the intention to treat analysis.

- The SCL 90 subscale did show significant differences between females and males in their rates of improvement. Comparisons on follow up scores to post intervention scores indicated males continued to show improvement (refer to table 22) over the follow up period whereas females showed a mild increase in their rates of depression (refer to table 17). This was again evident in the intention to treat analysis.

- Follow up and pre intervention comparisons for the SCL 90 subscale did not show a significant difference between females and males. The intention to treat analysis did however indicate the males did improve significantly more than females.

4.9   DSM-IV-TR Diagnosis

Analysis was also carried out using the DSM-IV. The purpose of these comparisons was to quantify the affect the acupuncture intervention had in regards to diagnostic status as opposed to symptom severity (as done in the depression scales). This diagnostic analysis considered changes in the status of depression in the subjects over each of the phases within the study. Subjects were deemed to have a "positive" diagnostic status if they satisfied the criteria for MDD. Conversely, if subjects were not able to satisfy the criteria for MDD they were deemed to have a "negative" diagnostic status. The DSM-IV structured clinical interview considers a total of nine diagnostic criteria. A total of five of these criteria need to be met to satisfy the diagnosis of MDD. When the diagnosis for an individual changes from a positive to a negative, this is considered to be remission.

The DSM IV analysis will be reported in tables 29-31. Table 29 will report the main analysis (when both sexes were considered as a homogenous group) and tables 30 and 31 will report on the analysis for females and males respectively.

Each table breaks the analysis up into five comparisons. In order of appearance, these are:
- The intervention phase for those in the treatment and the control group combined;
- The treatment group when they received the intervention;
- The control group over the wait period;
- The control group when they received the intervention; and
- The follow up period which combined those in the treatment and the control group.

There are a total of eight calculations on each table. Three of them are repeated on a pre and post basis. These were determined in respect to which phase of the study was being analysed. The eight calculations are:

- The number of those who had a positive diagnosis at the start of the phase in question;

- The average number of DMS-IV criteria which were met before the phase in question;

- The percentage of those who received a positive diagnosis before the phase in question;

- The number of those who had a positive diagnosis at the conclusion of the phase in question;

- The average number of DMS-IV criteria which were met after the phase in question;

- The percentage of those who received a positive diagnosis after the phase in question;

- The number of those who changed their diagnostic status during the phase in question (this is the number of subjects who achieved remission); and

- The percentage of those who changed their diagnostic status (this a rate of remission).

4.9.1 Main Analysis

Table 29: Depression – Main Analysis – DSM IV.

	Pre +'ve diagnosis	Pre average number of criteria met	Pre % +'ve diagnosis	Post +'ve diagnosis	Post average number of criteria met	Post % +'ve diagnosis	# of diagnosis changed	% of diagnosis change
Treatment & Control Gps: Intervention (n=66)	63	6.68	95.45	8	2.48	12.12	55	87.30
Treatment Gp: Intervention (n=33)	33	7.18	100.00	4	2.61	12.12	29	87.88
Control Wait Period: control Gp (n=33)	33	6.70	100.00	30	6.55	90.91	3	9.09
Control Gp: Intervention (n=33)	30	6.55	90.91	4	0.20	12.12	26	86.67
Treatment & Control Gps: Follow Up (n=66)	8	2.48	12.12	14	2.58	21.21	-6	-10.34

- Before the acupuncture intervention, 63 out of the 66 subjects (95.45%) were diagnosed as suffering from MDD according to DSM-IV-TR. Three subjects in the control group experienced remission over the wait period. Across these 66 subjects, the pre intervention diagnoses saw subjects meet an average of 6.68 out of the nine criteria for MDD. After the intervention, 8 subjects (12.12%) maintained a positive diagnosis with an average of 2.48 of the diagnostic criteria being met. This means 55 of the 63 subjects (who were initially diagnosed with MDD) no longer satisfied the criteria. This is a remission rate of 87.30%.

- In the treatment group, all 33 subjects (100%) fulfilled the diagnostic criteria of MDD before the intervention with an average of 7.18 of the diagnostic criteria being met. Post intervention, only 4 subjects (12.12%) continued to maintain their diagnosis with an average of 2.61 of the diagnostic criteria being met across the 33 subjects. This is a remission rate of 87.88% of the subjects.

- Before the wait period, all 33 subjects (100%) in the control group were diagnosed with major depressive disorder with an average of 6.70 of the diagnostic criteria being met. After the wait period, 30 subjects (90.91%) in the control group continued to satisfy the diagnosis with an average of 6.55 of the diagnostic criteria being met. Three subjects were no longer diagnosed with MDD, this is a remission rate of 9.09%.

- Due to 3 subjects no longer fulfilling the diagnostic criteria after the wait period, the control group began the intervention with 30 out the 33 subjects being diagnosed with MDD. After the intervention, only 4 out of the remaining 30 clinical subjects continued to maintain a positive diagnosis of MDD. The degree of improvement was so high that on average only 0.20 of the criteria were met. Since 26 subjects no longer satisfied the diagnosis there was a remission rate of 86.67% of subjects.

- Over the follow up period, 10.34% of subjects fell out of remission. This means an extra 6 subjects once again met the criteria for MDD. After the

acupuncture intervention 8 subjects remained depressed. When these 8 subjects were combined with the 6 who regressed, a total of 14 subjects were diagnosed as depressed at the conclusion of the wait period. In percentage terms, 21.12% of patients satisfied the diagnostic criteria for MDD at the conclusion of the follow up period.

4.9.2 Female Subjects

Table 30: Depression – Females – DSM IV.

	Pre +'ve diagnosis	Pre average number of criteria met	Pre % +'ve diagnosis	Post +'ve diagnosis	Post average number of criteria met	Post % +'ve diagnosis	# of diagnosis changed	% of diagnosis change
Treatment & Control Gps: Intervention (n=33)	33	7.24	100.00	5	2.97	15.15	28	84.85
Treatment Gp: Intervention (n=16)	16	7.75	100.00	3	3.44	18.75	13	81.25
Control Wait Period: control Gp (n=17)	17	6.65	100.00	17	6.76	100.00	0	0
Control Gp: Intervention (n=17)	17	6.76	100.00	2	2.53	11.67	15	88.23
Treatment & Control Gps: Follow Up Period: (n=33)	5	2.97	15.15	9	3.18	27.27	-4	-14.29

- Before the acupuncture intervention, all 33 female subjects (100%) were diagnosed as suffering from MDD according to DSM-IV. Across these 33 subjects, the pre intervention diagnoses saw subjects meet an average of 7.24 of the criteria. After the intervention, 5 female subjects (15.15%) fulfilled a positive diagnosis with an average of 2.97 of the diagnostic criteria being met. This means 28 out the initial 33 female subjects who were initially diagnosed with MDD no longer satisfied the criteria. This is a remission rate of 84.85% for female subjects.

- In the treatment group, all 16 female subjects (100%) fulfilled the diagnostic criteria for MDD before the intervention with an average of 7.75 of the diagnostic criteria being met. Post intervention, only 3 subjects (18.75%) continued to fulfill the diagnosis with an average of 3.44 of the diagnostic criteria being met. With 13 female subjects changing their diagnosis status, this is a remission rate of 81.25% for female subjects.

- Before the wait period, all 17 female subjects (100%) in the control group were diagnosed with MDD with an average of 6.65 of the diagnostic criteria being met. After the wait period, all 17 subjects (100%) in the control group continued to satisfy the diagnosis with an average of 6.76 of the diagnostic criteria being met. Since no female subjects changed their diagnosis status the remission rate for females in the wait period was 0.00%.

- Since no female subjects changed their diagnostic status during the wait period, the control group began the intervention with all 17 female subjects being diagnosed with MDD. After the intervention, only 2 of the subjects continued to maintain their depressed status. On an average, the number of criteria met was reduced down to 2.53. A total of 15 female subjects no longer satisfied the diagnosis for MDD. This is a remission rate of 88.23% for female subjects in the control group when they received the acupuncture.

- Over the follow up period, 14.29% of female subjects fell out of remission. This means an extra 4 subjects once again met the criteria for MDD. After the acupuncture intervention 9 subjects remained depressed. When these 9 subjects were combined with the 4 who regressed, there were a total of 13 subjects who were diagnosed as depressed at the conclusion of the wait period. In percentage terms, 27.27% of female patients satisfied the diagnostic criteria for depression at the conclusion of the follow up period.

4.9.3 Male Subjects

Table 31: Depression – Males – DSM IV.

	Pre +'ve diagnosis	Pre average number of criteria met	Pre % +'ve diagnosis	Post +'ve diagnosis	Post average number of criteria met	Post % +'ve diagnosis	# of diagnosis changed	% of diagnosis change
Treatment & Control Gps: Intervention (n=33)	30	6.48	90.91	3	2.00	9.09	27	90.00
Treatment Gp: Intervention (n=17)	17	6.65	100.00	1	1.82	5.88	16	94.12
Control Wait Period: control Gp (n=16)	16	6.35	100.00	13	6.31	81.25	3	18.75
Control Gp: Intervention (n=16)	13	6.31	81.25	2	2.19	12.50	11	84.62
Treatment & Control Gps: Follow Up Period: (n=66)	3	2.00	9.09	5	1.97	15.15	-2	-6.67

- Before the acupuncture intervention, 30 out of the whole 33 male subjects (90.91) were diagnosed as suffering from MDD according to DSM-IV across these 33 subjects, the pre intervention diagnoses saw male subjects meet an average of 6.48 of criteria for MDD. After the intervention, three male subjects (9.09%) maintained their positive diagnosis with an average of 2.00 of the diagnostic criteria being met. This means 27 out of the initial 33 male subjects who were initially diagnosed with MDD no longer satisfied the criteria. This is a remission rate of 90.00% across all the male subjects.

- In the treatment group, all 17 male subjects (100%) fulfilled the criteria for MDD before the intervention with an average of 6.65 of the diagnostic criteria being met. Post intervention, only 1 male subject (5.88%) continued to maintain a positive diagnosis with an average of 1.82 of the diagnostic criteria being met across the 17 male subjects. With 16 subjects changing their diagnostic status, this led to a remission rate of 94.12% for male subjects in the treatment group.

- Before the wait period, all 16 male subjects (100%) in the control group were diagnosed with MDD with an average of 6.35 of the diagnostic criteria being met. After the wait period, 13 male subjects (81.25%) in the control group continued to maintain their diagnosis with an average of 6.31 of the diagnostic criteria being met (across the 16 subjects in this group). Since three subjects were no longer diagnosed as having MDD, this led to a remission rate of 18.75% for male subjects in the control group over the wait period.

- Due to 3 male subjects no longer fulfilling the diagnostic criteria after the wait period, the control group began the intervention with 13 out the 16 male subjects as being diagnosed with MDD. After the intervention, only two (12.50%) subjects maintained their diagnosis of MDD with an average of 2.91 of the criteria being met across all 16 subjects. In turn, 11 subjects no longer satisfied the diagnosis. This led to a remission rate of

84.62% for male subjects in the control group when they received the intervention.

- Over the follow up period, 6.67% of male subjects fell out of remission. This means an extra two male subjects once again met the criteria of MDD. After the acupuncture intervention, three male subjects maintained their status of MDD. When these three subjects were combined with the two who regressed, a total of five subjects were diagnosed as depressed at the conclusion of the wait period. In percentage terms, 15.15% of the male subjects satisfied the diagnostic criteria for MDD at the conclusion of the follow up period.

4.10 Anxiety

The following tables report the statistical outcomes from the anxiety analysis. The subjects involved in each of the three phases of the study were administered the State-Trait Anxiety Index for Adults (STAI) before and after each phase of the study. A requirement of the STAI stipulates administration mid way through a treatment program. This requirement was maintained. In turn the analysis for the anxiety data was extended to include two extra comparisons, these being pre intervention scores to mid intervention scores and mid intervention scores to post intervention scores. The anxiety subscale for the SCL 90 has been included in this section since it also is a measure of anxiety.

The STAI consists of a global scale which is the adjunctive of its two subscales. The two subscales measure two different aspects of anxiety. These are "state" based anxiety, which measures the short term characteristics of anxiety, and "trait" based anxiety, which measures the long term characteristics of anxiety. The "state" based anxiety is measured in the Y1 scale and the "trait" based anxiety is measured in the Y2 scale.

In each table, the mean scores and their corresponding standard deviations from both before (PRE) and after (POST) the relevant intervention are reported alongside each measure. Please be aware these categorisations cannot be maintained in the comparisons which involve the mid assessments. In these situations the mid assessment assumes either the status of being either being "pre" or "post". This will be dependent on what is being compared. When the pre assessments are compared to the mid assessment, the mid assumes the "post" status. When the mid assessments are compared to the post assessment, the mid assumes the "pre" status. Also listed adjacent to each scale are the P scores. These give the statistical significance in the change in the corresponding mean scores for that measure during the respective phase of the study.

4.10.1 Main Analysis

4.10.2 Treatment Group

Table 32: Anxiety – Main Analysis – Treatment Gp – Intervention.

	n	M PRE	SD PRE	M POST	SD POST	
TEST						**P**
STAI	33	112.79	18.631	84.18	22.851	0.000
STAI PRE-MID	33	112.79	18.631	90.48	21.470	0.000
STAI MID-POST	33	90.48	21.470	84.18	22.851	0.024
Y1	33	52.76	10.689	38.27	12.122	0.000
Y1 PRE-MID	33	52.76	10.689	42.21	12.173	0.000
Y1 MID-POST	33	42.21	12.173	38.27	12.122	0.040
Y2	33	59.94	10.235	45.91	11.870	0.000
Y2 PRE-MID	33	59.94	10.235	48.27	10.771	0.000
Y2 MID-POST	33	48.27	10.771	45.91	11.870	0.032
SCL ANX	33	13.12	7.105	5.76	6.713	0.000

- Pre and post intervention comparisons for the intervention group showed statistically significant reduction in the degree of anxiety on the STAI and SCL 90 anxiety subscale. Statistically significant reductions in anxiety were also evident on the two subscales of the STAI. Both state (Y1) and trait (Y2) anxiety were significantly reduced.
- Pre to mid treatment and mid to post treatment comparisons on the STAI as well as the Y1 and Y2 subscales. All of these comparisons showed significant improvement for the intervention group.

4.10.3 Control Group

Table 33: Anxiety – Main Analysis – Control Gp – Wait Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
STAI	33	113.24	17.618	111.52	15.625	0.545
Y1	33	52.82	10.684	51.36	10.046	0.782
Y2	33	60.15	9.169	59.15	8.113	0.500
SCL ANX	33	14.36	9.413	12.94	8.653	0.081

- Subjects in the control experienced no significant change in their anxiety scores over the wait period. The STAI and its Y1 and Y2 subscales as well as the SCL anxiety subscale all indicated no significant change in anxiety.

Table 34: Anxiety – Main Analysis – Control Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
STAI	33	111.52	15.625	83.15	16.938	0.000
STAI PRE-MID	33	111.52	15.625	89.06	19.640	0.000
STAI MID-POST	33	89.06	19.640	83.15	16.938	0.018
Y1	33	52.36	10.046	37.82	9.492	0.000
Y1 PRE-MID	33	52.36	10.046	40.61	11.233	0.000
Y1 MID-POST	33	40.61	11.233	37.82	9.492	0.099
Y2	33	59.15	8.113	45.30	9.619	0.000
Y2 PRE-MID	33	59.15	8.113	48.45	10.226	0.000
Y2 MID-POST	33	48.45	10.226	45.30	9.619	0.005
SCL ANX	33	12.94	8.653	5.94	5.69	0.000

- Pre and post intervention scores for the control group indicate a statistically significant reduction in the degree of anxiety on the STAI and SCL anxiety subscale. This again was also the case for the STAI Y1 and Y2 subscales.
- Pre to mid treatment and mid to post treatment comparisons showed a significant reduction in anxiety on the STAI and the Y2 subscales
- The STAI Y1 subscale mid to post treatment comparison showed no significant change in anxiety for the control group, the pre to mid treatment comparison did however show a significant improvement.

4.10.4 Follow Up

As in the depression analysis, the reporting of the results from the follow up period differs in how it is presented since it involves two tables. The first table will report the results over the follow up period. This will involve comparisons between scores at the **end** of the intervention to those at the **end** of the follow up period. The second table will report the comparisons between scores **before** the intervention to those at the **end** of the follow up.

Table 35: Anxiety - Main Analysis – Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
STAI	66	83.67	19.964	85.12	24.150	0.536
Y1	66	38.05	10.805	40.33	12.813	0.104
Y2	66	47.11	10.912	44.20	11.835	0.006
SCL ANX	66	5.85	6.175	5.73	7.061	0.862

- Follow up comparisons for the STAI and the SCL 90 anxiety subscale showed no significant change in the levels of anxiety over the follow up. This was also the case for the STAI Y1 subscale.

- The comparison for the Y2 subscale did indicate subjects continued to show significant improvement in their trait based anxiety over the follow up period.

Table 36: Anxiety - Main Analysis – Pre Intervention to Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
STAI	66	112.15	17.073	85.12	24.150	0.000
Y1	66	52.56	10.294	40.33	12.813	0.000
Y2	66	59.55	9.172	44.20	11.835	0.000
SCL ANX	66	13.03	7.856	5.73	7.061	0.000

- The follow up to pre intervention comparison on all measures showed significant improvement in anxiety.

- These results indicate subjects either stayed unchanged or improved after the intervention (refer to table 32) and were also able to remain significantly less anxious than their pre intervention assessments.

4.11 Stratification Based on Sex

The following data shows the results stratified according to sex. The results for the female treatment and control groups and follow up analysis are reported first, followed by the same analysis for the males. The stratified results are reported using the same tables and structure as the main analysis. The same labeling and abbreviations also apply. Following the consideration of the male subjects is the analysis on sex interactions. This compares the relative improvement of females and males during the intervention phase and over the follow up period, on the STAI and it Y1 and Y2 subscales, as well as the SCL 90 subscale.

4.11.1 Female Subjects

4.11.2 Treatment Group

Table 37: Anxiety – Females – Treatment Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
STAI	16	119.06	17.023	94.00	21.364	0.001
STAI PRE-MID	16	119.06	17.023	93.94	20.521	0.001
STAI MID-POST	16	93.94	20.521	94.00	21.364	0.982
Y1	16	54.63	11.020	43.00	12.453	0.006
Y1 PRE-MID	16	54.63	11.020	43.13	11.272	0.006
Y1 MID-POST	16	43.13	11.272	43.00	12.453	0.949
Y2	16	64.44	6.909	51.00	10.973	0.000
Y2 PRE-MID	16	64.44	6.909	50.81	10.980	0.000
Y2 MID-POST	16	50.81	10.980	51.00	10.973	0.887
SCL ANX	16	14.94	6.894	8.50	7.394	0.007

- Pre and post intervention scores indicated female subjects in the treatment group showed significant improvement in the STAI and the Y1 as well as the SCL 90 anxiety subscale.

- Pre to mid treatment comparisons on the STAI and the Y1 and Y2 subscales showed significant improvement in anxiety scores for females.

- Mid to post treatment comparisons showed no significant change in anxiety scores on the STAI and the Y1 and Y2 subscales.

4.11.3 Control Group

Table 38: Anxiety – Females – Control Gp – Wait Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
STAI	17	113.06	19.107	111.35	14.164	0.706
Y1	17	52.29	10.947	52.06	9.277	0.925
Y2	17	60.76	9.660	59.29	6.771	0.508
SCL ANX	17	13.18	8.925	10.94	8.678	0.048

- The STAI and the Y1 and Y2 subscales indicated no significant change in the anxiety suffered by females in the control group over the wait period.

- The SCL anxiety subscale did show a significant improvement in the level of anxiety suffered by females over the wait period.

Table 39: Anxiety – Females – Control Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
STAI	17	111.35	14.164	79.59	16.789	0.000
STAI PRE-MID	17	111.35	14.164	87.18	22.965	0.000
STAI MID-POST	17	87.18	22.965	79.59	16.789	0.023
Y1	17	52.06	9.277	35.35	8.299	0.000
Y1 PRE-MID	17	52.06	9.277	39.00	13.811	0.000
Y1 MID-POST	17	39.00	13.811	35.35	8.299	0.139
Y2	17	59.29	6.771	44.24	10.562	0.000
Y2 PRE-MID	17	59.29	6.771	48.18	10.927	0.000
Y2 MID-POST	17	48.18	10.927	44.24	10.562	0.002
SCL ANX	17	10.94	8.678	3.47	3.608	0.002

- The female subjects in the control group showed significant improvement in their pre and post intervention assessments on the STAI and the Y1 and Y2 subscales as well as the SCL anxiety 90 subscale.
- The pre to mid intervention comparisons for females in the control group showed significant improvement in anxiety scores for the STAI and the Y1 and Y2 subscales.
- The mid and post intervention comparisons for females in the control group showed significant improvement in anxiety scores for the STAI and the Y2 subscale.
- The Y1 subscale showed no significant improvement in anxiety scores for females in the control group in the mid to post intervention comparison.

4.11.4 Follow Up

Table 40: Anxiety - Females – Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
STAI	33	86.58	20.208	92.24	23.652	0.096
Y1	33	39.06	11.054	43.94	12.542	0.021
Y2	33	49.15	10.776	47.21	12.532	0.194
SCL ANX	33	5.91	6.217	7.64	8.764	0.105

- The STAI and the Y2 subscale a well as the SCL 90 anxiety subscale showed no significant change in the levels of anxiety for female subjects over the follow up.
- The Y1 subscale did indicate a statistically significant increase in the level of state based anxiety as experienced by females over the eight week follow up period.

Table 41: Anxiety - Females – Pre Intervention to Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
STAI	33	115.09	15.857	92.24	23.652	0.000
Y1	33	53.30	10.082	43.94	12.542	0.000
Y2	33	61.79	7.219	47.21	12.532	0.000
SCL ANX	33	12.88	8.003	7.64	8.764	0.000

- All measures indicated a significant improvement in the anxiety experienced by female subjects after the follow up period compared to the pre intervention assessments.

- These results indicated female subjects either maintained or improved their anxiety scores after the intervention (refer to table 37) and in turn were able to remain significantly less anxious than their pre intervention assessments.

4.11.5 Male Subjects

4.11.6 Treatment Group

Table 42: Anxiety – Males – Treatment Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
STAI	17	106.88	18.604	74.94	20.720	0.000
STAI PRE-MID	17	106.88	18.604	87.24	22.449	0.006
STAI MID-POST	17	87.24	22.449	74.94	20.720	0.008
Y1	17	51.00	10.386	33.82	10.254	0.000
Y1 PRE-MID	17	51.00	10.386	41.35	13.252	0.021
Y1 MID-POST	17	41.35	13.252	33.82	10.254	0.018
Y2	17	55.71	11.201	41.12	10.891	0.000
Y2 PRE-MID	17	55.71	11.201	45.88	10.319	0.003
Y2 MID-POST	17	45.88	10.319	41.12	10.891	0.004
SCL ANX	17	11.41	7.072	3.18	4.927	0

- Male subjects in the treatment group showed significant improvement on pre and post intervention comparisons for the STAI, the Y1 and Y2 subscales as well as the SCL 90 anxiety subscale.
- Male subjects showed significant improvement in their anxiety scores for both pre to mid intervention and mid to post intervention comparisons according the STAI as well as the Y1 and Y2 subscales.

4.11.7 Control Group

Table 43: Anxiety – Males – Control Gp – Wait Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
STAI	16	113.44	16.512	111.69	17.515	0.632
Y1	16	53.38	10.726	52.69	11.104	0.760
Y2	16	59.50	8.884	59.00	9.564	0.808
SCL ANX	16	15.63	10.039	15.06	8.370	0.643

- Male subjects in the control group experienced no significant improvement in their anxiety scores over the wait period. The STAI, the Y1 and Y2 subscales as well as the SCL 90 anxiety subscale all indicated no significant change.

Table 44: Anxiety – Males – Control Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
STAI	16	111.69	17.515	86.94	16.783	0.000
STAI PRE-MID	16	111.69	17.515	91.06	15.876	0.000
STAI MID-POST	16	91.06	15.876	86.94	16.783	0.286
Y1	16	52.69	11.104	40.44	10.224	0.001
Y1 PRE-MID	16	52.69	11.104	42.31	7.726	0.001
Y1 MID-POST	16	42.31	7.726	40.44	10.224	0.438
Y2	16	59.00	9.564	46.44	8.702	0.000
Y2 PRE-MID	16	59.00	9.564	48.75	9.774	0.000
Y2 MID-POST	16	48.75	9.774	46.44	8.702	0.230
SCL ANX	16	15.06	8.370	8.56	6.398	0.002

- Male subjects in the control group experienced significant improvement in their post intervention scores compared to their pre intervention scores on the STAI and the Y1 and Y2subscales as well as the SCL 90 anxiety subscale.
- Males in the control group also showed significant improvement on pre to mid intervention comparisons on the STAI and the Y1 and Y2 subscales.
- The mid to post intervention assessments across all measures indicated males in the control group did not demonstrate a statistically significant change in their anxiety.

4.11.8 Follow Up

Table 45: Anxiety - Males – Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
STAI	33	80.76	19.590	78.00	22.816	0.393
Y1	33	37.03	10.623	36.73	12.220	0.870
Y2	33	45.06	10.822	41.18	10.427	0.021
SCL ANX	33	5.79	6.229	3.82	4.119	0.023

- Male subjects experienced no significant change in anxiety scores over the follow up period according to the STAI and the Y1 subscale.
- Males did continue to show significant improvement in their anxiety scores over the follow up period on the Y2 subscale as well as the SCL 90 anxiety subscale.

Table 46: Anxiety - Males – Pre Intervention to Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
STAI	33	109.21	17.967	78.00	22.816	0.000
Y1	33	51.82	10.605	36.73	12.220	0.000
Y2	33	57.30	10.412	41.18	10.427	0.000
SCL ANX	33	13.18	7.828	3.82	4.119	0.000

- Follow up to pre intervention comparisons for male subjects showed significant improvement in anxiety scores across all measures.
- These results indicated male subjects either maintained or improved their anxiety scores after the intervention (refer to table 42) and in turn were able to remain significantly less anxious than their pre intervention assessments.

4.12 Sex Interactions

4.12.1 Intervention Phase

Table 47: Anxiety - Sex Interactions - Intervention.

TEST	n	ANOVA
STAI	66	0.484
Y1	66	0.382
Y2	66	0.664
SCL ANX	66	0.271

- Over the intervention phase of the study, males and females demonstrated no significant difference in their rates of improvement. This was the case on pre and post intervention comparisons for the STAI and the Y1 and Y2 subscales, was well as the SCL 90 anxiety subscale.

4.12.2 Follow Up Phase

Table 48: Anxiety – Sex Interactions – Follow Up Period.

TEST	n	ANOVA PRE - FU	ANOVA POST - FU
STAI	66	0.152	0.071
Y1	66	0.066	0.061
Y2	66	0.376	0.155
SCL ANX	66	0.028	0.007

- Over the follow up period, there were no significant differences in the rate of improvement in anxiety scores between females and males. This was the case according to the STAI as well as the Y1 and Y2 subscales when comparing follow up scores to both the pre and post intervention scores.

- The SCL 90 anxiety subscale did indicate there was a difference between the sexes. It indicated male subjects improved at a greater rate than female subjects on both the follow up to post intervention as well as the follow up to pre intervention comparisons.

4.13 SCL – 90

The following tables report the statistical outcomes from the analysis on general mental health. The subjects involved in each of the three phases of the study were administered the Symptomatic Checklist 90 (SCL 90). The SCL 90 consists of a global scale which is an adjunctive of its nine subscales. The nine subscales are somatization, obsessive compulsive, interpersonal sensitivity, phobic anxiety, anxiety, depression, hostility, paranoid ideation and psychoticism. The depression and anxiety subscales have been addressed in the previous analysis. The remaining seven subscales will be addressed below.

In each table, the mean scores and their corresponding standard deviations from both before (PRE) and after (POST) the relevant intervention are reported alongside each measure. Also listed adjacent to each of the scales are the P scores (P). These give the statistical significance in the change in the corresponding mean scores for that measure during the respective phase of the study.

4.13.1 Main Analysis

4.13.2 Treatment Group

Table 49: SCL 90 – Main Analysis – Treatment Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	33	128.09	53.834	66.09	55.731	0.000
SOM	33	13.58	7.945	6.55	5.386	0.000
OC	33	18.03	7.764	8.88	7.088	0.000
IS	33	15.30	8.305	9.91	7.974	0.000
P.ANX	33	5.12	5.098	2.61	4.710	0.001
HOST	33	7.36	5.606	3.24	3.391	0.000
PID	33	8.00	4.677	5.58	5.172	0.010
PSYCH	33	9.03	6.322	5.27	6.181	0.000

- Subjects in the treatment group demonstrated significant improvement in scores from the intervention on the SCL 90 global scale as well as its somatization, obsessive compulsive, interpersonal sensitivity, phobic anxiety, hostility, paranoid ideation and the psychoticism subscales.

4.13.3 Control Group

Table 50: SCL 90 – Main Analysis – Control Gp – Wait Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	33	140.85	57.118	136.79	62.905	0.537
SOM	33	14.58	8.570	15.73	8.552	0.230
OC	33	20.42	6.694	19.91	8.210	0.627
IS	33	16.91	7.679	16.94	8.664	0.978
P.ANX	33	6.24	6.666	6.30	6.650	0.935
HOST	33	7.27	4.843	7.39	5.402	0.831
PID	33	8.42	5.624	9.27	6.135	0.159
PSYCH	33	11.03	6.232	10.52	6.699	0.451

- Over the wait period, subjects in the control group demonstrated no significant change in scores on the SCL 90 global scale and on all its subscales.

Table 51: SCL 90 – Main Analysis – Control Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	33	136.79	62.905	64.55	45.189	0.000
SOM	33	15.73	8.552	6.36	5.561	0.000
OC	33	19.91	8.210	10.52	7.129	0.000
IS	33	16.94	8.664	8.09	5.876	0.000
P.ANX	33	6.30	6.650	2.39	3.665	0.000
HOST	33	7.39	5.402	3.33	3.461	0.000
PID	33	9.27	6.135	4.85	4.790	0.000
PSYCH	33	10.52	6.699	5.27	5.252	0.000

- During the acupuncture intervention phase, subjects in the control group demonstrated significant improvement in scores on the SCL 90 global scale and on all its subscales.

4.13.4 Follow Up

As in the depression and anxiety analysis, the reporting of the results from the follow up period differs in how it is presented since it involves two tables. The first table will report the results over the follow up period. This will involve comparisons between scores at the **end** of the intervention to those at the **end** of the follow up period. The second table will report the comparisons between scores **before** the intervention to those at the **end** of the follow up.

Table 52: SCL 90 – Main Analysis – Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	66	65.32	50.349	64.24	57.459	0.820
SOM	66	6.45	5.433	6.65	6.331	0.749
OC	66	9.70	7.101	9.35	7.641	0.591
IS	66	9.00	7.010	8.45	8.302	0.472
P.ANX	66	2.50	4.189	2.58	4.727	0.817
HOST	66	3.29	3.400	3.39	3.965	0.786
PID	66	5.21	4.960	4.55	4.950	0.150
PSYCH	66	5.27	5.691	5.15	6.087	0.835

- At the conclusion of the eight-week follow up period, subjects experienced no significant change in scores in the SCL 90 global scale as well as all its subscales.

Table 53: SCL 90 – Main Analysis – Pre Intervention to Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	66	132.44	58.259	64.24	57.459	0.000
SOM	66	14.65	8.262	6.65	6.331	0.000
OC	66	18.97	7.985	9.35	7.641	0.000
IS	66	16.12	8.461	8.45	8.302	0.000
P.ANX	66	5.71	5.909	2.58	4.727	0.000
HOST	66	7.38	5.463	3.39	3.965	0.000
PID	66	8.64	5.451	4.55	4.950	0.000
PSYCH	66	9.77	6.506	5.15	6.087	0.000

- Compared to the pre intervention scores, subjects showed statistically significant improvement after the follow up period in SCL 90 global scale as well as on all its subscales.
- This means subjects showed no significant decline in scores across all the SCL 90 scales during the follow up period (refer to table 48) and were able to maintain significantly better scores than their pre intervention scores.

4.14 Stratification Based on Sex

Following is the analysis of the results for general mental health which take into consideration the stratification of subjects based on sex. The results for females in the treatment and control groups as well as the follow up analysis are reported first, followed by the same analysis for the males. The stratified results are reported using the same tables and structure as in the main analysis. The same labeling and abbreviations also apply. Following the consideration of the male subjects is the analysis on sex interactions. This compares the relative improvement in scores between females and males during the intervention phase as well as over the follow up period on SCL 90 scales.

4.14.1 Females Subjects

4.14.2 Treatment Group

Table 54: SCL 90 – Females – Treatment Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	16	150.94	53.686	92.50	58.518	0.002
SOM	16	15.56	8.278	8.38	5.726	0.005
OC	16	20.75	8.355	12.38	7.473	0.004
IS	16	19.50	7.832	14.31	8.014	0.020
P.ANX	16	6.25	5.520	4.13	5.807	0.081
HOST	16	9.50	5.854	4.56	4.320	0.003
PID	16	9.94	5.196	7.94	5.131	0.206
PSYCH	16	10.69	5.907	7.38	5.943	0.018

- Female subjects in the treatment group demonstrated significant improvement in scores due to the intervention on the SCL 90 global scale as well as the somatization, obsessive compulsive, interpersonal sensitivity, hostility and the psychoticism subscales.
- Female subjects in the treatment group did not significantly improve on their phobic anxiety and paranoid ideation scores over the intervention period.

4.14.3 Control Group

Table 55: SCL 90 – Females – Control Gp – Wait Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	17	143.71	59.228	133.76	69.645	0.251
SOM	17	15.71	9.150	15.71	7.928	1.000
OC	17	22.00	7.331	21.29	9.143	0.613
IS	17	17.24	8.437	17.00	9.434	0.898
P.ANX	17	5.65	7.441	5.18	7.577	0.674
HOST	17	7.00	4.228	7.47	5.949	0.533
PID	17	7.71	5.665	8.35	6.133	0.399
PSYCH	17	11.53	6.597	10.29	7.523	0.206

- Female subjects in the control group demonstrated no significant change over the wait period on the SCL 90 global scale as well as all of its subscales.

Table 56: SCL 90 – Females – Control Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	17	133.76	69.645	53.18	43.189	0.000
SOM	17	15.71	7.928	6.82	6.095	0.000
OC	17	21.29	9.143	9.82	7.384	0.000
IS	17	17.00	9.434	7.00	6.461	0.000
P.ANX	17	5.18	7.577	1.88	3.998	0.026
HOST	17	7.47	5.949	2.41	2.320	0.000
PID	17	8.35	6.133	3.65	4.443	0.000
PSYCH	17	10.29	7.523	3.47	4.474	0.001

- Female subjects in the control group demonstrated significant improvement due to the intervention on the SCL 90 global scale as well as all its subscales.

4.14.4 Follow Up

Table 57: SCL 90 – Females – Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	33	72.24	54.186	81.76	65.567	0.205
SOM	33	7.58	5.879	8.67	7.427	0.248
OC	33	11.06	7.424	12.24	8.277	0.226
IS	33	10.55	8.047	11.24	9.004	0.565
P.ANX	33	2.97	5.009	3.55	5.772	0.222
HOST	33	3.45	3.554	4.55	4.671	0.041
PID	33	5.73	5.192	5.36	5.396	0.646
PSYCH	33	5.36	5.522	6.36	6.749	0.271

- Over the follow up period females showed no significant change in their SCL 90 global scores as well as somatization, obsessive compulsive, interpersonal sensitivity, phobic anxiety, paranoid ideation and the psychoticism subscales.
- Females did however show a significant increase in their levels of hostility over the eight week follow up period.

Table 58: SCL 90 – Females – Pre Intervention to Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	33	142.09	62.066	81.76	65.567	0.000
SOM	33	15.64	7.972	8.67	7.427	0.000
OC	33	21.03	8.637	12.24	8.277	0.000
IS	33	18.21	8.652	11.24	9.004	0.000
P.ANX	33	5.70	6.579	3.55	5.772	0.005
HOST	33	8.45	5.901	4.55	4.671	0.000
PID	33	9.12	5.667	5.36	5.396	0.000
PSYCH	33	10.48	6.685	6.36	6.749	0.000

- After the follow up period, females showed a significant improvement in their scores on SCL 90 global scale as well as all its subscales compared to the pre intervention scores.
- This means female subjects showed no significant decline in scores across all the SCL 90 scales during the follow up period (refer to table 53) and were able to maintain significantly better scores than their pre intervention scores.

4.14.5 Male Subjects

4.14.6 Treatment Group

Table 59: SCL 90 – Males – Treatment Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	17	106.59	45.618	41.24	40.592	0.000
SOM	17	11.71	7.372	4.82	4.558	0.000
OC	17	15.47	6.385	5.59	4.938	0.000
IS	17	11.35	6.800	5.76	5.414	0.000
P.ANX	17	4.06	4.575	1.18	2.877	0.005
HOST	17	5.35	4.676	2.00	1.458	0.004
PID	17	6.18	3.340	3.35	4.242	0.013
PSYCH	17	7.47	6.472	3.29	5.892	0.000

- Male subjects in the treatment group demonstrated significant improvement due to the intervention on the SCL 90 global scale as well as the somatization, obsessive compulsive, interpersonal sensitivity, phobic anxiety, hostility, paranoid ideation and the psychoticism subscales.

4.14.7 Control Group

Table 60: SCL 90 – Males – Control Gp – Wait Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	16	137.81	56.561	140.00	56.978	0.831
SOM	16	13.38	8.024	15.75	9.434	0.169
OC	16	18.75	5.698	18.44	7.080	0.852
IS	16	16.56	7.042	16.88	8.074	0.788
P.ANX	16	6.88	5.909	7.50	5.489	0.538
HOST	16	7.56	5.549	7.31	4.949	0.777
PID	16	9.19	5.659	10.25	6.181	0.277
PSYCH	16	10.50	5.989	10.75	5.939	0.799

- Male subjects in the control group demonstrated no significant change over the wait period on the SCL 90 global scale as well as its subscales.

Table 61: SCL 90 – Males – Control Gp – Intervention.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	16	140.00	56.978	76.63	45.437	0.000
SOM	16	15.75	9.434	5.88	5.084	0.001
OC	16	18.44	7.080	11.25	7.010	0.000
IS	16	16.88	8.074	9.25	5.132	0.001
P.ANX	16	7.50	5.489	2.94	3.316	0.001
HOST	16	7.31	4.949	4.31	4.222	0.007
PID	16	10.25	6.181	6.13	4.951	0.000
PSYCH	16	10.75	5.939	7.19	5.468	0.003

- Male subjects in the treatment group demonstrated significant improvement due to the intervention on the SCL 90 global scale as well as all of its subscales.

4.14.8 Follow Up

Table 62: SCL 90 – Males – Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	33	58.39	45.980	46.73	42.126	0.036
SOM	33	5.33	4.774	4.64	4.227	0.383
OC	33	8.33	6.594	6.45	5.729	0.024
IS	33	7.45	5.489	5.67	6.551	0.051
P.ANX	33	2.03	3.177	1.61	3.181	0.351
HOST	33	3.12	3.286	2.24	2.716	0.115
PID	33	4.70	4.740	3.73	4.389	0.053
PSYCH	33	5.18	5.940	3.94	5.166	0.084

- Over the follow up period, male subjects experienced no significant change in scores on the SCL subscales of somatization, interpersonal sensitivity, phobic anxiety, hostility, paranoid ideation and the psychoticism subscales.
- Males did however continue to significantly reduce their SCL 90 global scores and the obsessive compulsive subscale scores over the follow up period.

Table 63: SCL 90 – Males – Pre Intervention to Follow Up Period.

TEST	n	M PRE	SD PRE	M POST	SD POST	P
GLOBAL	33	122.79	53.384	46.73	42.126	0.000
SOM	33	13.67	8.550	4.64	4.227	0.000
OC	33	16.91	6.793	6.45	5.729	0.000
IS	33	14.03	7.844	5.67	6.551	0.000
P.ANX	33	5.73	5.258	1.61	3.181	0.000
HOST	33	6.30	4.838	2.24	2.716	0.000
PID	33	8.15	5.269	3.73	4.389	0.000
PSYCH	33	9.06	6.344	3.94	5.166	0.000

- Comparisons between follow up scores and pre intervention scores indicated males showed significant improvement on the SCL 90 global scale as well all its subscales.
- This means male subjects showed no significant decline in scores across all the SCL 90 scales during the follow up period (refer to table 58) and were able to maintain significantly better scores than their pre intervention scores.

4.15 Sex Interactions

4.15.1 Intervention Phase

Table 64: SCL 90 - Sex Interactions - Intervention.

TEST	n	ANOVA
GLOBAL	66	0.391
SOM	66	0.525
OC	66	0.581
IS	66	0.765
P.ANX	66	0.372
HOST	66	0.263
PID	66	0.568
PSYCH	66	0.238

- Females and males showed no significant difference in their rates of improvement on the SCL 90 global scale as well as on all its subscales. This occurred despite females in the treatment group (refer to table 49) not experiencing a significant change on phobic anxiety and paranoid ideation when males did show significant improvement in both these subscales (refer to table 55 and 57).

4.15.2 Follow Up Phase

Table 65: SCL 90 – Sex Interactions – Follow Up Period.

TEST	n	ANOVA PRE - FU	ANOVA POST - FU
GLOBAL	66	0.194	0.023
SOM	66	0.241	0.147
OC	66	0.381	0.017
IS	66	0.470	0.100
P.ANX	66	0.049	0.125
HOST	66	0.889	0.010
PID	66	0.548	0.513
PSYCH	66	0.407	0.055

- The pre intervention to follow up score comparison on the phobic anxiety subscale showed males improved at a significantly greater rate than females (refer to tables 55 and 59).

- Post intervention to follow comparisons showed scores for males improved at a significantly greater rate than females on the SCL 90 global scale and the obsessive compulsive subscale (refer to tables 54 and 58).

- Since female hostility scores increased over the follow up period (refer to table 53) and the male scores remained unchanged, the males were able to improve at a statistically greater rate.

- There were no sex differences in scores between post intervention to follow up comparisons on the somatization, interpersonal sensitivity, phobic anxiety, paranoid ideation and the psychoticism subscales.

- There were no sex differences on pre intervention scores to follow up score comparisons on the SCL 90 global scale as well as the somatization, obsessive compulsive, interpersonal sensitivity, hostility, paranoid ideation and the psychoticism subscales.

Chapter 5: Discussion

The main aim of this is study was to test the effectiveness of acupuncture as an adjunct to SSRI's to treat depression, so the implications of the depression results will carry the greatest weight in this chapter. To create a context, first a summary of the main depression findings from this study will be given. The depression outcomes will be subsequently discussed in greater detail relative to the three depression scales (i.e. the BDI, the Ham-D and the SCL 90 depression subscale). Subsequent to this, the depression results will also be discussed within the context of the key elements of design which were used in this study. This will involve separate sections discussing the results relevant to the use of acupuncture as an adjunct treatment to antidepressant medication and the use of a wait list group as a means of control. To clarify and better understand the long-term benefits of using acupuncture to treat depression, the results over the follow up period will then be addressed.

This study also examined sex differences. The literature demonstrated a lack of reporting analysis based on sex. The depression outcomes from this study for both females and males will be concurrently discussed in reference to each of the respective phases in this study (i.e. the intervention, wait and follow up phases). This will be done for both sexes in the treatment group and the control group. Further to this, will be the discussion of the statistical analysis which considered the interaction in results between the sexes.

The intention to treat analysis using the depression scales closely replicated the findings in the standard analysis. This section of the chapter will therefore only highlight the main trends and identify the few differences to the standard analysis. The analysis considering the effectiveness of the treatment in respect to changes in the rate of remission will be discussed at the end of the section on depression. This analysis involves the outcomes relating to the diagnosis of depression as outlined in the structured clinical interview for MDD. This was

done by considering the changes in the diagnostic status of the subjects in reference to the respective phases of the study.

To conclude the discussion on the statistical outcomes, the results for anxiety and general mental health will then be discussed in turn. This will involve separate consideration of the results from the STAI and then the SCL 90. The limitations of this study will be addressed and so too will the implications of the study in reference to the relevance to liver qi stagnation.

5.1 Depression

The general trends in the results from this study indicated acupuncture is an effective adjunct therapy to SSRI's in the treatment of depression in those who have liver qi stagnation. Across all depression scales (i.e. the BDI, the Ham-D and the SCL 90 depression subscale) subjects demonstrated significant reductions in depression scores over the acupuncture intervention phases. This trend was evident in the treatment group and when those in the control group received the acupuncture intervention at the conclusion of the wait period. All three depression scales yielded the same statistical results (α = 0.05, p = 0.000). Encouragingly, this was also the case for the intention to treat analysis.

More interesting were the results over the wait period. There was a large contrast between the changes in scores over the intervention phase of the study compared to those over the wait period. No significant change in the scores from any of the depression scales occurred in the wait period. The respective P scores across the measures indicated varying levels of statistical insignificance. This was also the case in the intention to treat analysis. This suggests that there was no improvement based on time alone and hence, the improvement in scores during the intervention phases of the study may be attributable to the acupuncture treatment.

5.2   Beck Depression Inventory

5.2.1   P Scores

The main empirical focus of this study was on the BDI. The results from the BDI analysis indicate acupuncture is an effective adjunct therapy to SSRI's in the treatment of MDD for those with liver qi stagnation. The results from the main analysis indicated large statistically significant reductions in BDI scores over the acupuncture intervention phase. This was true for those in treatment (α = 0.05, p = 0.000) as well as for those in the control group (α = 0.05, p = 0.000) when they received the acupuncture after the wait period.

This statistically large effect is consistent with the studies in the literature. Three studies in the literature used the BDI to independently assess the effectiveness of using acupuncture to treat depression. Blitzer (2004) treated the extra ordinary vessels and found a statistically significant improvement in BDI scores, so too MacPherson (2004) though the effect size was not as large. These results were also consistent with the Quah-Smith (2005) laser acupuncture clinical trial.

Two other studies also used the BDI as the primary measure but their results cannot be compared. Neither Manber (2004) nor Allen (2006) carried out independent analysis on an acupuncture treatment group. Instead, they compared their intervention data to their control groups which implemented non–specific acupuncture. The BDI results between these two studies were also inconsistent. Whilst Allen (1998) found active acupuncture out performed the non–specific acupuncture, Manber (2004) in contrast found no significant difference between the two groups over the first month and second month of treatment. Manber (2004) did however find that acupuncture was able to reduce depression scores significantly more than an additional control group which used massage therapy. The difference in the outcomes between these two

studies may not be necessarily attributable to the effectiveness of the acupuncture intervention. Instead, the variance in the outcomes may be a reflection on the use of the non-specific acupuncture. As previously discussed, the use of non-specificity acupuncture is problematic. It is difficult to claim that any non-specific treatment is outside the domain of acupoint specificity and in turn elicits no therapeutic benefit. In the case of Manber (2004), this concern may be valid. It could be argued the use of massage as a method of control provided a greater contrast to the active acupuncture intervention. Hence, it may have been a more neutral and more suitable control.

5.2.2 Average Scores

Given the large statistical significance in BDI scores for the treatment group's acupuncture intervention it stands to reason this would be characterised by considerable improvements in BDI scores. BDI scores however, can also be interpreted depending on classifications of being either "minimal", "mild", "moderate" or severe" depending on predetermined parameters. These classifications are designed to interpret the severity of the depression (as reported by the subjects). The BDI scores in this study averaged 27.39 before the acupuncture intervention and were reduced to 12.03 after the intervention. This is an average reduction of approximately 56%. According to the BDI classification on severity, on average, the subjects in the treatment group entered the study classified as being "moderately" depressed. Beck's scale range for moderate depression is between 20-28 points. It is also interesting to note that an average score of 27.39 is on the high side of the "moderate" classification and very close to the classification of "severe". In contrast, the post intervention average scores for the treatment group met the classification of "minimal" depression. Beck deems BDI scores between 0-13 to be representative of minimal depression. This means the average scores were reduced from a "moderate" classification down to a "minimal" classification, bypassing the classification of "mild". A similar pattern of average score changes

was observed when the control group received the acupuncture intervention. Their score went from 28.76 ("severe") down to 12.53 ("minimal").

Three out of the five studies in the literature, which used the BDI, reported average scores. The study which had the results most similar to this study was that of Quah-Smith (2005). In the Quah-Smith (2005) study, subjects improved their scores by an average of 16.1 points. Pre intervention scores averaged 22.8 ("moderate") and were reduced to 6.7 ("minimal"). This is similar to the average reductions in BDI scores which were found in this study. Those in the treatment group reduced their scores by an average of 15.36 points and those in the control group reduced their scores by 16.23 when they received the acupuncture intervention.

Two other studies in the literature which reported average BDI scores were Allen (1998) and MacPherson (2004). Allen did not report average scores before and after the intervention, but did report an average reduction in BDI scores over the acupuncture intervention of 10.7 points. This is comparatively less than what was experienced in this study. Average DBI scores in MacPherson (2004) were reduced from 22.2 to 16.0. What is interesting to note here is that subjects were not able to improve the severity of their depression beyond the "moderate" level they had reported before the treatment. This again is considerably less of a change than experienced in this study.

It is difficult to account for the similarities and variances in average BDI scores found between this study and those in the literature. The most salient difference is how this study better maintained the integrity of the CM model. As discussed previously, this study focused on subjects with liver qi stagnation and administered a principle of treatment in accordance with that CM diagnosis. Since Allen (1998) did not report any steps to follow the CM paradigm, this may account for the reduced effectiveness of the acupuncture used in that study. MacPherson (2004) did make an attempt to identify the patterns of disharmony

within the sample of subjects. He did not however fully consider the CM model and included multiple unrelated patterns of disharmony as a homogenous group within his statistical analysis. No consideration was given to particular diagnostic patterns that could best account for depression or if it could be more effectively treated. Nor did he find consensus on suitable acupoint prescriptions for each pattern of disharmony. Any or a combination of these shortcomings may account for the difference in outcomes between this study and MacPherson (2004).

Alternatively, the differences in the scores may be related to the extent to which acupuncture can reduce BDI scores. This study, along with MacPherson (2004), were both not able to reduce scores lower than the 15-16 point mark. This could imply that acupuncture may only be able to reduce BDI scores up to a certain point. The pre intervention averages were higher in this study than they were in MacPherson (2004). This difference may explain why MacPherson (2004) was not able to have as large a reduction on average scores. Although this may appear to be plausible, the results for Quah-Smith (2005) do not support this. Conversely, it is interesting to note, Quah-Smith (2005) used laser acupuncture. This could be the reason why average scores were able to be so low after the intervention. This may be an interesting consideration for another study in which laser acupuncture can be tested as an adjunct treatment to SSRI's to reduce depression in sufferers who present with liver qi stagnation.

5.2.3 Wait List Comparison

In order to determine the effectiveness of the acupuncture intervention, a comparison was made to the scores from the control group over the wait period. The length of time for the wait period was over eight weeks and equal to that of the acupuncture intervention. The main analysis saw little change in BDI scores over the wait period and no statistical significance was observed (α = 0.05, p = 0.371). Average scores were slightly reduced from 28.30 to 27.15. This means the severity of the depression, as experienced by the subject over the wait

period, remained in the classification of "moderate" with no improvement in BDI scores while taking SSRI medication alone. Since the only difference between the treatment group and the control group over the wait period was that the treatment group received the acupuncture intervention, it may be concluded the change in scores for the treatment group was attributable to the acupuncture intervention.

5.2.4 Intervention comparison

When the subjects in the control group received the acupuncture intervention the change in their BDI scores was comparable to those experienced by the treatment group. This change yielded the same large statistical significance (α = 0.05, p = 0.000). The average scores were reduced from 27.15 ("moderate") to 12.12 ("minimal"). This is similar to the average change in scores experienced by those in the treatment group. The difference in the average change in scores in both the intervention phases in this study compared to those over the wait period, highlight the improvement in depression severity when the acupuncture was administered in contrast to when it was not administered.

5.3 Hamilton Rating Scale for Depression

5.3.1 P Scores

In addition to the BDI, this study administered the Ham-D as part of the assessment regimen. The use of both depression measures offered a more detailed understanding of the effectiveness of using acupuncture to treat depression. It would allow for the consideration of any differences in sensitivity to varying outcome instruments and increase opportunities for comparison with the majority of the studies in the literature. Like the BDI, the statistical analysis using the Ham-D suggested the acupuncture intervention was significantly effective in reducing the severity of depression. The outcomes from the main

analysis using the Ham-D indicated that acupuncture had a large statistical effect over the intervention phases of the study (α = 0.05, p = 0.000).

This statistically significant effect is consistent with the studies in the literature. Despite that the Ham-D was used by the majority of the studies in the literature, only three studies used it for a within group analysis. Luo (1985), Yang (1994) and Han (2004) all experienced changes in Ham-D scores with large statistical significance. Although each of these studies used differing approaches in how they implemented the acupuncture, these studies serve to provide a consensus which suggests acupuncture is effective in treating depression.

Other studies in the literature which used the Ham-D did not independently assess the effectiveness of using acupuncture to treat depression using within group analysis. Instead, they used between group comparisons where the effectiveness of acupuncture was assessed relative to a range of controls. As explained in the literature review, there are difficulties with finding an appropriate method of control for acupuncture studies. This is problematic since between group analysis may not offer a true indication of the effectiveness of an acupuncture intervention. This may have consequently led to these studies to report conflicting results. Nevertheless, since these studies did not report any statistical outcomes which have measured the effectiveness of the acupuncture treatment alone, they have little comparative relevance to the statistical outcomes of this study.

5.3.2 Wait List

As stated above, the outcomes over the acupuncture intervention indicated statistically significant reductions in Ham-D scores. In contrast, the main analysis over the wait period, for those in the control group, showed no change. No statistical significance was found in Ham-D scores over the wait period (α = 0.05, p = 0.3.13). The only difference between the treatment group and the

control group, over the wait period, was that those in the treatment group received the acupuncture intervention. As the case for the BDI, the results according to the Ham-D may indicate that acupuncture is an effective adjunct therapy to SSRI in the treatment of depression for those with liver qi stagnation.

5.3.3 Intervention comparison

At the conclusion of the wait period, the subjects in the control group also received the acupuncture intervention. Analysis using the Ham-D suggested the effect of the intervention was the same as experienced by those in the treatment group ($α = 0.05$, $p = 0.000$). This means the subjects in both groups experience large statically significant improvements. These outcomes alleviate the potential problem of the significant difference which was found in the base statistics between those in the treatment and control groups at entry to the study. More importantly, these Ham-D outcomes from both the treatment group and the control group highlight the changes in depression scores when the acupuncture was administered, compared to when it was not. Like the BDI, no change was observed when the subjects in the control group received only the SSRI medication during the wait period. Then when the acupuncture was later administered in conjunction with the SSRI's, significant reductions in the severity of depression were observed.

5.3.4 Averages

Unlike the BDI scores, the average score of the Ham-D have no diagnostic significance. They are however of particular interest in this study. A limitation of this study is the inexperience of the assessor in administering the Ham-D. Despite being in a standardised structured interview format, the Ham-D requires a sophisticated user. By reviewing the average scores, before and after the acupuncture intervention, it will be evident if the lack of experience in administering the Ham-D was a limiting factor in this study.

The percentage change in depression score was very similar on the Ham-D compared to the BDI. According to the BDI, subjects reduced the average severity of their depression by 56.97% in the treatment group and by 55.35% in the control group when they received the intervention. The average percentage improvement for the Ham-D scores was 52.02% for the treatment group and by 50.12% for the control group when they received the intervention. Given both measures yielded the similar improvement in scores, this may imply the Ham-D was administered in a manner which correctly reflected the severity of depression as it was reported by the subjects. This can be further evidenced by what was reported in the literature.

The studies of Luo (1985), Luo (1990), Luo (1998) and Yang (2004) reported their average Ham-D scores. These averages however have limited relevance for comparison to this study. Each of these studies used hospital inpatients. This population is vastly different to that in this study. The studies in the literature which did report average Ham-D scores and are however more comparable to this study are Allen (1998), Allen (2000) and Manber (2004). All three of these studies used assessors who were expert in administering the Ham-D structured interview. The Allen (1998) and the Allen (2000) studies both found the same average improvement in Ham-D scores over the acupuncture intervention. The average improvement was 11.7 points. These outcomes are not dissimilar to the average change in scores experienced by Manber (2004).

The average improvement in this study was almost the same as those experienced by Allen (1998) and Allen (2000). The average change in Ham-D scores in the main analysis for those in the treatment group was an improvement of 11.68 points. When Ham-D scores over the intervention phase for both the treatment group and the control group were combined, the average of the pre intervention score was 21.0 points. Once the acupuncture treatment had been received the average Ham-D score was reduced to 9.4. This is an average reduction of 11.4 points. This again is similar to the studies in the

literature. Although it remains unknown where within the spectrum of possible Ham-D scores the Allen (1998) and Allen (2000) average scores lie, these comparisons offer encouraging support for the accuracy of the Ham-D assessments in this study.

5.4   SCL 90 Depression Subscale

As discussed in the method chapter, the SCL 90 was included in this study to serve as an alternate and secondary measure. Amongst many other subscales, the SCL 90 has a depression subscale. Like the BDI and the Ham-D, the analysis of the SCL 90 depression subscale indicated statistically significant reductions in depression scores for the treatment group over the acupuncture intervention phase of the study ($\alpha = 0.05$, $p = 0.000$). In contrast, the SCL 90 depression subscale indicated there was no significant change on scores over the wait period ($\alpha = 0.05$, $p = 0.210$). In turn, when those in the control group received the acupuncture intervention, they too showed statistically significant reductions in their SCL 90 depression subscale scores ($\alpha = 0.05$, $p = 0.000$). These outcomes are consistent with the BDI and the Ham-D. Collectively they suggest the improvement in the mood of the subjects may be due to the acupuncture intervention.

The SCL 90 depression subscale offers a slightly different insight into how levels of depression are reported. The SCL 90 measures how subjects were feeling at the immediate time of the assessment. In contrast, the BDI and the Ham-D both consider how the subjects felt over a period of time (albeit a short term horizon). This is interesting as the SCL 90 allows researchers to get an indication as to how the subjects felt at a precise moment compared to short term trends in mood. The results of the SCL 90 depression subscale analysis suggest the acupuncture intervention was able to influence depression scores with a synergy between both short term trends (in the BDI and the Ham-D) and the immediate perception of a depressed mood (in the SCL 90). This insight is even more

interesting when it is considered relative to the long term nature of the depression suffered by the subjects in this study. It would otherwise be expected that the improvement of long term chronic depression would occur with greater variance across time horizons than what has been suggested by the results in this study.

There were no studies in the literature which have tested the use of using acupuncture in the treatment of depression and have used the SCL 90 depression subscale in their analysis. No comparison can therefore be made directly to any particular study. Though within this study, it is of interest to note the close similarity in the trends on the SCL 90 depression subscale compared to the more recognised BDI and Ham-D scales.

5.5   Acupuncture as an Adjunct to Medication

The majority of the studies in the literature administered acupuncture as a stand-alone treatment for depression. There were however three studies which were designed to test the use of acupuncture as an adjunct treatment to antidepressant medication. These were the studies of Luo (1990), Luo (1998) and Roschke (2000). Roschke (2000) and phase 1 of Luo (1998) used the same design as this study. In the studies of Luo (1990) as well as phase 2 of Luo (1998) acupuncture was administered as an adjunct to a placebo medication. The outcomes in both these instances have limited comparability to this study.

Before the results of this study are compared to those in the literature, it must be stressed that the previous studies did not use the same antidepressant medication. Luo (1998) used amitriptyline and Roschke (2000) used mianserin. Both of these antidepressant medications are based on different pharmacological actions compared to SSRI's. Acupuncture when combined with an SSRI needs to be acknowledged as a different treatment compared to when

it is administered in conjunction with either of the drugs from these studies. Any comparison with the literature serves only to provide a general insight into the effectiveness of acupuncture when used in conjunction with an antidepressant.

Phase 1 of Luo (1998) was the only analysis in the literature which undertook a within groups analysis to test the effectiveness of using acupuncture in conjunction with a medication. The statistical analysis in the remaining literature was mostly focused on relative comparisons to control groups. Nevertheless, Ham-D analysis in phase 1 of Luo (1998) indicated that when acupuncture is combined with the use of the antidepressant medication, there was a significant reduction in depression scores. This outcome is consistent with the results from the Ham-D analysis in this study.

Roschke (2000) used a method most similar to this study. In both studies, acupuncture was used as an adjunct to medication. He however only carried out statistical analysis in reference to the control groups (non-specific acupuncture and a wait list). No within group analysis was carried out which solely focused on measuring the effectiveness of using acupuncture combined with the medication. The outcomes from his results mostly indicated insignificant differences between the groups.

In addition, Roschke (2000) took a unique approach in his choice of measures. Despite his use of multiple measures, none of the depression measures were used in this study (nor anywhere else in the literature). This adds to the difficulty in making comparisons between the two studies. There are interesting points of difference in his choice of depression scales. In contrast to providing a broad understanding of depression, each of the scales he chose focused on specific aspects of depression. For example, the GAS focused specifically on functioning and the BRMS focused specifically on melancholia. Each of these scales does offer insight into an aspect of a depressive mood, but they do not fully reflect on

MDD. In contrast, researchers in this study included scales which were designed to measure a broader understanding of depression.

The differences in the choice of scales may account for the variance in outcomes between Roschke's study and this study. It is however interesting to note, the scale he used, which is most similar to those used in this study, was the CGI Item 2. This scale measured global improvement and in doing so considered a broader understanding of the symptoms associated with depression. Consistent with the results of this study, this CGI Item 2 did show a statistically significant difference between the acupuncture combined with the medication compared to the control groups. It is possible that using broad measures of depression can lead to a consistency between outcomes across studies. This highlights the need to find a consensus as to which are the most suitable and valid measures for a clinical trial which uses acupuncture to treat depression. This may potentially be an area for further research.

Alternatively, the outcomes in his study may potentially be attributable to how acupuncture was applied in accordance with the CM model. As per previous discussion, the way in which Roschke applied the use of acupuncture was problematic. Little attempt was made to maintain the integrity of the CM model. Within his approach there was no inclusion of a CM diagnosis according to patterns of disharmony. Hence, acupoint selection was not based on a suitable principle of treatment. It may then be fair to suggest his acupuncture treatments may not have been indicative of the potential effectiveness of acupuncture as it is used in clinical practice. This may be evidenced by the insignificant differences between the active acupuncture and the non-specific acupuncture across all of the measures used in the study.

The design of this study, along with those of Luo (1998) and Roschke (2000), not only used acupuncture as an adjunct treatment to antidepressant medication, they also made comparisons to a wait list control group. Both

elements of design complement each other and they are commonly used together. Further analysis comparing this study to those in the literature which used acupuncture as an adjunct treatment will be addressed within the discussion concerning the use of wait list control groups.

5.6   Wait List

Over the wait period, subjects in this study showed no significant change in their depression scores. This was evident across all three depression measures; BDI (α = 0.05, p = 0.371); Ham-D (α = 0.05, p = 0.313); and the SCL 90 depression subscale (α = 0.05, p = 0.210). This indicated the subjects in the control group experienced no significant change in their levels of depression when they only received the SSRI antidepressants. In contrast, when the acupuncture intervention was used in conjunction with the SSRI's, in both the treatment group and the control group, significant score reductions occurred across all the depression measures (α = 0.05, p = 0.000). This highlights the difference in effectiveness when acupuncture is administered with the SSRI's.

The outcomes in this study were consistent with the outcomes for the wait list control group in Allen (2006) but not so for Allen (1998) and Allen (2000). More specifically relevant to this study are those studies in the literature which combined the design of a wait list along with the use of acupuncture as an adjunct treatment to medication. Roschke (2000) carried out his statistical analysis so that the use of acupuncture as an adjunct treatment was considered separate to analysis involving the wait group. The results will therefore differ between each consideration. The outcomes involving the wait group indicated the results from two of the scales were consistent with the outcomes in this study. The two other measures in Roschke (2000) as well as the result from Luo (1998) in contrast, indicated acupuncture combined with an antidepressant was not able to significantly improve scores more than those who just received the medication.

Of particular interest are the Allen (1998) and Allen (2000) studies. Neither of these studies offered any statistical analysis specifically relating to the wait list control group. They did however report on the change in the average scores using the Ham-D. First, it must be acknowledged that both of these studies reported the same statistical outcomes and average scores throughout their analysis. This suggests both studies were based on the same Ham-D data. Nevertheless, the change in average scores over the wait period was reported to be an improvement of 6.1 points. This is much greater than the average improvement of 0.7 points experienced by the subjects in this study. The area of interest between these studies is that Allen (1998) and Allen (2000) did not include the use of acupuncture as an adjunct treatment to antidepressants, in contrast, this study did. So, those in the wait group in Allen (1998) and Allen (2000) received no treatment at all. This may suggest that either the change in the scores over the wait period for Allen (1998) and Allen (2000) were unusually large, or the average change in this study was small. Expectation would suggest that subjects who participate in a wait period who are taking antidepressant medication would be more likely to experience greater improvement than those who receive no treatment at all. Given the outcomes on each of the control groups, this expectation appears to not be the case.

It is difficult to account for why the subjects in the wait periods in Allen (1998) and Allen (2000) differed to those in this study. The reason for the differences may be due to the type of subjects who were recruited into the studies. The average length of the depressive episode at the time of entry into the Allen studies was 9.2 months [according to Allen (1998)]. This is much shorter than the average of 3.18 years for those who were recruited into this study. Since the subjects in this study were experiencing longer episodes of depression and they had already been taking their SSRI, their levels of depression may have plateaued. In contrast, the subjects in the wait group in Allen (1998) and Allen (2000) may have been more susceptible to improvements in scores over the wait period because their depressive episode was less chronic. It is reasonable

to assume those who are experiencing a shorter episode of MDD would have a greater propensity for improvement.

It is also interesting to note that the researchers in this study did notice a recurrence of subjects who reported on the long-term nature of their depression. This was often accompanied by a report that they were interested in finding a new treatment since nothing they had tried seemed to be effective (including their SSRI medication). This situation was in effect created by the inclusion criteria in this study. It stands to reason that those who were taking SSRI's and who were experiencing improvements in their mood, would not feel the need to seek out different treatment. In turn, those who were no longer experiencing a benefit from their SSRI's would be more compelled to join the study. This effect may have been further exacerbated by the majority of recruitment for this study being the result of media coverage. The media reports were sensational by nature and suggested hope for those who have been unable to find effective treatment.

The large change in scores over the wait period may account for why the acupuncture treatment in both Allen (1998) and Allen (2000) was not able to significantly outperform the wait list. The lack of significant difference in the between group analysis may not be the result of an under performing acupuncture treatment, but from an unexpectedly large improvement by those in the wait group. The relative change in average scores between the Allen studies and this study support this hypothesis. The change in average scores for those in the treatment group in the Allen studies was an improvement of 11.7 points. This is much the same as the 11.68 improvement experienced in this study. Given this, it may be therefore reasonable to assume the acupuncture in Allen (1998) and (2000) may have been effective if a within group analysis were to be considered. In turn, the absence of a statistically significant difference in the between group analysis may be attributable to the unusually large improvement over the wait period. It must be noted this claim is a postulation. More

information is required than average changes in score to correctly reflect on statistical significance.

The notion of an over performance in the wait period may be further evidenced by the outcomes in Allen (2006). This study used the same methodological design as Allen (1998) and (2000), yet it reported statistically significant differences between the acupuncture intervention and the wait period on both the BDI and the Ham-D. Again, Allen (2006) only reported between group statistical analysis. There was no statistical outcomes nor were the average change in scores given for each of the groups individually. There was however the inclusion of a chart which represented the change in scores. It is of interest to note the change in Ham-D scores for those who received the intervention was similar to this study (approximately 12 points). Moreover, the change in scores over the wait period was more comparable to this study (approximately 2 points).

5.7   Wait List Controls using Medication

More relevant to the design of this study, are those in the literature which used acupuncture as an adjunct to antidepressant medication. Given this, these studies also made comparisons to those in a wait group who also continued to take antidepressant medication. Phase 1 of Luo (1998) and Roschke's (2000) both used this design.

Luo (1998) found acupuncture and the drug amitriptyline were able to significantly improve depression scores using Ham-D analysis. Though, no difference was found between those who received the acupuncture and the medication to those who just received the medication alone.
This outcome of Luo (1998) is different to what was experienced in this study. There are a number of variables which may account for this difference. This study used (manual) acupuncture compared to the electro-acupuncture used in

Luo (1998). It may be plausible that manual acupuncture may simply be better at treating depression than electro-acupuncture. This is difficult to determine and specific research would be needed to test this difference. More likely, the difference could be attributed to the foundations on which the acupuncture treatments were based. This study employed an acupoint prescription which was based on a CM diagnosis of pattern differentiation and the corresponding principle of treatment. Luo (1998) on the other hand used electro-acupuncture between the acupoints of Du20 and Yintang. These acupoints imply this prescription is based on the principle of calming shen. The symptoms associated with depression are however broader then the simple pattern of disharmony of shen disturbance. The researchers in this study more suitably applied the CM model by isolating the most relevant principle of treatment. This may be why this study offered outcomes for the intervention which had greater contrast to those experienced over the wait period.

Luo (1998) did not report a within groups analysis on the change in scores for those who only received the medication alone during the wait period. This makes comparisons with this study difficult. Though, assuming the acupuncture treatments from both studies were equal. The difference in between group comparisons would then be attributable to the difference in the medications. If this were to be the case, the amitriptyline may then be considered to be possibly more effective in the treatment of depression than the SSRI's. This explanation is however less likely. The medical and psychiatric professions prescribe SSRI more frequently than amitriptyline. This implicitly indicates expert opinion considers SSRI's to be more effective than amitriptyline. The difference between the two studies most likely lies within the different application of the CM model.

The subjects in the Roschke (2000) study received the antidepressant mianserin during their involvement in the wait period. This study administered four scales, two of which indicated those who received acupuncture (active and non-specific) and the medication showed significantly greater improvement than those, who

only received the medication. The significant outcomes were on scales which specifically focused on levels of functioning and severity of illness. No difference was found on the two scales which respectively considered global improvement and melancholia. These findings are conflicting and are not consistent with the findings of this study.

Comparisons between the Roschke (2000) and this study are not straightforward. Both studies used different medications and different scales. Roschke (2000) also focused on between group comparisons and did not report any analysis on each of the groups independently. As previously argued, when making comparisons with Luo (1998), the difference between the results of the two studies is less likely to be due to the variance in medication. In the case of Roschke (2000) this would require mianserin to be more effective at treating depression than SSRI's. This is not likely since SSRI antidepressants are again considered by expert opinion to best treat the modern understanding of depression. The interpretation of depression is however a crucial point of interest. Roschke (2000) used different scales than what were used by the researchers in this study. Although both studies used multiple scales, none of the scales were used in both studies. As discussed previously, each of the scales used in Roschke (2000) specifically measured different characteristics of a depressed mood. This may also account for why Roschke (2000) had conflicting results within his study, and may also be the reason why his outcomes differ to those in this study.

Nevertheless, as previously discussed, the difference between the results of this study and those of Roschke (2000) is more likely to be attributable to the difference in how each of the studies maintained the integrity of the CM model. Roschke (2000) did not report a diagnosis based on patterns of disharmony. Nor was there any consideration made with respect to a principle of treatment to formulate the acupoint prescription. In contrast the researchers in this study better maintained the integrity of the CM model.

5.8 Follow Up

The results for the follow up analysis in the study are encouraging. After receiving the acupuncture intervention, subjects were able to maintain their improved depression scores across all three depression measures for a further eight weeks. The BDI (α = 0.05, p = 0.618), the Ham-D (α = 0.05, p = 0.267) and the SCL-90 depression subscale (α = 0.05, p = 0.748) indicated no significant change in depression scores (post intervention compared to the follow up period). Of particular interest is how well the subjects were able to maintain their scores. They were much less depressed at the end of the follow up period than they were before they received the acupuncture intervention. This was again evident across all three depression measures. The BDI (α = 0.05, p = 0.0000), the Ham-D (α = 0.05, p = 0.0000), and the SCL-90 depression subscale (α = 0.05, p = 0.0000) all indicated large statistically significant improvement in scores.

The results from the studies in the literature, which incorporated independent assessments over a follow up period, were consistent with the findings in this study. Using Ham-D analysis, Gallagher (2001) also found no significant change in depression scores over a follow up period. It is interesting to note that the follow up assessment in his study was six months after the conclusion of the acupuncture intervention. This is much longer than the eight weeks used in this study. The outcome in Gallagher (2001) may indicate there is long-term effectiveness of using acupuncture to treat depression. Given this study experienced such statistically large improvements in depression scores, it would be interesting to measure how long the benefits of the acupuncture intervention would have lasted. No such allowance was however made in the design of this study. Additional follow up assessments over extended time frames would have offered information which is more valuable from a clinical perspective.

The findings in this study were also supported by the follow up results of Manber (2004). She found that, 10 weeks after the intervention had concluded, Ham-D depression scores were still significantly less than before the acupuncture was administered. Despite the similarities between the results of this study and those in the literature, it must be noted that both Gallagher (2001) and Manber (2004) focused their studies on pregnant women. Consequently, direct comparisons between the studies need to take this into account.

The between group comparisons made in the follow up analysis of Luo (1990), Manber (2004) and Allen (2006) are not directly relevant to outcomes in this study. These studies used different designs. Quah-Smith (2005) does however offer some interesting information. Her study compared active laser acupuncture to sham laser treatments. Comparisons were made between follow up scores and the base scores before the intervention. She found no statistical difference in the between group comparisons at a four-week follow up. Though, she did find a difference at the 12 week follow up. The subjects who received the active laser treatment were significantly less depressed than those who received the sham laser. This is similar to the follow up outcome in this study. More interestingly were her average changes in BDI scores. These indicated those who received the active laser had experienced an average reduction in scores of 15.8 points over the period of 20 weeks. This is very similar to the average reduction in BDI scores of 14.69 experienced in this study over a 16 week period (which includes the eight week intervention and the eight week follow up period)..

Compared to the studies in the literature, this study was able to offer more detailed insight into the long-term effectiveness of an acupuncture intervention in the treatment of depression. This study was the first which analysed follow up scores in reference to scores both before and after the acupuncture intervention. Given this, more studies with similarly more detailed longitudinal analysis are needed to add to the current literature. When broadly considering the results of

this study and those of Gallagher (2001), Manber (2004) and Quah-Smith (2005), the inconsistencies and variations in the elements of design in each study should be considered to be of minor importance. Collectively, between these studies there is a growing body of literature which suggests acupuncture does have long-term benefits in the treatment of depression.

5.9 Sex Differences

The data in this study was stratified based on sex to facilitate between and within group comparisons. Data on all three of the depression scales was collected. ANOVA analysis was also conducted over the acupuncture intervention as well as over the follow up phase of the study.

5.9.1 Treatment Group

The outcomes for the acupuncture intervention group were similar to those in the main analysis. Both females and males demonstrated highly significant improvement in their depression scores. This was evident on the BDI, Ham-D and the SCL 90 depression subscale (refer to tables 17 and 22). The degree of significance in the changes in scores was the same for both sexes on the Ham-D. Females on the BDI and the SCL 90 depression subscales experienced slightly less significant outcomes. This difference was very small since the P scores only varied by 0.001. This difference may add support to the BDI being a more sensitive measure of depression for the females in the treatment group as seen in the base statistics. This increased sensitivity was also evident in the base statistics according to the Ham-D for females in the treatment group. However the analysis of the outcome data indicates these differences had little or no influence on the outcome of the intervention phase of the study.

There are few studies in the literature which take into account the sex of the subjects. There were a small number of studies which assessed females. Allen

(1998) and Allen (2000) focused their studies on depressed females and Manber (2004) more specifically studied depression in pregnant females. In contrast, there were no studies in the literature which exclusively studied treating depressed men. In turn, there is no available data within any of the studies in the literature which compared differences between the sexes.

None of the studies which considered depressed females offered a within groups analysis. Instead the analysis was based on between group comparisons. This means the statistical analysis only considered how an acupuncture intervention performed relative to a control group. Few of these outcomes are therefore directly comparable to the results from this study. Nevertheless, all three studies, which focused on female subjects, showed encouraging results suggesting acupuncture is an effective treatment for depression in females.

In contrast, Manber's (2004) outcomes were less definitive. In her study, the females who received an active acupuncture, showed no significant difference in changes in depression scores compared to those who received non-specific acupuncture. This was the case according to the Ham-D and as well as on the BDI over the first month and the second month of treatment. It is important to note there was however a significant improvement in scores when all three groups were considered together (active acupuncture, non-specific acupuncture as well as a control who received massage). This again was the case on both the Ham-D and the BDI. As discussed earlier, the inability of the active acupuncture to outperform the non-specific acupuncture may be due to problems associated with the application of the non-specific treatments. This is likely to be the case for Manber's (2004) study. She employed non-specific acupuncture treatments which involved the needling of locations in the same general area as those used in the active treatments. Hence the non-specific acupuncture may have elicited a therapeutic benefit.

5.9.2 Control Group – Wait

The wait list control groups showed no change in depression scores according to sex over the eight-week wait period. No statistical significant change was observed for both females and males (refer to tables 18 and 23). These outcomes indicate the SSRI antidepressant medication alone was not able to significantly improve depression scores in both female and male subjects. This creates a contrast where both the females and males who received the acupuncture significantly reduced their depression scores, compared to the females and males who showed no change when they continued to take only the SSRI's over the wait period. This suggests that acupuncture may be effective in treating depression in both sexes.

The results for females in this study over the wait period are different to those in the literature. The studies of Allen (1998) and Allen (2000) also made comparisons to a wait list group. Their Ham-D analysis found no statistically significant differences between the improvement in scores for those who received acupuncture compared to those who received no treatment. These outcomes suggest the improvement in the depression scores could be attributable to a natural rate of improvement over time. This raises the possibility that the improvement in depression may not be attributable to the effectiveness of the respective acupoint prescriptions.

Subjects in this study continued to take their SSRI medication over the wait period. In contrast the females in the Allen studies did not continue with medication. Assuming the SSRI medication would have a therapeutic benefit over the wait period, it would appear those in the Allen studies would be more likely to improve their depression scores over the wait period. This however was not the case. The Allen studies did not report any within groups' analyses, therefore only the changes in average scores can be considered. The females over the wait period in the Allen studies improved their Ham-D score by 6.1. This

is much more than the average improvement of 0.53 which was experienced by the females in this study over the wait period.

The female subjects in this study and Allen's studies have interesting points of difference. Both of the Allen studies used female subjects who were experiencing a shorter duration of depressive episode. Hence, their depression was characterised as being of a more acute nature. This may have given their subjects a greater chance of natural improvement over the wait period. In addition, there is a question over how well the integrity of the CM model was upheld in Allen's studies. It may be the case that an ineffective use of acupuncture may have restricted the active treatment from outperforming the wait list control. There are a number of possible explanations as to why the results of this study vary so greatly to those in the literature. Further research is needed to provide clarity over this discrepancy.

5.9.3   Control Group – Treatment

Similar to the outcomes for the treatment groups, when the females and males in the control group received the acupuncture intervention, they too demonstrated statistically significant reductions in depression scores on all three depression measures (refer to tables 19 and 24). The outcomes for the females were slightly more significant on the BDI and the SCL 90 depression subscale compared to the results from the females in the treatment group. The outcomes for the males were similarly as significant as those experienced by the males in the treatment group. These results in turn reflect on the existing literature for female sufferers of depression in the same manner as previously discussed for the females in the treatment group.

5.9.4 Sex Differences – ANOVA Interventions.

As already discussed, both females and males demonstrated highly significant improvements in their depression scores over the intervention period according to all the depression measures. To further understand the differences between females and males an analysis of variance was conducted. An ANOVA gives insight into the comparative rate of change in scores.

The data used in the ANOVA considered all females when they received the acupuncture intervention irrespective of them being in the treatment group or the control group. The same consideration was also made for the males. The results indicated there were no significant differences in the rates of change in depression scores between females and males over the acupuncture intervention phases of the study. This means the depression scores for females reduced at the same rate as did the depression scores for males. This outcome was consistent across all the depression measures (refer to table 27).

5.9.5 Follow Up

In the follow up analysis, females and males demonstrated mostly the same outcomes as found in the main analysis. Females were able to maintain their improved depression scores over the eight-week follow up period. No statistically significant change to scores was found for the female subjects on all three depression measures (refer to table 20). Similarly, the males were also able to maintain their improved depression scores over the follow up period, according to the BDI and the Ham-D. The analysis using the SCL 90 depression subscale however indicated males continued to significantly improve their scores over the follow up period despite the discontinuation of the acupuncture treatment (refer to table 25). There appears to be no clear reason as to why there was a difference between the sexes according to the SCL 90 depression

subscale. This will be discussed in more detail when the analysis of variance is later addressed.

These results are consistent with those of Gallagher (2001). Using the Ham-D, Gallagher (2001) found females were able to experience no statistically significant change in scores over a six month follow up period. Like this study, the female subjects were able to maintain their improved status of depression after the treatment had concluded. It must be noted that the follow up periods in each study were different. This study used an eight-week follow up period whereas Gallagher's (2001) was over six months. This makes the results of Gallagher (2001) even more noteworthy. Further study is needed to gain a better understanding of the long-term benefits of using acupuncture to treat depression. To provide the most informative results, this may be best done by implementing multiple follow-up assessments at set intervals.

When stratified scores at the conclusion of the follow up period were compared to the base scores before the acupuncture intervention, the same trends as found in the main analysis were evident. Both females and males demonstrated statistically significant improvement in depression scores at the end of the follow up period compared to their base scores before they received the acupuncture. The statistical outcomes were the same for both sexes across all three of the depression measures (refer to tables 21 and 26). These outcomes are consistent as experienced by the female subjects in Manber (2004).
Using Ham-D analysis, Manber (2004) found women were able to still be significantly less depressed at the conclusion of the follow up period, compared to how they felt before receiving the acupuncture. The results from the follow up analysis for females in this study support her outcomes. In turn it can be said, females are able to remain significantly less depressed eight to 10 weeks after the cessation of an acupuncture treatment compared to before they received the treatment. It is interesting to note, the design for the follow up period in her study and this study are similar. There is only two weeks difference in the length of the

follow up period. This study used an eight week follow up period compared to her 10 week follow up. The similarity in outcomes may be attributable to this common element of design.

Though it must be noted, in her study the female subjects who received the acupuncture showed no significant differences when compared to those who received either non-specific acupuncture or massage. This was the case using analysis based on the BDI and the Ham-D. This outcome may cast doubt over the effectiveness of the active acupuncture in her study. In turn, it may also raise concerns over the legitimacy of the involvement of acupoint specificity in the treatment. The researchers in this study used a different approach. Non-specific acupuncture was not used as a method of control. As a consequence, the results of this study cannot serve to reflect on her findings with respect to her between group outcomes for females. More research is needed into the long-term effectiveness of using acupuncture to treat depression for both sexes.

5.9.6 Sex Differences - ANOVA Follow Up

This study produced mixed outcomes on the ANOVAs which considered differences between sexes over the follow up period. No significant differences were found in the rate of change in scores between the sexes on the BDI and the Ham-D. This indicates females and males were equally able to remain less depressed at the conclusion of the follow up period. As highlighted earlier, this however is not consistent with outcomes according to the SCL 90 depression subscale. This indicated males were able to continue to reduce their scores over the follow up period. This improvement was so great that males demonstrated a statistically significant improvement in their scores over the follow up period (refer to table 25). In contrast, females demonstrated no statistically significant change in SCL 90 depression subscale scores (refer to table 20). This difference between the sexes was reflected in the ANOVA which indicated males improved

their SCL 90 depression subscales scores at a greater rate than females over the follow up period (refer to table 28).

The reason for the different outcome on the SCL90 depression subscale is not clear. It may be possible the SCL 90 subscale may have a greater sensitivity in measuring depression in males. This may be evidenced by how females in the treatment group experienced slightly less significant improvement in scores compared to the males in the treatment group (refer to tables 17 and 22). This theory is however not supported by the results of the females in the control group when they received the acupuncture intervention. They experienced the same significant improvement in scores as did the males (refer to table 19). There are two possible explanations which may account for this difference between the sexes. The depression subscale may be more sensitive in measuring depression in males or it may be due to small sample sizes. This may have lead to statistical anomalies. If this were the case, however, the difference would have also occurred on the BDI and the Ham-D. Despite this conflicting outcome, it needs to be acknowledged that the BDI and the Ham-D are considered to be a more valid and reliable measure of depression, hence more credence should be placed in the outcomes from these two measures.

Further ANOVA comparisons between females and males were made which considered scores at the end of the follow up period compared to scores before the commencement of the acupuncture intervention. In other words, this measured the change in depression scores through the entirety of the subjects' involvement in the study. No statistically significant difference was found between the rates of improvement in depression scores. This was the case across all three depression measures (refer to table 28). This analysis indicated females and males experienced the same rate of improvement in their depression scores at the conclusion of the follow up period compared to before the acupuncture intervention was received. This is of particular interest in regards to the SCL 90 depression subscale. As previously discussed, the SCL

90 subscale indicated males improved at a greater rate than did the females over the follow up period. Though, this result shows both sexes improved at the same rate across the whole study. This would suggest improvement for the males over the follow up period was neutralised by a similar greater rate of improvement for the females over the males in the intervention phase of the study.

The above analysis serves as an initial indication of the differences between the sexes. Females and males were similarly able to significantly reduce their levels of depression when they received the acupuncture intervention. In contrast, both females and males showed no improvement in their level of depression when they only received the SSRI antidepressants over the wait period. This indicated both sexes responded favourably to the acupuncture, and both sexes did not respond when just receiving the SSRI's. Following this, these results suggest that acupuncture is an effective adjunct treatment for depression for both females and males who are experiencing liver qi stagnation. Also interesting were the results on the long-term benefits of the acupuncture intervention. Both sexes were able to maintain their less depressed mood for eight weeks after the conclusion of the acupuncture treatment. These less depressed moods were maintained well enough by both sexes so that they were significantly less depressed than they were before they received the acupuncture. These results are very encouraging as they suggest acupuncture treatments have similarly long-term benefits for both female and male sufferers of depression when it is used an as adjunct to SSRI antidepressant medication.

5.10 Intention to Treat Analysis

As previously discussed, an intention to treat (ITT) analysis was undertaken. This analysis was done using the BDI, Ham-D and the SCL 90 depression subscale. It was also extended across all phases of the study i.e. the wait period, acupuncture interventions and the follow up periods. Consideration was

made for those who opted out of the study by assuming their level of depression remained unchanged for the subsequent assessments in which they did not take part. This consideration was done irrespective of what phase of the study the subject was in when they opted out.

When taking into consideration those who chose to no longer participate in this study, a total of 81 subjects were involved. This was divided into 39 in the treatment group and 42 in the control group. This was lower than the rate of withdrawal which was expected by the researchers at the inception of the study. Considering the subjects in the study were experiencing significant levels of depression, anxiety and a range of other psychological symptoms, it was thought the withdrawal rate would be higher. A total of 15 subjects chose to withdraw their participation in the study. Six of these were from the treatment group, and nine from the control group. It was expected the control group would have a greater amount of withdrawals. The design of the study required them to undertake an extra eight weeks of involvement due to their added requirement to take part in the wait period.

The majority of those who chose to end their involvement in the study were female. A total of 13 females withdrew compared to only two males. Six of these females were in the treatment group and 7 were in the control group. The two males who withdrew from the study were both in the control group. This means a total of 46 females were involved in the ITT analysis. 22 females being in the treatment group, and 24 were in the control group. A total of 35 males were involved in ITT analysis. 18 of these were in the control group whilst the treatment group remained to the standard analysis with 17 male subjects.

The results from the ITT analysis mostly followed the same trends which were found in the standard analysis. This was most evident in analysis of the acupuncture intervention. The outcomes across all three depression scales indicated the intervention was able to significantly improve depression scores.

This was evident when the ITT treatment group received the acupuncture intervention as well as when the ITT control group received the acupuncture at the conclusion of the wait period.

The ITT data over the intervention phase of this study was also stratified to consider differences between the sexes. This produced outcomes which again indicated the same trends as seen in the standards analysis for both sexes on the BDI and the Ham-D and the SCL 90 depression subscale (refer to tables 17 and 22). Interaction between the sexes was found to show no significant differences on the BDI and the Ham-D. Though, there was an interesting point of difference to the standard analysis according to the SCL 90 depression subscale. This indicated males were able to improve at a greater rate than did the females (refer to table 27). This different outcome between the ITT analysis and the standard analysis is likely to be a function of the difference in the number of dropouts between the sexes. Only two males left the study compared to 13 females. This means there were fewer unchanged scores carried through for the males compared to the females. This allowed the ITT analysis for the males to show a stronger trend towards score reductions. For all three scales, the P scores for the ITT sex interactions were lower than experienced by the standard analysis. Each experienced a similar proportional decrease in P scores from the standard analysis to the ITT analysis. The difference for the SCL 90 depression subscale was that its P score for the standard analysis was much lower. This made it more susceptible for the ITT analysis to indicate a significant difference between the sexes.

The ITT analysis of the wait period mostly replicated the results from the standard analysis. This was the case according to the BDI and the Ham-D (refer to table 13). This was also the case for the ITT results on the BDI and the Ham-D when the data was stratified to consider differences between the sexes (refer to tables 18 and 23). The trends in outcomes using the SCL 90 subscale were mostly the same for the standard analysis and the ITT analysis. The ITT main

analysis and the ITT analysis for males was the same as found in the standard analysis. There was however a difference in the ITT SCL 90 depression subscales outcome for females. They were found to significantly reduce their depression scores over the wait period (refer to table 18). This is again due to the large amount of females who dropped out of the study. This meant more unchanged scores were consistently carried forward. As seen in the intervention phase of the study, this affected all the depression measures. All scales experienced a reduction on P scores from the standard analysis to the ITT analysis. Again, The SCL 90 depression subscale was the only measure to change its status of significance in the ITT analysis, since it had a comparatively lower P score in the standard analysis than the other two scales.

The ITT results over the follow up period were mostly the same as the results in the standard follow up analysis. Subjects were able to maintain their improved moods over the follow up period and show no statistically significant change in scores (refer to table 15). Again, the scores at the conclusion of the follow up period were still significantly lower than base scores before the acupuncture intervention (refer to table 16).

The same trends in the follow up analysis was also evident when the ITT analysis stratified data to consider differences based on sex. The ITT trends for both females and males replicated the standard analysis not only over the follow up period (refer to tables 20 and 25) but also when the follow up scores were compared to base scores before the acupuncture intervention (refer to tables 21 and 26). The ITT ANOVAs which considered interactions between the sexes over the follow up period, also mostly replicated the standard analysis. The results on the BDI and the Ham-D showed no difference between the sexes over the follow up period as well as when the follow up scores were compared to base scores. As in the standard analysis, the SCL 90 depression subscale ITT analysis indicated males improved their scores at a greater rate than did the females over the follow up period. The SCL 90 depression subscale outcomes did however differ when the follow up scores were compared to base scores.

ITT analysis indicated males were able to improve at a greater rate than the females. This was not the case in the standard analysis (refer to table 28).

5.11 DSM-IV: Analysis of the Diagnostic Criteria

Analysis using measures of depression can only offer insight into changes in the severity of the symptoms associated with depression. In contrast, the following analysis is based on diagnostic criteria. This type of analysis allows for a better understanding as to whether changes in the severity of depression, as it is experienced by the subjects, results in a change of diagnostic status. The researchers in this study have undertaken this analysis by considering the diagnostic criteria for Major Depressive Disorder as outlined in the Diagnostic and Statistical Manual for Mental Disorders IV (Third Revision). From the above discussion, it is clear the acupuncture as used in this study, is an effective adjunct therapy in reducing depression in those who present with liver qi stagnation and are taking SSRI medication. The following analysis will however determine if these large reductions in symptomology were able to translate into a status of remission from the diagnosis of MDD. The status of remission is achieved once the inclusion criteria for MDD is no longer met.

5.11.1 Main Analysis

The changes in diagnostic status in this study (refer to table 29) followed similar patterns of change which were seen in the analysis based on the depression scales. A comparison before and after the acupuncture intervention resulted in an 87.30% rate of remission from MDD. This calculation included all the subjects who were involved in the study. This can be specifically divided into an 87.88% rate of remission for the treatment group and an 86.67% rate of remission for the control group when they received the acupuncture treatment at the conclusion of the wait period. Of particular note is the consistency between the remission rates of the treatment and control groups. These outcomes offer encouraging

evidence for the use of acupuncture in the treatment of depression. Not only does acupuncture help reduce the severity of the symptoms, which are associated with depression, but it also assists in helping sufferers of depression achieve remission.

Although the studies in the literature commonly used the DSM, it was mostly used for the purposes of inclusion. The only study which did report a rate of remission which considers both females and males combined, was Blitzer (2004). This was a pilot study which had eight subjects complete the treatment program. After the acupuncture intervention, 100% of these subjects were able to achieve remission according to DSM-IV. The remission rates of Blitzer (2004) are higher than those experienced in this study. More than likely, this would be due to variations which are a result of the small sample size used in Blitzer (2004).

5.11.2 Females

Remission rates for the females in this study were similar to those experienced in the main analysis refer to table (30). When all females in the treatment group and the control group were considered together, they experienced a remission rate of 84.85% over the intervention phase of the study. More specifically, these remission rates could be divided up into 81.25% for those in the treatment group and 88.25% for the control group.

The remission rates for the females in this study are very similar to the findings of Manber (2004). She found the females in her study achieved a remission rate 85.7% when they received the active acupuncture intervention. In contrast, these rates of remission for female subjects were comparatively much higher than the other studies in the literature. Allen (1998) and Allen (2000) used the same design and each used DSM –IV to assess depression in females mostly for inclusion purposes. Hence the main remission rate which was reported

included the subjects across all three groups (active acupuncture, non-specific acupuncture and a wait list group). Both of these studies reported rates of remission of 64%. This is much lower than what was experienced in this study and Manber (2004).

Allen (1998) offered more detailed analysis. He specifically reported those who received the active acupuncture experienced a remission rate of 42%. This outcome is unexpected. It would be assumed those who received the active acupuncture would be more likely to have a higher remission rate than those who received either the non-specific acupuncture or who took part in the wait period. This unexpected outcome may be attributed to how well Allen (1998) designed the acupuncture intervention. This study and to a lesser extent Manber (2004) better maintained the integrity of the CM model. In contrast, in Allen (1998) and Allen (2000), poor standards of reporting created the appearance that the measures, which best maintain the integrity of the CM model, were not implemented. For example, Allen (1998) did not report information surrounding the use of pattern differentiation to perform a relevant CM diagnosis. Hence the chosen acupoint prescriptions may not have been correctly based on a corresponding principle of treatment. This may be why this study and Manber (2004) were both able achieve to higher rates of remission for females. Since this study and the three studies from the literature all used the same treatment program, the difference in the remission rates must be attributable to how the acupuncture intervention was carried out.

5.11.3 Males

The analysis of the male subjects in this study also showed similarly high rates of remission (refer to table 31). When all males involved in the study were considered together, the remission rate was 90.00%. The rate of remission for the males in the treatment group was a high 94.12%. This was greater than 84.62% remission rates experienced by the males in the control group when

they received the acupuncture treatment after the wait period. There is a reasonable degree of variation between the remission rates experienced by the males in the treatment group compared to those in the control group. This may be due to the differing number of subjects. Males in the control group numbered only 13 compared to the treatment group of 17. This means the remission rates for the control group were more sensitive relative to each subject who failed to achieve remission. In absolute terms the control group had two subjects who failed to achieve remission compared to one subject in the treatment group. It would be fair to assume the difference in remission rates for the males between the two groups is due to the small sample size in the control group. There were no studies in the literature which specifically focused on using acupuncture to treat depression in males and hence there is no data which could facilitate a direct comparison.

5.11.4 Comparisons Between the Sexes.

The remission rates experienced in this study for both females and males are high. The total rate of remission for females was 84.85% compared to 90.00% for males (refer to tables 30 and 31). The outcomes are similarly close when comparisons are made with the treatment group as well as the control group. The females in the treatment group experienced a remission rate of 81.25% compared to 94.12% in males. Likewise, the females in the treatment group experienced a remission rate of 88.25% compared to 84.62% in males. None of these differences are large enough to warrant a chi square analysis. All differences would clearly turn out to be insignificant. Therefore it can be concluded that females and males in this study experienced similar rates of remission over the acupuncture intervention phase of this study.

5.11.5 Wait Period

Subjects in the control group experienced a low rate of remission during the wait period. The main analysis, which included females and males together, indicated a remission rate of 9.09%. Putting this into context, this means only 9.09% of subjects were able to change their MDD diagnosis when they only received the SSRI antidepressants. This is a very low rate of remission compared to the 87.30% when the acupuncture was administered in addition to the SSRI's. This suggests the acupuncture treatment was an effective adjunct to SSRI's drug therapy in the treatment of depression. It was determined that a chi square analysis comparing the rates of remission between those who received the acupuncture to those who undertook the wait period would not be informative. The result would clearly indicate a significant difference between the groups.

The differences experienced between the remission rates for the treatment group and the wait group in this study is inconsistent with the literature. Allen (1998) was the only study which reported remission rates for those who took part in a wait list group. He found that 20% of the subjects who took part in the wait group experienced remission. Both of these studies used the same time horizon for each of the wait periods. Interestingly, those in the wait list group in Allen (1998) received no treatment, whereas those in this study continued to receive their SSRI antidepressants. Given this, it could be assumed this study would experience a higher rate of remission than Allen (1998). The difference between the two studies is most likely attributable to the subjects who were recruited into each study. This study had no exclusion criteria based on the length of MDD. Consequently the demographic of subjects included in this study were characterised by many long term sufferers of depression. Moreover, because the subjects in this study were also required to be prescribed and taking SSRI's, they were more likely to be experiencing a plateau in their treatment. This may account for the low remission rate over the wait period in

this study. In contrast, Allen (1998) excluded subjects who experienced MDD for a period of time greater than two years. This means the depression experienced by the subjects in his study was less chronic. This may give them a high propensity towards a natural rate of improvement.

Of particular interest was the difference between the sexes over the wait period. No females went into remission over the wait period (i.e. a remission rate of 0.00%). This is markedly different from the males who had a much larger rate of remission of 18.75%. There are a number of possible explanations which could account for this outcome. More than likely it is an anomaly caused by the smaller sample sizes used in the sex comparisons. In the control group there were only 17 females and 16 males. Of the 16 males, three of them went into remission over the wait period. In absolute terms three is not too many but in comparative terms it causes a seemingly large effect. Alternatively, although it is unlikely, it could be postulated that males respond better to SSRI therapy than females. Though, the most likely cause of the variance would be the small sample sizes. Again, no chi square analysis was undertaken to compare the rates of remission between females and males over the wait period. The rates of remission were too close to yield a significant outcome.

5.11.6 Follow Up

The outcomes of the DSM-IV structured interviews at the conclusion of the eight-week follow up period, did suggest some of the patients were unable to maintain their status of remission. There was a relapse rate of 10.34% of the subjects in the main analysis over the follow up period. This means 10.34% of subjects who had achieved remission at the end of the intervention, once again met the criteria for major depressive disorder, at the conclusion of the follow up period. As expected, some relapse in mood was experienced over the eight-week follow up period.

The number of subjects in this study who were able to maintain their remission status until the end of the follow up period is interesting. A total of 76.96% of subjects were able to achieve remission as a result of the acupuncture intervention and were able to maintain their status of remission until the end of the follow up period. This was calculated by subtracting the relapse rate at the end of the follow up period from the remission rate of all subjects at the end of the acupuncture intervention (87.30% remission as at the end of the treatment phase, less the 10.34% relapse rate over the follow up period). This outcome suggests acupuncture is an effective adjunct to SSRI drug therapy in the treatment of depression for those who are also experiencing liver qi stagnation.

There were evident differences between the sexes in relapse rates over the follow up period. Females were found to have a greater rate of relapse than did the males. 14.29% of females experienced a relapse over the follow up period. In contrast the males experienced a relapse of 6.69% over the follow up period. This means that fewer males than females relapsed over the wait period. Conversely, it also means the males were better able to maintain their status of remission.

The differences between the sexes are accentuated when the final remission rates were considered at the conclusion of the study (i.e. at the end of the follow up period). These final rates of remission were calculated by subtracting the relapse rate over the follow up period from the remission rate from the intervention phase. The females experienced a final remission rate of 70.56%. (84.85% less 14.29%) compared to 83.33% (90.00 less 6.67) for the males. More males experienced remission over the intervention period, and fewer males experienced relapse over the follow up period. Both these factors combined to indicate more males were able to have a status of remission at the conclusion of the study. In terms of percentages, there are differences between females and males, these differences are however small. They are too small to

warrant a chi square analysis since no statistically significant differences would be detected between the sexes.

The relapse in depression in the females in this study over the follow up period is less than reported in the literature. The females in Gallagher (2001) were reported to have a relapse rate of 24% over the follow up period. The reason for Gallagher's (2001) higher relapse rates may be due to the observation of a six month follow up period instead of the eight weeks used in this study. In contrast, the relapse rates in this study were higher than experienced by Manber (2004). At the end of the follow up period, those who received the active acupuncture intervention experienced a remission rate of 85.7%. Though, this is close to the final remission rates experienced by the males in this study. Despite the differences in terms of percentages, the differences between the remission rates for females in each of the studies would not be considered to be of significance if they were subjected to a chi square analysis.

5.11.7 Conclusion

The results from this study indicate acupuncture is able to treat the symptoms of depression well enough to effectively assist depression sufferers to achieve a status of remission. Moreover, the few instances of relapses which were observed over the follow up period, indicate the acupuncture treatments can effectively maintain a status of remission over the long term. These trends were consistent for both females and males. The above diagnostic based analysis, suggests acupuncture is an effective treatment to help depression sufferers who are taking SSRI's and who are also experiencing liver qi stagnation.

5.12 Anxiety

As noted in the literature review, anxiety is a comorbidity commonly seen with depression. Anxiety scores are a secondary area of interest in this study The

STAI was consequently administered as a measure of anxiety. This consists of a main scale (or global scale) which provides a general indication of anxiety levels. This global scale is further divisible into measures of state based anxiety, labeled as Y1, and the trait based anxiety, labeled as Y2. The state based anxiety is a measure of the intensity of anxiety related feelings, and the trait based anxiety is a measure of the frequency of anxiety related feelings. The STAI was administered before and after each phase of the study. In addition it was administered midway through the intervention phase of the study. This was recommended in the instruction manual. The SCL 90 anxiety subscale was also scored separately. This was included in the analysis to serve as a secondary and supporting measure.

5.12.1 Treatment Group

Subjects in the acupuncture treatment group demonstrated a statistically significant improvement in their levels of anxiety at the conclusion of the acupuncture intervention. Scores on the STAI global scale indicated a significant reduction in anxiety over the eight week intervention period. This is consistent with the change of scores on the SCL 90 anxiety subscale. The analysis which involved the mid treatment assessment was also encouraging. Significant improvements in anxiety scores were evident over both the first four weeks of the intervention as well as the second four weeks of the intervention. Refer to table 32.

Collectively, the STAI and the state and trait subscales, as well as the SCL 90 anxiety subscale, all indicated statistically significant reductions in anxiety scores as a result of the acupuncture intervention. More specifically, this means the subjects not only experienced less intense feelings of anxiety, they also experienced the feelings associated with anxiety less frequently. These outcomes are encouraging evidence to suggest acupuncture is an effective

treatment to reduce anxiety levels in depression sufferers with liver qi stagnation.

Interestingly were the changes in scores over the second four weeks of the study. The improvement across all three STAI scales, although statistically significant, was not as large as those experienced in the first four-week period. There is more than one possible explanation for this. Either, the greatest level of improvement had been reached not long after the fourth week, and subsequently the improvement plateaued. Or, the frequency of the treatment protocol was the determining factor. In the first four weeks the treatments were administered twice per week. In contrast, the treatments in the second four weeks were administered only weekly. The greater frequency of treatment in the first four weeks may have led to a proportionately greater improvement on anxiety scores. This makes an interesting point of consideration when the follow up results are discussed. If the frequency of treatment was a determining factor, then it would imply anxiety scores may increase over the follow up period.

5.12.2 Control Group

In contrast, the subjects in the control group showed no significant change in anxiety scores over the wait period. This was the case for the STAI global scale and the state and trait subscales, as well as the SCL 90 anxiety subscale. No statistically significant difference in scores was found on the STAI as well as the state and trait based anxiety. A less insignificant change in scores occurred on the SCL anxiety subscale (refer to table 33). These unchanged anxiety scores over the wait period serve as a method of control to which the comparisons can be made to the scores from the treatment group. As previously reported, when the acupuncture intervention was administered there was a significant reduction in anxiety. In contrast when there was no acupuncture administered, there were no changes in anxiety scores. This suggests that acupuncture may be an effective treatment for anxiety. Moreover, the clinical implications of these

results suggest it is better for depression sufferers with comorbidities of anxiety to be treated with acupuncture because it may reduce the intensity and the frequency of the anxiety they experience.

5.12.3 Control Treatment

When the subjects in the control group were later administered the intervention, similar outcomes were observed to those of the treatment group. Across the eight-week intervention, the acupuncture intervention led to statistically significant improvements in anxiety scores on the STAI the state and trait subscales, as well as the SCL 90 anxiety subscale. Again, similar to the outcomes from the treatment group, significant reductions for all scales were evident over the first four weeks of the intervention. These statistical outcomes were similarly as large as those experienced by the subjects in the treatment group. Refer to table 34.

In contrast to the results from the treatment group, the statistical analysis over the second four weeks of the intervention differed. Like the subjects in the treatment group, the subjects on the control group also showed significant reductions in anxiety according to the STAI and the trait subscale. The state subscale for those in the control group however did not indicate a significant reduction in anxiety. This means, for those in the control group, the state based anxiety stabilised over the second half of the intervention. The reason for the difference for state based anxiety scores between those in the treatment group to those in the control group is unclear. There is no salient explanation which could suggest why the state based anxiety measure would be more sensitive for one group over the other. It may well be that acupuncture has a greater consistency in being effective at improving trait based anxiety. Conversely, acupuncture may not consistently be able to effectively reduce state based anxiety. The most probable explanation for the difference may well be due to slight differences in timing when each of the groups experienced a plateau in

their progress. The control group may have just improved slightly faster and plateaued sooner than those in the treatment group. This is difficult to determine since data on the STAI was not collected on a weekly basis. Though it does however, appear the difference between the two groups is a statistical anomaly. This conclusion is supported by the fact that the longer term change in scores over the eight weeks was similar in both groups. In addition to this, the statistical significance of the control group's change in state based scores over the second four weeks of the intervention was not overwhelmingly large.

5.12.4 General Conclusion

These results indicate that anxiety scores were significantly reduced over the acupuncture intervention. In contrast, anxiety scores for the control group remained unchanged over the wait period. This suggests acupuncture is an effective treatment for anxiety for this population. The clinical significance of this outcome is that acupuncture may be an effective treatment for anxiety as experienced by sufferers of depression. Not only were their general levels of anxiety shown to improve, but more specifically, the frequency and the intensity of their anxious feelings benefited from the acupuncture treatment. These outcomes also serve to highlight the close co-morbidity between depression and anxiety. In addition, they are also informative from a CM perspective. These outcomes may offer insight into the broader implication of liver qi stagnation in mental health. Since it appears the principle of treatment of dispersing liver qi stagnation is also effective for treating anxiety.

5.13 Sex Differences

5.13.1 Treatment

Both females and males in the treatment group demonstrated statistically significant reductions in anxiety scores due to the acupuncture intervention. The scores for both sexes significantly improved on the STAI, the state subscale and trait subscale as well as the SCL 90 anxiety subscale (refer to tables 37 and 42). Interesting differences between the sexes emerged when reductions in scores over the first four weeks of the intervention were compared to those over the second four weeks. There were no differences between the sexes over the first four weeks of the intervention. Though in the second four weeks, females and males demonstrated vastly different responses to the treatment. According to the STAI and it's subscales, females showed no change in anxiety scores over the second four weeks of the acupuncture intervention. In contrast, males were able to continue to demonstrate improvement. This meant males were able to significantly improve their STAI scores in both the first and second weeks of the treatment.

It is difficult to account for this difference between the sexes. There are no acupuncture studies in the literature which compare sex differences in the treatment of anxiety as experienced by sufferers of depression. As discussed earlier, the only studies which considered the sex of subjects were Allen (1998), Allen (2000) and Manber (2004). All three of these studies focused on treating women. No studies in the literature considered the treatment of males. Moreover, the studies in the literature which did focus on females, did not administer any anxiety measures. This study is the first to make provision for such comparisons. The outcomes in this study have created a platform from which further research can be based. There remain some interesting implications concerning differences between the sexes which are yet to be explored. Further studies may consider if sex differences in depression sufferers

create variances in the sensitivity in the change in anxiety scores due to an acupuncture treatment. Or conversely, the profile of the various mental health symptoms associated liver qi stagnation may have varying sensitivities in their response to acupuncture treatment. This study did not administer the depression scales mid treatment so no comparison between depression and anxiety scores can be provided for this time point.

5.13.2 Control – Wait Period

Both sexes demonstrated similar trends in anxiety scores over the wait period. Females and males alike showed no significant change in anxiety scores on the STAI as well as the state and trait subscales. Females did however demonstrate a significant improvement in SCL 90 anxiety subscale scores over the wait period. No significant change was however evident for males on the SCL 90 anxiety subscale (refer to tables 38 and 43).

There is no clear reason as to why only females demonstrated significant improvement in anxiety scores over the wait period according to the SCL 90 anxiety subscale. Though, this outcome has to be put in context. The STAI is considered to be a more valid and reliable measure of anxiety. This is why within this study, the SCL 90 anxiety subscale is being used as a secondary scale. Consequently, this discrepancy between females and males on the SCL 90 subscale over the wait period should be considered within this context. Moreover the statistical significance in the change in the females score was marginally significant ($\alpha = 0.05$, p = 0.048). Taking these points into consideration, it is valid to conclude that females and males did not substantially differ over the wait period.

5.13.3 Control – Treatment

Results for females and males when the control group received the acupuncture intervention followed the same trends as when both sexes in the treatment group received the intervention. Consequently, both sexes demonstrated significant improvement in anxiety scores across all scales. Again similar to the outcomes experienced by those in the treatment group, both sexes in the control group also showed significant reductions in anxiety scores during the first four weeks of treatment (refer to tables 39 and 44).

The females and males in the control group also showed differing trends between the sexes on the STAI over the second four weeks of treatment. As discussed earlier, the females in the treatment group showed no change in anxiety over the second four weeks of treatment, whereas males did significantly improve. In contrast, the results in the control group were the reverse. The males did not show any change over the second four weeks of treatment but the females did. For the males, no significant change in scores was found on the STAI and the state and trait subscales. This was also true for the females according to state based analysis. The females however were able to show significant reductions in their scores in the STAI and the trait based analysis. This further exacerbates the uncertainty previously discussed in reference to the results from the treatment group. Moreover, these outcomes add further weight to the need to find better clarity as to why these sex differences have occurred. The anxiety results in this study have been inconsistent over the second half of the intervention and further study is warranted.

5.13.4 ANOVA

Analysis of variance over the entire eight-week treatment phase indicated no differences between the sexes. This analysis combined the data from the treatment group and control group to form respective singular groups of females

and males alike. Scores before the acupuncture intervention were compared to those after the acupuncture intervention. The results indicated that female and male anxiety scores improved at the same rate across all scales (refer to table 47).

5.13.5 Sex Comparison Conclusion

The ANOVA results indicated females and males improved at the same rate over the eight-week intervention period. As previously discussed, other results indicated this rate of change may not have been consistent between the two sexes throughout the eight weeks. This was shown by the STAI analysis when comparing the sexes over the second four-week period of the intervention for both the treatment group and the control group. Despite the variations which occurred within the intervention period, it is interesting to note the differences had evened out at the conclusion of the eight weeks period. This may give us some insight to suggest the discrepancies were due to the time horizon. It may be hypothesised that the four week time frame used in the mid treatment comparisons was too short and hence it facilitated anomalies in the statistical outcomes. These discrepancies should however be considered to be of minor importance. According to the major trends in the study, both females and males were able to equally experience significant reductions in anxiety scores over the full eight weeks of the intervention. This may therefore suggest that both female and male sufferers of depression in this population can effectively and equally have their anxiety treated using acupuncture.

5.14 Follow Up

The results from the follow up analysis indicated subjects were able to maintain their improved anxiety scores eight weeks after the conclusion of the treatment. Anxiety scores remained unchanged on the STAI, the state subscale and the SCL 90 anxiety subscale. More encouragingly, the trait based subscale

indicated subjects were able to continue to significantly improve their scores during the follow up period despite the discontinuation of the acupuncture intervention (refer to table 35). The improved anxiety was so well maintained over the follow up period that subjects were able to remain significantly less anxious than they were before they received the acupuncture intervention. This was found on the STAI, the state subscale and the trait subscale, as well as the SCL 90 anxiety (refer to table 36). The implication of these results indicate subjects were able to maintain their significantly reduced anxiety over the eight week follow up period and still be significantly less anxious than before they were treated. These are encouraging outcomes since they suggest acupuncture can assist depression sufferers to continue to feel less anxious after an acupuncture treatment is no longer being administered.

An additional point of interest is to discuss the follow up results in reference to an issue which was previously raised. This involved the postulation as to if the frequency of treatments affected the changes in anxiety scores. Given the above outcomes over the follow up period, it would indicate that the removal of the acupuncture intervention did not negatively affect the anxiety scores. This therefore makes it less likely that the reduction from two treatments per week to one per week over the second half of the intervention phase would account for the difference in improvement. This in turn would suggest that subjects most likely experienced a plateau in their improvement in their anxiety during the second four weeks of treatment.

5.14.1 Follow up - Sex Differences - Within Group

When the results over the follow up period took into consideration the stratification of the sexes, variations on the general trend emerged. Females demonstrated no statistically significant change in anxiety scores over the follow up period on the STAI the trait subscale and the SCL 90 anxiety subscale. They did however show significant regression in their levels of anxiety according to

the state subscale (refer to table 40). On the other hand, males were able to maintain, if not improve their anxiety score over the follow up period. The males maintained their anxiety scores from the intervention on the STAI and the state subscale. They continued to show improvement on the trait subscale and the SCL 90 anxiety subscale (refer to table 45).

Despite the variations between the sexes over the follow up period, both females and males managed to be significantly less anxious at the conclusion of the follow up period compared to how they felt before they received the intervention. Both sexes showed equally high statistical effects across each of the anxiety measures (refer to tables 41 and 46).

5.14.2 Follow Up - ANOVA Sex Comparison - Between Group

Contrary to some of the within group differences between females and males over the follow up period, analysis of variance results mostly suggested both sexes showed the same rate of change on their scores. Over the follow up period, no statistical difference was found in the rate of change in anxiety scores between the sexes on the STAI as well as the state and trait anxiety subscales. The SCL 90 anxiety subscale did however suggest that males did improve at a greater rate than the females. This trend in outcomes was repeated when follow up scores were compared to base scores at entry to the study. No significant difference was found in the rate of change between the sexes according to the STAI, the state subscale, and the trait subscales. Similarly, on the SCL 90 anxiety subscale, males improved their scores at a significantly greater rate than did the females. Refer to table 45.

The SCL 90 anxiety subscale results indicated males were able to continue to improve their anxiety scores after the treatment had concluded. This enabled them to be significantly less anxious than females at the conclusion of the follow up period. Though, this observation on the SCL 90 anxiety subscales needs to

be put into context. The STAI and its subscales are considered a more reliable and valid measures of anxiety. More weight in the analysis should therefore be given to the outcomes from the STAI which suggested both sexes were able to equally maintain their improved anxiety scores over the follow up period. Both sexes were also equally less anxious at the conclusion of the follow up period compared to how they felt before they received the acupuncture. This implies acupuncture is equally as effective for depressed females and males in maintaining reduced intensity and frequency of anxious feelings at the cessation of the treatment.

5.15 SCL-90-R Subscales

The Symptomatic Checklist 90 R (SCL 90) is a general indictor of psychological status. It has been included within the analysis of this study to investigate the co-morbidity of liver qi stagnation with other psychological primary dimensions. The SCL 90 consists of nine separate dimensions or subscales. Each of the subscales were designed to reflect the understanding of each syndrome from a medical and psychiatric understanding as well as from the more common perspective as understood in the community. The subscales include somatization, obsessive-compulsive, interpersonal sensitivity, hostility, phobic anxiety, paranoid ideation and psychoticism. The remaining two subscales are depression and anxiety. The outcomes from these two subscales have already been addressed in the previous analysis on depression and anxiety. Following is an explanation of global scale and each of the subscales not previously covered.

5.15.1 Somatization (SOM)

The somatization subscale is an indicator of physical discomfort. It measures the degree of distress which is caused by the perception of dysfunction within the body. The items specifically focus on perceived disruptions in autonomic functions. They include symptoms associated with the cardiovascular,

gastrointestinal and respiratory systems. Musculoskeletal discomfort and the somatic symptoms which are associated with anxiety, are also included in the subscale. (Derogatis, 1994)

This definition of somatization can be likened to the common understanding of liver qi stagnation. The manualisation process which was undertaken by the researchers in this study, identified multiple symptoms of liver qi stagnation which are by nature associated with somatization. In general terms, these symptoms include hypochondriac stuffiness and distention, chest stuffiness, abdominal/epigastric distention. On closer consideration it is evident that eight of the 12 items used to measure somatization on the SCL 90 are closely linked to liver qi stagnation. These SCL items and their matched symptoms of liver qi stagnation are:

- Item 12: Pains in the heart or chest – chest/hypochondriac stuffiness/distension;
- Item 42: Soreness of your muscles - pain in arms and shoulders when walking;
- Item 48: Trouble getting your breath – chest stuffiness;
- Item 53: A lump in your throat - plumb stone throat or difficulty swallowing;
- Item 58: Heavy feelings in your arms or legs - pain in arms and shoulders when walking.

The SCL item which is closely associated with liver qi stagnation combined with wood invading earth is Item 40. This item is "nausea or upset stomach". This is consistent with a number of wood invading earth symptoms. These are:

- Abdominal distension and pain;
- Fullness of epigastrium/bloating;
- Sour regurgitation;
- Belching;
- Nausea;

- Vomiting, and
- Sour vomiting.

With the somatization subscale there are also two items which reflect on liver qi stagnation with heat. These items and their liver qi stagnation matches are:
- Item 1: Headaches – headaches;
- Items 49: Hot or cold spells – subjective feelings of heat. Although this symptom did not feature in the manualisation process, it is an assumed and self-evident symptom consistent with the CM understanding of heat.

It can be argued that many of the above items of the SCL 90 item listed above are also representative of other CM patterns of disharmony. Though given subjects in this study were included based on the presentation of liver qi stagnation, and given the principle of treatment and acupoint prescription treated liver qi stagnation, it can be assumed changes in the scores for each of these items would hence reflect on liver qi stagnation in this instance. For the purpose of this study, the somatization subscale can serve as an informative measure for key diagnostic criteria of liver qi stagnation.

5.15.2 Obsessive-Compulsive (OC)

The obsessive-compulsive subscale is designed to measure the severity of the symptoms associated with the clinical understanding of the syndrome of obsessive-compulsive disorder. These symptoms are characterised as being unremitting and irresistible by nature. The items in this subscale are designed to reflect the thoughts, impulses and actions of the subjects. This subscale also includes the experiences and behavior which are associated with general cognitive performance deficit. (Derogatis, 1994)

5.15.3 Interpersonal Sensitivity (IS)

Interpersonal sensitivity is a measure of how the subjects perceive themselves relative to their understanding of others. The items in this subscale reflect on symptoms such as feelings of inadequacy and inferiority. These symptoms are consistent with those which occur during interpersonal interactions, such as self-depreciation self-doubt and marked discomfort. Individuals who score highly for interpersonal sensitivity offer reports of experiencing acute self-consciousness, negative expectations surrounding their interpersonal behavior with others, as well as the negative perceptions of others which are directed at them. (Derogatis, 1994)

5.15.4 Phobic Anxiety (P.ANX)

Phobic anxiety is characterised by a persistent response of fear. This is an additional measure to the anxiety subscale discussed earlier. Phobic anxiety is more specifically defined as being in response to a person, place, object or situation. The fear response is irrational by nature and the degree of the response is disproportionate to the real threat of the stimulus. This often leads to either avoidance or escape behavior. The items on this subscale are reflective of the more pathological and disruptive manifestations of phobic anxiety. Given this, the subscale measure symptoms which are also closely associated with conditions such as agoraphobia and phobic-anxiety depersonalization syndrome. (Derogatis, 1994)

5.15.5 Hostility (HOST)

The hostility subscale measures symptoms which are the result of a state of anger. The items in the subscale represent the three modes of expression. These being, the thoughts feelings and actions which are associated with anger.

The subscale reflects on symptoms such as aggression, irritability, rage, and resentment. (Derogatis, 1994)

The hostility subscale has a common characteristic with the conventional understanding of liver qi stagnation. As seen in the manualisation process, amongst the symptoms which are associated with liver qi stagnation, are frustration and irritability. The six items in the hostility subscale collectively share characteristics which can be generally described to reflect on levels of frustration and irritability. This means the hostility subscale can be used as relative indicators of these aspects of liver qi stagnation.

5.15.6 Paranoid Ideation (PID)

The paranoid ideation subscale is designed to measure the behaviour consistent with the primary reflections of this disorder. The symptoms of paranoid ideation include: characteristics of projective thought; hostility; suspiciousness; grandiosity; centrality; fear of loss of autonomy and the delusions which are consistent with this disorder. (Derogatis, 1994)

5.15.7 Psychoticism (PSYCH)

The psychoticism subscale represents symptom on two levels. Firstly it considers symptoms which are generally synonymous with the understanding of a schizoid lifestyle, such as being withdrawn and isolated. Secondly, it considers symptoms which are explicit inclusion criteria such as hallucinations and thought control. (Derogatis, 1994)

5.15.8 Global Scale

The nine subscales are summarized by a global scale, which is simply created from the aggregate of all the subscales. This global scale offers greater flexibility

as an indicator of the subject's psychopathological status. It also serves as a general indicator of symptomology and psychological distress. The SCL 90 global scale has been used in a number of studies as a reliable and valid measure of stress (Derogatis, 1994). This is interesting in the context of the study. According to the CM model, a stressful lifestyle is one of the etiologies which is commonly understood to cause liver qi stagnation. This suggests the SCL 90 global scale can be used as an indicator on the stress component of liver qi stagnation. For the purposes of this study, the global scale will in turn reflect on the severity of a broad range of psychological symptoms. This may offer insight into the psychological benefits of dispersing liver qi.

5.15.9 Trends in the SCL 90

The acupuncture intervention, led to significant improvement in all the psychological dimensions considered in the SCL 90 (refer to table 49). In contrast, when the control group only received the SSRI antidepressants over the wait period, no corresponding improvement occurred (refer to table 50). Then, consistent with those in the treatment group, significant improvements according to all the psychological dimensions were also found when those in the control group were administered the acupuncture intervention (refer to table 51). This suggests that acupuncture is effective in reducing a broad range of psychological symptoms which are experienced by depression sufferers. Further to this is the implication on liver qi stagnation. These results suggest the treatment principle of dispersing liver qi is effective in improving a range of psychological symptoms.

According to the SCL 90, there were few differences found between the sexes over the intervention phase of the study. When both sexes in the treatment group and control group received the acupuncture intervention, significant improvements in scores were experienced on the SCL 90 global scale and most of the SCL 90 subscales (refer to tables 54, 56, 59 and 61). Differences between

the sexes were only evident in the treatment group on the phobic anxiety and the paranoid ideation subscales. The females in the treatment group showed no significant change in scores on both these subscales. On the remainder of the stratified analysis it suggested both sexes in the treatment group and the control group experienced significant reductions in scores on the global scales and the remainder of the subscales. The sex based analysis for the control group over the wait period followed the same trends as seen in the main analysis. No significant change in scores was experienced over the wait period (refer to tables 55 and 60).

It is difficult to account for the sex differences on the phobic anxiety and the paranoid ideation subscales. There appears to be no explanation as to why females and males in the treatment group differed. Moreover, it is also difficult to explain why the females in the treatment group differed from those in the control group. The only salient explanation may attribute these differences to the small sample sizes in the subgroups when the data from the control group was stratified. This theory is supported by the results from the ANOVAs. The data used for the ANOVAs combined the data from the treatment groups and the control groups (relative to each sex) over the intervention phase of the study. This created larger sample sizes. The results from the ANOVAs suggested there were no differences between the sexes. Both females and males improved their scores at the same rate on the SCL global scale and all the subscales. This included the phobic anxiety and the paranoid ideation subscales. (Refer to table 64).

5.15.10 Follow Up

The main analysis for the SCL 90 followed the same trends as seen in the depression and anxiety scales. No significant change in scores occurred on the global scale and the subscales over the follow up period (refer to table 52). Again consistent with the previous analysis, the subjects were able to maintain

their scores over the follow up period to remain significantly better than before the acupuncture intervention was administered (refer to table 53).

The majority of the outcomes, which separately considered females and males over the follow up period, replicated the trends in the main analysis (refer to tables 57 and 62). There were however some sex differences over the follow up period. According to the global scales the males were able to continue to improve their scores even though they were no longer receiving the acupuncture intervention. Males showed significant improvement in their global scales over the follow up period. In contrast the females demonstrated no significant change. This was reflected in the ANOVA which indicated males were able to improve at a greater rate than females in the global scales (refer to table 65).

The stratified results for the obsessive-compulsive subscale replicated the outcomes from the global scale. The males were able to continue to significantly improve their scores over the follow up period (refer to table 62). In contrast the females showed no significant change in scores (refer to table 57). This again was reflected in the ANOVA which suggested males were able to improve scores at a greater rate than females (refer to table 65).

The final difference between the sexes over the follow up period was found on the hostility subscale. Females experienced a worsening in their levels of hostility. Their scores significantly increased (refer to table 57). In contrast the hostility scores for the males indicated no significant change refer to table 62). This was reflected in the results of the ANOVA. The rates of change between the sexes were found to be of statistical significance. The females increased their hostility scores at a greater rate than the males (refer to table 65).

When the stratified scores at the conclusion of the follow up period were compared to those before the acupuncture intervention, the trend in outcomes was mostly consistent with the main analysis. Both sexes were able to maintain

their significant improvement in scores and remain better off than before they received the acupuncture intervention (refer to tables 58 and 63). The ANOVAs also suggested there were no differences between the sexes. Both females and males experienced the rate of change on their scores on the global scales and most of the subscales. The only exception was the phobic anxiety subscale (refer to table 65).

The analysis on the phobic anxiety subscale offered an interesting outcome. Both females and males experienced significant changes in scores over the follow up period (refer to tables 58 and 63). The ANOVA however suggested males were able to improve their scores at a greater rate than females, though the difference in the rate of change between the sexes, was not highly significant ($\alpha = 0.05$, $p = 0.049$). This was possible because the females experienced approximately half the rate of improvement than what was experienced by the males.

It is difficult to account for any of the differences which have arisen between the sexes in the follow up analysis using the SCL global scale and its subscales. This is the first acupuncture study designed to treat depression which has considered a broader understanding of mental health as a part of the analysis. No other study, which has used acupuncture to treat depression, has previously administered the SCL 90. Hence, no precedents exist within the literature to serve as a comparison. Further research is needed to add to the SCL 90 outcomes found in this study.

5.16 Limitations

5.16.1 Design – Wait Group

Standards of best practice in the application of scientific method demand the implementation of sound design. In practice, suitable designs for acupuncture

clinical trials are not easy to construct. They are often fraught with difficulties and controversy, much of which is focused on finding an appropriate method of control. Those with an unsophisticated understanding of acupuncture encourage the use of non-specific or "sham" acupuncture. They reason these would best serve as a test of acupoint specificity and hence offer a better reflection on the effectiveness of an acupuncture treatment. This reasoning is however flawed and it is not consistent with a deeper understanding of the CM and acupuncture paradigm.

The researchers in this study chose to compare an active acupoint prescription to a wait list control group. This is believed to be the better fit when the complexities of acupuncture are taken into consideration. wait list control groups however do have inherent problems. Most notably, the potential positive biases such as the placebo effect cannot be fully negated. Using a wait list group as a means of control omits the opportunity to specifically test the therapeutic benefit of the individual acupoints. Consequently, this study cannot critique the effectiveness of the acupoint prescription. Inferences can only be drawn in a more broad sense and be attributed to the acupuncture experience in general. The broader experience of an acupuncture treatment may additionally involve a number of positive biases, such as patient subject interaction. The implication is that this study can only make a true inference on the question of whether it is better to treat or not to treat using acupuncture. Although the researchers have made every attempt to construct a suitable design for an acupuncture clinical trial of this nature, the inherent shortcomings in the use of a wait list as a method of control remain. This leaves the outcomes of this study to be susceptible to the associated positive biases. This must be duly noted as a limitation of this study.

5.16.2 Design - Blinding Subjects

The design in this study did not allow for the full blinding of subjects. By nature, the design was simple. It posed the question of to treat or not to treat. No

provisions were implemented which would prevent the subjects from this realisation. For example, no suggestion was made as to the possibility of receiving a sham treatment. The omission of such a step was withheld for the design of this study following advice based on ethical considerations. Consequently, when the subjects in both the treatment group and the control group received the acupuncture intervention, they had no grounds to doubt whether they were receiving an acupuncture intervention. This may have provided a basis for a positive bias.

More fundamentally, subjects in the control group, despite being blind to the fact they were in a wait group, had no change in the treatment regimen. They therefore had no reason to suspect their condition would change. In contrast, those in the intervention phase of the study, received a clear change in treat regimen in the form of the introduction of the acupuncture treatment. This would induce an expectation of improvement in the conditions as experienced by the subjects. This design may have led to a situation, which may have exacerbated the relative expectations in each of the groups. On reflection, it may have been more informative if some form of extra treatment was added over the wait period. This would have changed the fundamental direction of the study but it would have helped to reduce the expectations of subjects.

5.16.3 Design – Blinding Researchers

Best practice in scientific method requires blinding of the researchers. In order to maintain the highest scientific integrity, this study ideally would have required three researchers. This would have allowed one researcher to perform the assessments, another to assign the subjects into groups and a third to administer the treatments. The study however was undertaken for the purpose of post-graduate research. This carries with it inherent financial limitations. This limitation of funds prevented the involvement of the preferred number of researchers. It could be argued that not enough precautions were implemented

to minimize the possible influence of an "informed" researcher. This leaves open the potential to influence the outcomes of the study. Although the post graduate researcher in this study carried out his duties with the highest intentions to uphold scientific integrity, an implicit risk still remained.

5.16.4 Randomisation

Randomisation is an important element of any clinical trial design. Due to the inherent limitations of this post graduate study in terms of funds and time for completion as well as the nature of the trial design, a method of simple randomisation was chosen over more complex systems. It could be argued that more sophisticated methods could have been employed, but this would be more suited to a similar study with a much larger sample size.

5.16.5 Practitioner Expertise – Assessments

Elements of this study required knowledge and expertise in psychological assessment. Expertise was needed when performing the DSM-IV structured interview to diagnose MDD and when administering the Hamilton Rating Scales for Depression. Both of the DSM IV and the Ham-D take the form of a structured interview. Although both were designed in a manner to facilitate standardisation and consistency in their application, they do require some level of expertise and knowledge. The researcher in this study had a bachelor degree majoring in psychology, but was not a practicing nor registered psychologist. Consequently he cannot be considered to be an expert in diagnosis or assessing depression.

One of the founding co-supervisors in this study was a psychologist and considered to be an expert in depression. He provided training to the student for quality assurance purposes. This training involved a theoretical understanding and practical proficiency in both the DSM IV and the Ham-D structured interviews. It is interesting to point out the changes in Ham-D scores in this

study were similar to those which were reported in the literature by Allen (1998), Allen (2000) and Manber (2004). Despite the consistency with the studies in the literature, it should be acknowledged the post graduate researcher was not considered to be expert in the diagnosis and assessment of depression. Consequently, this should still be noted and considered to be a limitation of the study.

5.16.6 Recruitment

The majority of subjects were recruited via the media. This allowed for sufficient numbers to be found in a timely manner. It also created a wide spread awareness of the study which allowed the researchers to find numerous suitable subjects. This however came at a cost due to the nature of stories in the media. Some of these stories were sensational and made claims that acupuncture may be a solution for those who were not experiencing benefits from conventional treatment.

This media coverage may have contributed to a positive bias in the study since subjects may have joined the study with the expectation of improving their levels of depression from the acupuncture treatment. In addition, the style of publicity inherently attracted many long-term sufferers of depression. Amongst them were many who were continuing to receive SSRI's despite no longer finding any therapeutic benefit. This may have influenced the lack of change on depression scores for those in the control group. In turn, this may have also served to exacerbate the differences in effectiveness between the intervention and wait period phases in this study.

5.17 Relevance of Liver Qi Stagnation

The researchers in this study focused the treatment protocols on treating liver qi stagnation. This was founded on the assumption that liver qi stagnation,

according to the model, can be considered to be the root energetic imbalance which gives people the propensity to feel depressed. Following this assumption, the principle of treatment used in the acupuncture intervention was based on dispersing liver qi. In order to increase subject inclusion in the study, the additional liver qi related patterns of disharmony of wood invading earth and liver qi stagnation with heat were added. Subjects falling under each of the respective patterns of disharmony were administered the correct and corresponding principle of treatment.

It must be duly noted that this study was not designed to test the relevance of liver qi stagnation to MDD compared to other CM patterns of disharmony. Nor are the researchers claiming that when using acupuncture, the most appropriate principle of treatment for all cases of depression is to disperse liver qi. Instead, the aim of this study was to test the effectiveness of using acupuncture to treat depression as an adjunct treatment to SSRI's in those who presented with liver qi stagnation. The focus on liver qi stagnation was included as an element of design to better facilitate the integrity of the CM model. As discussed earlier, the CM model requires an acupoint prescription which is based on a principle of treatment to rectify a pattern of disharmony. The researchers were of the opinion that the symptoms of liver qi stagnation would most likely account for depression in the population from which the study sample would be drawn. If this study was designed to test the best principle of treatment for depression, additional treatment groups based on alternative CM diagnoses would have been needed. Each treatment group would have to be administered the respective acupoint prescription relative to the different principles of treatment. For example, as done in Luo (1985), Luo (1990) and Luo (1998), one group may have been administered GV20 and Yin Tang to perform the principle of treatment of *calming shen*.

As discussed in the limitations of this study, a consequence of the design means an important question remains unanswered. Was it dispersal of the liver qi which

led to the improvement in depression scores? Or, was it due to the experience of receiving the acupuncture? The latter proposition raises the question of the involvement of a placebo effect. It may well be argued, that the virtue of allowing the subjects to relax on the treatment tables for 20 minutes at each treatment, may have given them relaxation time which they may not otherwise have engaged in. This may have led them to improve their depression scores. Alternatively, the improvement in depression could have simply been the result of expectation on the part of the subjects. Other explanations such as the patient-practitioner interaction are equally as feasible. Since in this study, non-specific acupuncture treatments were not used as a method of control, it is yet to be conclusively determined if the acupoint specificity (which dispersed liver qi) is responsible for the improvement in depression scores.

Given the large statistically significant outcomes witnessed in the results of this study, it appeals to reason however, that liver qi stagnation must have a significant association with depression. The size of the effect appears to be more than would be expected from a positive bias. It must also be highlighted that compared to the studies in the literature, this study better upheld the integrity of the acupuncture and CM models. This may be why this study, compared to some studies, was able to produce greater effect sizes and outperform the method of control. Those with acupuncture clinical experience, although anecdotally, are aware that by isolating a pattern of disharmony and applying an associated principle of treatment, more effective treatments would follow.

This study raised another point of interest concerning the understanding of liver qi stagnation and its relevance to mental health in general. From a CM perspective, one of the key psychological diagnostic criteria for liver qi stagnation are feelings of irritability and frustration. These are often expressed physically in the form of frequent sighing. Interestingly, a practical or clinical understanding of "irritability and "frustration" can more generally and more

simply be deemed to be "mental unrest" or "stress". Following this reasoning, the symptoms which are associated with liver qi stagnation, can be explained as the physical manifestations of psychogenic causes. This can be evidenced by the list of symptoms of liver qi stagnation which were used as inclusion criteria for this study and commonly associated with stress.

Part of the analysis involved in this study can serve as evidence for the effectiveness of treating liver qi stagnation to reduce stress. The SCL 90 has a history of being used as a reliable and relevant measure of stress. In effect, its nine primary dimensions offer varying insights into varying qualities of stress. Based on the SCL 90 results from this study, it can be concluded that stress can be reduced from an acupuncture intervention based on dispersing liver qi. Moreover, given this study was able to produce such large statistically significant improvements, it may be fair to suggest acupuncture is a very effective modality in the treatment of stress.

Further to this, the diagnostic criteria for MDD, also has an association with stress. The diagnostic criteria in DSM-IV is characterised by symptoms which can be construed as being psychogenic by nature. These include changes in appetite, sleep disturbances, fatigue / loss of energy and difficulty in concentrating. It is also theorised that depression is often caused from long term suffering of anxiety (which inherently is closely related to the common understanding of stress). The disempowerment and helplessness from the long-term anxiety is said to lead to despondency, and in turn lead to depression.

Interesting to note is the feedback offered by subjects both during and at the conclusion of their involvement in the study. Some subjects reported the acupuncture made them feel more relaxed after each of the treatments. They also reported this made them feel less irritable and frustrated, which also helped them have better quality sleep. This in turn enabled them to have more energy, which then allowed them to approach their everyday tasks with renewed vigour.

This gave them a sense of empowerment and made them feel more capable and have improved levels of functioning. This all resulted in allowing them to better deal with their depression. In effect, the dispersal of the liver qi, freed them from the signs and symptoms which caused the sense of restriction, which in turn led them to feel depressed. Feedback such as this is consistent with reports from acupuncture clinical practice when patients are treated for depression.

The results and patient feedback in this study indicate there is a relationship between liver qi stagnation and depression. Future clinical trials, which are designed to test if acupuncture can treat depression, need to be designed to offer better clarification. Ideally, these studies would need to address two areas of contention. Firstly, any doubt over dispersing liver qi as being an effective principal of treatment to reduce depression needs to be resolved. This could be achieved by comparing an acupoint prescription based on dispersing liver qi to a non-specific acupuncture prescription. As noted earlier, this design would have implicit difficulties. Alternately it may be possible to include a third treatment group which would receive a principle of treatment which is designed to *calm shen*. This would clarify if the improvement in depression was due to the experience of the acupuncture or if it was result of applying a principal of treatment. Secondly, if confirmation can be found to suggest dispersing liver qi is the most effective principle of treatment to reduce depression, then the most effective combination of liver qi dispersing points which best treat depression would need to be determined.

5.18 Conclusion

From the results in this study, it can be concluded that acupuncture is an effective adjunct therapy to SSRI antidepressants in the treatment of depression in the liver qi stagnation population. It still remains to be fully determined if the principle of treatment of dispersing stagnant liver qi is the most effective

principle of treatment to reduce depression. Though, this study does provide evidence to suggest it is better to treat sufferers of depression (using acupuncture) than not to treat. Moreover, this study also serves to give evidence that acupuncture can also treat anxiety and improve the general mental health in sufferers of depression who also present with liver qi stagnation.

5.19 Recent Additions to the Literature

Over the seven years since the start of this study, there have been a number of publications which consider the use of acupuncture in the treatment of depression. Interestingly, this study introduced a number of changes in design which were not consistent with the literature at that time. Encouragingly, some of the improvements employed in this study are now appearing in the recent literature. This is especially the case in the use of acupuncture as an adjunct therapy to SSRI antidepressants.

5.19.1 Acupuncture's Effectiveness in Treating Depression

In the recent literature, two meta-analysis articles have been published, which consider the effectiveness of using acupuncture to treat depression. Wang (2008) focused on the results from eight clinical trials. The studies considered within this analysis used either or both the BDI and the Ham-D. This makes for an interesting comparison to the results from this study. The effectiveness of the collective acupuncture interventions was found to be statistically significant (p =0.02). This confirms the findings from this study. In contrast, Wang (2008) carried out extra analysis which found that the acupuncture interventions had a poor response rate (p = 0.25) and poor remission rate (p = 0.53). These outcomes are not consistent with the findings in this study. Wang (2008) however, did stress that the poor standards used in the studies included in the meta-analysis raised concerns over the validity of the findings. Smith (2010) updated the Smith (2004) systematic review by adding another 23 studies to the

analysis. She concluded there is still insufficient evidence to suggest that acupuncture is effective in the treatment of depression. She too stressed the results are limited by methodological flaws. These were reported to create frequent high-risk biases within the majority of the clinical trials. The inconsistency in outcomes between Wang (2008) and Smith (2010) are consistent with the mixed outcomes which were discussed earlier in the literature review. More importantly, both studies further highlighted the lack of sound design in clinical trials which have used acupuncture to treat depression.

Ernst (2011) conducted a "systematic review" of the systematic reviews. He concluded that the effectiveness of acupuncture in the treatment of depression still remains unproven. He reasoned that the main contributing factor which led to this conclusion was the high number of less than reliable studies from China. The Chinese studies were said to forgo important elements of design. He concluded that these methodological deficiencies cast doubt over their outcomes and in turn, the effectiveness of using acupuncture to treat depression is yet to be satisfactorily determined.

Schroer (2011) carried out a review and identified a range of areas of concern which frequently occurred within the literature. Recommendations were also offered to circumvent each problem. Four main areas of concern were identified. These were: the poor standards of reporting on the selection of subjects and the associated inclusion/exclusion criteria; poorly rationalised and implemented acupuncture protocols; therapeutically inappropriate acupuncture; and an over emphasis on short term outcomes. These shortcomings are characterised by flaws in both the application of scientific method as well as the CM model.

The opinions of Schroer (2011) are more aligned with the researchers in this study. Much of the critique in the literature is focused on improving designs to be better aligned with the scientific method. This approach however denies the CM model. As a result important elements of design are overlooked. Schroer (2011),

along with the researchers in this study, have stressed the importance of maintaining the integrity of both models. Only when both paradigms are satisfied, via the creation of a synthesis model, will outcomes be truly informative. Wu (2012) also shared this opinion and highlighted the implementation of poor acupuncture interventions which did not uphold the CM model. Wu (2012) was of the opinion that acupuncture is an effective mono-therapy in the treatment of depression. It was reasoned that it is only the absence of a well-designed study which is preventing this realisation.

5.19.2 The New Approach to Design

Despite seven years of further research, there continues to be problems associated with finding a consensus on an appropriate design for RCTs which use acupuncture to treat depression. The studies in China tend to favour maintaining the integrity of the CM model over sound scientific method. Conversely, the studies in the west continue to favour sound scientific method over the maintenance of the CM model. A synthesis model remains to be developed in which satisfactory application of the scientific model can be upheld whilst maintaining the integrity of the CM model.

Since the commencement of this study, there have been articles published which specifically addressed the problems associated with the inadequate designs in clinical trials which used acupuncture to treat depression. MacPherson (2007) made the first steps to address the issue. His suggestions acknowledged the idiosyncratic nature of acupuncture and CM. He suggested a series of recommendations which were designed to ensure acupuncture would be used in clinical trials in a manner more consistent with clinical practice. The essences of these suggestions were consistent with STRICTA recommendations from MacPherson (2001). Included was the need to base acupoint prescriptions on a principle of treatment stemming from a diagnosis

based on CM patterns of disharmony. This had already been identified in this study as a common design problem within the studies in the literature.

Schroer (2009) rightly pointed out that acupuncture is a complicated intervention which is difficult to standardise for a heterogeneous group of people. Schroer (2009) then contributed further steps in moving closer to solving this problem by suggesting the isolation of subgroups within the depressed population. Later Schroer (2012) further identified the need to design clinical trials which are more reflective of how acupuncture is used in a clinical setting. It was proposed that if a model could be established which better facilitates the CM model, then future studies will have greater validity.

The researchers in this study are of the opinion they have proposed a suitable solution. During the inception of this study, the researchers ensured the adherence to the STRICTA recommendations of MacPherson (2001). These recommendations helped maintain the integrity of the CM model. They did not however solve the problem surrounding the incompatibility with the scientific model. The researchers in this study preempted Schroer (2009) and also identified the need to isolate a subset of the depressed population in accordance with the CM model. This was done by focusing on those who presented with liver qi stagnation. Consistent with Schroer's (2009) theory, this enabled the researchers in this study to achieve new levels in standardisation when applying the CM model. This facilitated the standardisation of a pattern of disharmony (with the associated inclusion/exclusion criteria), a principle of treatment and an appropriate acupoint prescription. Through the isolation of a subset of the depressed population, the idiosyncratic nature of the CM model was able to be neutralised. In turn, the key demands of each paradigm were able to co-exist within the one model.

5.19.3 New Focus on Chinese Medicine Diagnosis

There has been a recent development in the literature which was aimed at improving the CM understanding of depression. MacPherson (2013) investigated various patterns of disharmony to find which could best account for depression. It was found the patterns of disharmony which were most closely associated with depression were, liver qi stagnation and spleen qi deficiency. These patterns respectively accounted for 66% and 34% of the subjects who participated in his study. This is consistent with the opinion held by the researchers in this study at the time of this study's inception. The researchers in this study considered the symptoms of the greatest number of depression sufferers, were best explained by liver qi stagnation.

The understanding of pattern differentiation is, however open to interpretation. Consequently there are a number of different schools of thought within the CM community. Given this, it is possible some of the subjects who were diagnosed by MacPherson as having spleen qi deficiency, may have been diagnosed differently by another CM practitioner. Liver qi stagnation and spleen qi deficiency share common symptoms. This is especially evident when liver qi stagnation is extended to include the pattern of "wood invading earth". To avoid confusion over patterns of disharmony, the researchers in this study utilised a manualisation process. MacPherson (2013) chose not to do this, nor did he report how he chose to define each pattern of disharmony. Although MacPherson (2013) did include wood invading earth under liver qi stagnation, differing schools of thought may have led to variance in the allocation of the diagnoses. What one considers being spleen qi deficiency, another may consider being liver qi stagnation. This is why it is important to clearly define patterns of disharmony. Best practice involves a manualisation process and as done in this study, with the outcomes serving as inclusion criteria. In addition, full reporting of the process and its outcomes needs to be provided so as to offer clarity and to assist future replication of the study.

5.19.4 Acupuncture as an Adjunct Treatment

In the recent literature, there has been an increase in the number of RCT's which have used acupuncture as an adjunct treatment to antidepressant medication. This was also considered by the researchers of this study to be an important component of design. The new studies in the literature also identified the need to update the antidepressant medication to better reflect the modern treatment of depression. As seen in this study, the antidepressant which is most commonly being included in the new studies, are SSRI's.

In the literature which was published before this study, the analysis focused on comparisons between the effectiveness of an antidepressant to that of an acupuncture intervention. Within the recent literature, Wang (2013) was the only study to take this approach. He designed his study so a direct comparison was made between electro-acupuncture to the SSRI paroxetine. The results from this study indicated the subjects in both groups experienced significant improvement according to a range of depression scales, though no statistically significant differences were found between the two groups. The other studies in the recent literature have moved away from analysing the difference in effectiveness between acupuncture and antidepressant medication. Instead, these studies have chosen to combine the two treatments together and measure the benefits of using acupuncture as an adjunct treatment to the antidepressant medication. A consensus has emerged suggesting this provided more meaningful outcomes since it was more reflective of the clinical setting. This opinion then led a number of studies to adopt this design. This trend serves to validate the design used in this study.

Zhang (2010) carried out a review and meta-analysis which considered, amongst others, elements of design and the use of acupuncture as an adjunct treatment to antidepressant medication. He suggested there is sufficient evidence to consider acupuncture as being an effective monotherapy in the

treatment of MDD. But, he highlighted the presence of continuing doubt as to if acupuncture can further increase effectiveness beyond that expected from an antidepressant when it is taken alone. In response, he concluded that it is still yet to be determined if acupuncture is an effective adjunct therapy to antidepressant medication.

The new studies in the literature which used a design most similar to this study, were Zhang (2007), Qu (2013) and Duan. Both of these studies compared a treatment group, which received acupuncture plus an SSRI, to a control group which only received the SSRI. The results of Zhang (2007) and Qu (2013) were consistent with those found on this study. These indicated that acupuncture is an effective adjunct treatment to SSRI medication ($p < 0.05$ and $p = 0.004$ respectively). Interestingly, both of these studies chose to specifically administer paroxetine instead of a range of SSRI's. Nevertheless, given the outcomes from these two studies support those of this study, the three studies can combine to form an emerging consensus on the effectiveness of using acupuncture as an adjunct treatment to SSRI antidepressants.

In contrast, Duan (2009) used electro-acupuncture as an adjunct to the SSRI fluoxetine. His results indicated no significant difference in the change of scores when the electro-acupuncture was administered in addition to the fluoxetine compared to when the fluoxetine was taken alone ($p > 0.05$). Despite this statistically insignificant outcome, there are other interesting observations within his study. Those who received the acupuncture were found to experience more rapid improvement in their depression scores. The improvement in scores they experienced over a two week period required an additional two weeks to be matched by those who only received the fluoxetine alone. In effect it took those in the wait group twice as long to experience the same improvement in scores. This has important clinical implications as it may suggest that acupuncture can expedite the therapeutic benefits of SSRI's. This is important from a clinical perspective since the combination of both treatments may limit the time

depression sufferers experience potentially harmful symptoms such as suicide ideation.

The results from Zhang (2007), Duan (2009) and Qu (2013) collectively offer encouraging evidence for the use of acupuncture as an adjunct treatment to SSRI's. The outcomes from these studies suggest that depression sufferers would experience improved treatment if they were to receive acupuncture in conjunction with their SSRI antidepressants. Despite this, the design in each of these studies could have been improved. None of the studies made an attempt to define the depressed population with respect to the CM model. This would have better facilitated the isolation of a subset within the depressed population. This in turn may have assisted the implementation of a more suitable acupuncture intervention. Moreover, it would have offered a better understanding of the clinical implications.

The studies of Zhang (2009) and Zhang (2012) chose to use different design than used in this study. Despite that both studies used acupuncture as an adjunct treatment to fluoxetine, the inclusion of alternate elements of design rendered their outcome less informative. Zhang (2009) compared active acupuncture to sham acupuncture whilst all subjects received fluoxetine. The results indicated there were no significant differences between the two groups (p = 0.07). It must be stressed the design used in this study did not facilitate a direct comparison between the active and sham acupuncture. The subjects who were administered the active acupuncture, received a low dose of fluoxetine. In contrast, those who were administered the sham acupuncture received a clinical dose of fluoxetine. This renders the outcome less informative since there were two variables involved. Consequently, it does not serve to provide a clear comparison between the active and the sham acupuncture. Zhang (2012) used a unique approach where non-invasive electro-acupuncture were compared to dense cranial electro-acupuncture stimulation (DCEAS) whist all subjects were taking fluoxetine. The results indicated the use of the DCEAS combined with the

fluoxetine was more effective than the non-invasive electro-acupuncture with the fluoxetine.

The designs used in both Zhang (2009) and Zhang (2012) offered limited understanding of the effectiveness of using acupuncture as an adjunct to SSRI's. Each study chose to focus on a specific aspect of treatment. In effect, the outcome of Zhang (2009) was more indicative of differences in the dosage of fluoxetine. The outcomes in Zhang (2012) reflected more on different techniques of electro-acupuncture. Both studies did not clearly address the more fundamental question as to whether acupuncture is a valuable adjunct to medication. This line of enquiry was better addressed by the methodological design used in the studies by Zhang (2007), Duan (2009) and Qu (2013). All three of these studies used acupuncture as an adjunct to an SSRI and each made comparisons to when the SSRI was taken alone. These designs combined the use of acupuncture as an adjunct treatment with the use of a wait list control group and were the same core elements of design employed by the researchers in this study.

5.19.5 Upholding the Chinese Medicine Model

The study within the literature, which used a design most similar to this study, was Fu (2009). His study also chose to isolate a subset of the depressed population according to the CM model. Interestingly, as done in this study, Fu (2009) also decided to focus on those who presented with symptoms consistent with liver qi stagnation. A set acupuncture prescription was also administered as an adjunct treatment to SSRI's. Comparisons were then made between those who received the acupuncture as an adjunct to the medication, to those who took part in a wait group and only received the medication. There are however differences between the designs of this study and Fu (2009). The main difference is that Fu (2009) did not define his subjects in accordance with MDD. Instead, he chose to focus on depressive neurosis. Secondly, despite using a

principle of treatment to disperse liver qi stagnation, Fu (2009) used a different acupoint prescription. Lastly, Fu (2009) also used an SSRI, but in contrast to this study, focused on one specific SSRI - prozac. In addition to using a wait list control group Fu (2009) also used a non-specific acupuncture group as a second method of control. The outcomes of Fu (2009) were consistent with those found in this study. The active acupuncture was also significantly more effective at reducing depression scores than the medication alone ($p < 0.05$). Since the methodological design of Fu (2009) and this study are very similar, it is hoped this may be the start of achieving a consensus in relation to a suitable methodological design for studies which use acupuncture to treat depression.

5.19.6 Summation on Recent Additions to the Literature

The better designed studies in the recent literature have used similar elements of design as used in this study. There may not yet be an acknowledged consensus on the most preferred design for clinical trials which use acupuncture to treat depression, but trends are starting to emerge. The most evident trend is the use of acupuncture as an adjunct treatment to SSRI antidepressants. This has in turn facilitated increased value in the use of a wait list control group as a method of control. This may have inherent limitations, but it better reflects the clinical setting and it is the most compatible with the CM model. Improved steps have also begun to be made to better isolate a subset of the depressed population in accordance to the CM model. When the studies in the literature did use more suitable methodological designs, the results did indicate there is encouraging evidence to suggest that acupuncture may be an effective adjunct treatment for depression. Nevertheless, the studies in the literature are yet to match the rigour employed by the researchers in this study.

Chapter 6: Conclusion

This study has offered an interesting insight into the effectiveness of using acupuncture as an adjunct to SSRI's in the treatment of depression. This concluding chapter will provide a summation of the main findings and will highlight the key inferences and implications. These points of interest will be discussed in turn by addressing the outcomes related to each of the psychometric measures followed by the main elements of design. Finally recommendations for further study will be suggested.

6.1   The Main Trends

The general trends in the results from this study indicated acupuncture is an effective adjunct therapy to SSRI's in the treatment of depression in those who are identified as having the CM diagnostic pattern of liver qi stagnation. Across all depression scales (i.e. the BDI, the Ham-D and the SCL 90 depression subscale) subjects demonstrated significant reductions in outcome scores over the acupuncture intervention phases. This trend was evident in the treatment group and when those in the wait list control group received the acupuncture intervention at the conclusion of the wait period. No significant change in the scores from any of the depression scales occurred in the wait period. This suggests that there was no improvement based on time alone. The improvement in scores during the intervention phases of the study may therefore be attributable to the acupuncture treatment.

6.2   Beck Depression Inventory

The analysis on the BDI indicated statistically significant improvement in scores over the intervention phases of the study for those in the treatment group (α = 0.05, p = 0.000) and for those in the control group (α = 0.05, p = 0.000). In

contrast, when those in the wait list control group received their SSRI's over the wait period, there was no significant change in scores ($α$ = 0.05, p = 0.371). Of particular interest are the average changes in scores. Average scores of the BDI are classified within set parameters to serve as an indicator on the severity of depression. Average scores for those in the treatment group were reduced from a "moderate" classification down to a "minimal" classification, bypassing the classification of "mild". A similar pattern of average change in scores was observed when the wait list control group received the acupuncture intervention. This offers an interesting insight as it reflects on the perceived change in severity of the depression as it would be experienced by the subjects. This simulates a clinical setting where in the absence of psychometric measures, the patients might gauge their improvement using similar classifications of minimal, mild, moderate or severe. The improvement from a rating of moderate down to a minimal would suggest patients in a clinical setting would most likely consider the acupuncture to have been effective in treating their depression. This identified a possible additional area of study which was not covered in this trial i.e. the credibility of the treatment or the nature of the patient practitioner relationship and how that affects clinical outcomes.

6.3   Hamilton Rating Scale for Depressive Illness

The use of the Ham-D in addition to the BDI offered a more detailed understanding of the effectiveness of using acupuncture to treat depression. The outcomes from the main analysis using the Ham-D showed that acupuncture had a large statistical effect over the intervention phases of the study ($α$ = 0.05, p = 0.000). In contrast when those in the wait list control group received their SSRI's no change in scores was evident ($α$ = 0.05, p = 0.313). As the case with the BDI analysis, these outcomes suggest the acupuncture intervention may have been effective in reducing the severity of depression.

6.4   SCL 90 – Depression Subscale

Consistent with the results from the BDI and the Ham-D, the SCL 90 depression subscale also indicated the acupuncture intervention significantly improved depression outcome scores ($α$ = 0.05, p = 0.000). Similar to the other depression measures, no significant change was observed when those in the wait list control group received their SSRI's. A point of interest with the SCL 90 is that it measures the immediate perception of depression, in other words, how the subjects felt at the exact time of the assessment. This is in contrast to the BDI and the Ham-D which both measure short to medium term trends. The results however suggested all depression measures identified equally significant improvements in scores. This was unexpected as empirical observations suggest patients' perception of their state of depression normally fluctuate. Future studies in this area may consider the use of weekly assessments of at least one measure such as the SCL 90 to ascertain if there is any objective evidence of variance in scores which could reflect on the different time horizons considered for each measure.

6.5   Follow Up Analysis

Compared to studies in the literature, this study was able to offer greater insight into the long-term benefits of an acupuncture treatment for depression. Few studies in the literature included a follow up analysis. Those which did incorporate a follow up period, were not very informative, due to the wide range of time lapses between assessments within and across the studies. This study took a more structured approach by comparing scores at the conclusion of the follow up period to those at the conclusion of the acupuncture intervention as well as to those preceding the acupuncture intervention. This lead to two conclusions within the follow up analysis. Firstly, that subjects were able to maintain their improved outcome scores over the follow up period, and secondly,

they were still significantly less depressed at the conclusion of the follow up period than they were before they started the intervention.

The outcomes from the follow up analysis have important clinical implications. They demonstrate that a course of acupuncture may significantly reduce depression for patients on SSRI's with liver qi stagnation and that such patients may be able to maintain that effect for at least a period of up to eight weeks. This encouraging outcome highlights the need for further studies which incorporate a more meaningful follow up analyses. The outcomes from such studies would offer more clinically relevant information such as determining how long the effect of the treatment can be maintained. This could reflect on how frequently a course of acupuncture would need to be administered to maintain improvement in depression. Information of this nature would be valuable from a costing and policy perspective.

6.6   Sex Differences

This was the first acupuncture for depression study to stratify the data according to sex. The outcomes for each sex mirrored those when both sexes were considered together. Both sexes demonstrated significant improvement in depression outcome scores when they received the acupuncture intervention. In contrast both sexes experienced no change in the level of depression scores over the wait period. Both sexes also experienced no significant change in scores over the follow up period and remained significantly less depressed than before the intervention. These trends were confirmed via analysis of variance. According to all three depression scales, both sexes experienced the same rate of change across all phases of the study.

Given this is the first study to offer comparisons between the sexes, further research is needed. It is interesting to note that from a diagnostic perspective the sexes have a tendency to experience depression differently. Men are less

likely to report being depressed but will identify the physical symptoms associated with depression such as loss of weight, tiredness and irritability/anger (Beyond Blue, 2014). When it comes to the outcome of treatment however, there appears to be uniformity across the sexes in response to the treatment.

6.7 Intention To Treat Analysis

The purpose of an Intention to Treat (ITT) analysis is to better simulate how effective a treatment will be in a clinical setting. ITTs are undertaken to help avoid misleading artifacts that can arise through such factors and non-random attrition of subjects. The ITT analysis is based on the initial treatment assignment and includes analysis of all subjects assigned regardless of any deviations subjects may make from the protocol. Missing data is accounted for my using the last observation carried forward (LOCF) method, which retains the last available measurement for a subject at a time point prior to withdrawal and retains this in the analysis.

The ITT results in this study are encouraging since they replicate those from the standard analysis. This was also the case for the follow up analysis. In other words there is a reduced possibility of a type I error which if present would indicates a false positive result. This suggests that the positive improvements in depression noted by the subjects in this study were real and attributable to the acupuncture. Few studies in the literature included an ITT analysis and hence these study could be reporting false positive results. Future studies should either include an ITT analysis or otherwise report how they controlled for type I errors in their statistical analysis.

6.8 DSM-IV Diagnostic Criteria

The general trends in the analysis of diagnostic status of subjects followed similar patterns as seen in the analysis based on the scores from the depression scales. When all subjects were considered together over the intervention phases of the study, there was a remission rate of 87.3%. This means that 87.3% who received the acupuncture intervention no longer fulfilled the DSM-IV inclusion criteria for MDD at the conclusion of the treatment. This was the case for both females and males. Consistent with the previous analysis, the diagnostic status remained unchanged over the wait period when both sexes were considered together. There were however differences between the sexes. No females went into remission over the wait period however 18.75% of males did. This is most likely to be attributable to an anomaly caused by the smaller sample sizes used in the sex comparisons. It is however interesting to note this in the context of the earlier comment on differences relating to how males report depression and requires further study.

These outcomes indicate that the acupuncture intervention may have been effective in helping sufferers of depression achieve remission from MDD. If acupuncture is used in conjunction with SSRI's for subjects with liver qi stagnation, then this may lead to a reduction in the incidence of MDD. This is a key clinical implication of this study.

6.9 Design

This study employed a synthesis model of best fit which best upheld the rigour of scientific method whilst maintaining the integrity of the CM model. This approach was the key feature which differentiated this study from those in the literature. The main elements of design which differentiate this study include: the use of a wait list control group; acupuncture as an adjunct to SSRI's; and more

rigorous inclusion criteria which defined subjects from both a psychological and CM perspective.

6.10 Wait List Control

According to all three depression measures, the subjects in the wait list control group experienced no significant change in their levels of depression when they only received the SSRI antidepressants over the wait period. This outcome was not expected. Given subjects were prescribed and taking their SSRI over the wait period it was expected they would most likely experience some change in the severity of their depression. Hence it was thought the use of the SSRI's over the wait period, may create a positive bias. This however was not the case. The scores on all three depression measures over the wait period remained almost unchanged. The extent of consistency in the scores between the pre and post assessments of the wait period was of particular interest.

The researchers also noted that the majority of subjects had reported experiencing long-term depression. This was often accompanied by a clarification that they were interested in finding a new treatment since nothing they had tried seemed to be effective (including their SSRI medication). This situation was in effect created by the processes of recruitment used in this study. It stands to reason that those who were taking SSRI's and who were experiencing improvements in their mood, would not feel the need to seek out a different treatment. In turn, those who were no longer experiencing a benefit from their SSRI's would feel more compelled to find a new treatment and join the study. This effect may have been further exacerbated by the majority of recruitment for this study being the result of media coverage. The media reports were sensational by nature and suggested hope for those who have been unable to find an effective treatment. This may have contributed to the reason why the wait list control group results did not change and the levels of depression remained stable over the wait period.

6.11 Acupuncture as an Adjunct

The outcomes from this study suggest that acupuncture may be an effective adjunct therapy to SSRI antidepressant medication and more beneficial than taking SSRI's alone. This suggests that acupuncture has a potential to alleviate drug dependency and possibly help reduce the risk of the side effects associated with SSRI's. This has important clinical implications. SSRI's are considered to be a "first port of call" medication in the treatment of depression (Parker, 2006). Consequently, at any given time, there are a large number of sufferers of depression who are prescribed and taking SSRI's and acupuncture may be of benefit to a significant proportion of the community.

Further research is needed on the effectiveness of using acupuncture as an adjunct to SSRI medication. The majority of the studies in the literature administered acupuncture as a stand-alone treatment. Moreover, the studies which did use acupuncture as an adjunct therapy, used a range of different antidepressant medications. It is important to acknowledge that when acupuncture is combined with an SSRI it is a different treatment than when it is used as an adjunct to another antidepressant medication. It is yet to be determined if acupuncture is equally effective as an adjunct treatment irrespective of which antidepressant is used in combination. The recent trends in the literature have seen an increased use of acupuncture as an adjunct to antidepressants. Given the contemporary nature of these studies, SSRI's are being included. This new research will help define knowledge in this area, and it validates the decision of the researchers in this study to include subjects taking SSRI's.

6.12 Chinese Medicine Diagnosis

From a diagnostic perspective, the studies in the literature only used the diagnosis of MDD as inclusion criteria. This meant that depression sufferers

were treated as a heterogeneous group irrespective of how CM accounted for their depression. The researchers in this study included extra inclusion criteria in accordance with CM differential pattern diagnosis. In addition to satisfying the criteria for MDD, subjects needed to present with liver qi stagnation (or liver qi stagnation with wood invading earth or liver qi stagnation with heat). The signs and symptoms most commonly associated with these patterns were identified via a process of manualisation. Compared to the studies in the literature, this more detailed approach better isolated a specific subset of the depressed population. The researchers considered this to be an integral part of the design since it enabled a more thorough approach which satisfied both the pragmatic issues of CM and the scientific rigour required of clinical trials.

The more rigorous approach of restricting inclusion of subjects to a more defined diagnosis may explain why the acupuncture intervention used in this study was found to be more effective than those in the literature. The advantage of this approach is that it facilitated a correct matching of a standardised acupoint prescription to a specific pattern of disharmony. Moreover, it better reflects how acupuncture is used in a clinical setting.

6.13 Liver Qi Stagnation

Given the outcomes in this study there appears to be a close relationship between depression and liver qi stagnation. Moreover, this relationship has been shown to possibly extend further to include anxiety and other dimensions of mental heath included in the SCL 90. Further research is needed to better explore the relationship between CM patterns of disharmony and depression. It remains to be tested if an alternate pattern of disharmony and a matching acupoint prescription would treat depression as effectively. Focusing on other patterns of disharmony associated with depression may also offer important clinical implications. It would help create a more detailed understanding of depression from a CM perspective to be established. In turn, this may lead to

acupuncture being used to treat a broader range of depression sufferers and/or treat sufferers of other mental health conditions.

6.14 Anxiety

Anxiety outcome scores followed the same trends as the depression analysis. There were significant reductions in scores over the intervention phase of this study. This was the case for those in the treatment group and in the wait list control group. Analysis of the scores for the wait list control group remained unchanged over the wait period. The clinical significance of this outcome is that acupuncture may be an effective treatment for anxiety in those who also suffer from depression. It may also offer insight into the broader implication of liver qi stagnation in mental health.

More specifically, these outcomes indicated the subjects experienced less intense feelings of anxiety, less frequently. An interesting observation was that the greatest improvement was witnessed in the first four weeks of the intervention. During this time, treatments were administered twice a week. This may suggest that to treat anxiety effectively with acupuncture, more frequent treatment is required. This remains an area for further study.

Both sexes demonstrated mostly similar trends in their anxiety outcome scores. Unlike the females, the males did however continue to demonstrate significant improvement during the second four weeks of treatment. It is also interesting to note there were variations between the sexes which occurred within the intervention period. These differences however evened out at the conclusion of the full eight week period. This may suggest the discrepancies are due to differences in sensitivity to the time horizon of the mid treatment assessment.

6.15 SCL-90-R Subscales

The trends on the SCL 90 mostly followed the same patterns as the depression and anxiety analysis. When both sexes were considered together, statistically significant improvements were experienced on the global scale and all the respective subscales over the intervention phases. On analysis of the individual subscales however, there was a minor difference between the sexes. Unlike the males, the females did not exhibit improvement on the phobic anxiety and the paranoid ideation subscales. It is interesting to note this was only the case for the females in the treatment group. When the females in the wait list control group received the intervention, they showed the same trends as the males.

Whilst it is difficult to account for these subscale discrepancies between sexes, collectively, these outcomes suggest that acupuncture may be an effective treatment in improving a wide range of psychological dimensions in those who suffer from depression. This observation may have important clinical implications. It may offer insight into possible psychological dimensions which may have a comorbidity or association with depression. From a CM perspective, it may offer an interesting insight into the symptoms associated with liver qi stagnation and/or the benefits of dispersing liver qi.

The follow up analysis using the SCL 90 showed some interesting outcomes. On the global scale and the obsessive-compulsive subscale, males continued to demonstrate significant improvement over the follow up period. According to the analysis of variance, this enabled them to improve at a greater rate than females. Despite this, both sexes showed the same rate of improvement when scores at the end of the follow up period were compared to those before the intervention. This is similar to the outcomes on the hostility scale over the follow up period. In this instance the scores for the females significantly increased over the follow up period. This lead to a different rate of change between the sexes

over the follow up period but again no difference was found over the entirety of the study.

Another point of interest found over the entirety of the study was on the phobic anxiety subscale. Here the males demonstrated greater improvement than females yet no other differences were found when the shorter time horizons were considered. As noted earlier, it is difficult to account for why these differences between the sexes occurred. This is the first acupuncture study for depression which has used the SCL 90. Further research is needed to test if these outcomes are valid.

6.16 Recommendations for Further Study

There has been a recent increase in the amount of RCT's which use acupuncture to treat depression. There however still remain many areas for further research. Some areas of interest have received very little attention in the literature.

A greater understanding is needed in regards to the long-term benefits of using acupuncture to treat depression. This needs to be done through further studies which include a more detailed follow up analysis. Varying lengths of follow up periods need to be tested to ascertain the length of time the benefit of the acupuncture intervention can be maintained. This was also relevant to any further studies that would test anxiety as a comorbidity of depression or as a standalone condition. Evidence from this study suggested the frequency of treatment may affect anxiety and this could be further tested.

This study was the first to make comparisons based on sex. Further studies are needed to reflect on these outcomes. Similarly, more research in needed which considers a broader understanding of depression sufferers. This study increased this understanding by including both an anxiety measure and a general mental

health measure. More research could be done using further measures to expand on the understanding of depression and it's comorbidities. This could be extended to determine the effect of acupuncture when used in conjunction with drugs other than SSRI's.

The most important area for further research involves the method of diagnosis and treatment in CM. The clear outcomes in this study serve to support the need to isolate a subset of the depressed population according to CM diagnostic pattern differentiation. This facilitated the formation of an acupoint prescription which best treated the pattern of disharmony in question (liver qi stagnation). This in turn has enabled the researchers to add further evidence of the link between liver qi stagnation and depression. Additional research on points of interest can move in three different directions.

Firstly, other patterns of disharmony can be used to isolate other subsets of the depressed population to test if they too can be treated as effectively. For example some schools of thought consider lung qi deficiency to be associated with depression. A corresponding inclusion criterion would include subjects who satisfy DSM-IV criteria as well as criteria for lung qi deficiency.

Secondly, the theoretical assumption of the involvement of pattern differentiation could be tested. The effectiveness of the treatments in this study may well not be due to the focus on liver qi stagnation. It may have been the acupuncture process per se which facilitated the improvement in depression outcome scores. As discussed earlier, it is difficult to ascertain acupoint specificity in an RCT due to the inherent problems which are associated with including non-specific acupuncture. Another way around this may be to administer an acupoint prescription which is designed to calm shen. An ensuing design would have the same inclusion criteria as this study, but subjects would be assigned to either a treatment group where liver qi is dispersed or a treatment group which receives an intervention designed to calm shen with no liver qi clearing acupoints.

The third option for further research would involve determining which acupoints best disperse liver qi stagnation. This line of testing rests on the assumption that liver qi stagnation is the pattern of disharmony most associated with depression and that point specificity and related functional activity is determined. A study of this nature would use a range of depression measures to ascertain the effectiveness of differing acupoint prescriptions.

6.17 Conclusion

This study used a randomised, single blind, repeated measures clinical trial design to demonstrate that acupuncture can be used as an effective adjunct treatment to SSRI antidepressant medication in the treatment of depression in those who are experiencing liver qi stagnation. The effect of the acupuncture intervention was maintained over an eight week follow up period. This was the first study to stratify outcome data based on sex and found similar trends for both sexes.

These results were achieved by employing a synthesis model of best fit which best upheld the rigour of scientific method whilst maintaining the integrity of the CM model. This unique approach involved the isolation of a subset of the depressed population according to both a psychological/psychiatric and CM perspective which allowed for the standardisation of the acupuncture intervention. The researchers in this study believe that it was the merits of this particular aspect of the clinical trial design which accounted for the large effect size of the acupuncture.

Appendices

APPENDIX A: DSM IV Structured Interview for Major Depressive Disorder

INCLUSION CRITERIA

PART A

Five or more of the following symptoms must be present. Either item 1 or 2 above the line must be present.

☐   Depressed mood for most of the day. This can be either a subjective report or an observation made by another (eg looking sad)

☐   Loss of interest or pleasure in usual activities.

☐   Appetite disturbance or weight change. I.e. significant weight loss when not dieting; weight gain of more than a 5% increase in a month; decrease/increase in appetite nearly every day.

☐   Sleep disturbance. Insomnia or hypersomnia nearly everyday.

☐   Psychomotor agitation or retardation nearly every day. This must not just be a subjective feelings of restless or being slowed down. It must be observable by others.

☐   Fatigue or loss of energy nearly every day.

☐   Feelings of worthlessness or excessive or inappropriate guilt nearly every day. These feelings may be delusional and not merely self-reproach or guilt about feeling sick.

☐   Difficulty concentrating nearly every day. This can be either subjectively reported or observed by others. E.g. diminished ability to think or concentrate and indecisiveness.

☐   Suicidal ideation. Recurrent thoughts of death; recurrent suicidal ideation without a specific plan; suicide attempt; a specific plan for committing suicide.

PART B

Both of the following symptoms must be present.

☐   Clinically significant distress.

☐   Impairment in social, occupational, or other important areas of functioning.

EXCLUSION CRITERIA

PART A

Other explanation. Any of the following circumstances excludes Major Depressive Disorder as a diagnosis.

☐   Drug side effect. Symptoms are due to a direct physiological effect of a substance.

☐   General medical condition. Can the symptoms be explained by an underlying general medical condition?

☐   Bereavement. Symptoms of bereavement can last longer than 2 months; have marked functional impairment; morbid preoccupation with worthlessness; suicide ideation; or psychomotor retardation.

PART B

Manic Episode / Hypomanic Episode

Either one of the items above the line must be present, along with 3 or more of the items below the line.

- ☐ A distinct period of abnormal and persistently elevated, expansive, or irritable mood, lasting more than 1 week. (Manic)
- ☐ A distinct period of abnormal and persistently elevated, expansive, or irritable mood, lasting throughout at least 4 days. (Hypomanic)

- ☐ Inflated self esteem or grandiosity.
- ☐ Decreased need for sleep (feels rested after 3 hours sleep).
- ☐ Talkative. More than usual, or pressure to talk.
- ☐ Flight of ideas or subjective experience that thoughts are racing.
- ☐ Distractibility. Attention easily drawn to unimportant or irrelevant external stimuli.
- ☐ Increased goal activity (either work, social, sexual or psychomotor agitation).
- ☐ Excessive involvement in pleasurable activities that may have a high potential for painful consequences.

PART C

Schizophrenia, Schizoaffective Disorder, and Schizopheniform Disorder.
Two or more of the following symptoms must be present for a significant
portion of time over a 1 month period.

- ☐   Delusions.
- ☐   Hallucinations.
- ☐   Disorganised speech.
- ☐   Grossly disorganized or catatonic behaviour.
- ☐   Negative symptoms (i.e. affective flattening, alogia, or avolition).

PART D

Delusional Disorder

All 3 of the following need to be present.
- ☐   The above criteria for Schizophrenia has never been met.
- ☐   Nonbizarre delusions (i.e. involving situations that occur in real life).
- ☐   Apart from the impact of the delusion(s) or its ramifications, functioning is not markedly impaired and behaviour is not obviously odd or bizarre.

PART E

Psychotic Disorder Not Otherwise Specified).

Any of the following.
- ☐   Postpartum psychosis.
- ☐   Psychotic symptoms that have lasted less than 1 month but have not yet remitted.
- ☐   Persistent auditory hallucinations in the absence of any other features.

APPENDIX B : Liver Qi Stagnation Manualisation

Legend for Manualisation Tables	
Table Reference number	Source
1	Maciocia (1989)
2	Maciocia (2004)
3	Deng (1999)
4	Deng (Depression subsection) (1999)
5	Chen (2004)
6	Wiseman (1996)
7	Wiseman (1998)
8	Ross (1995)
9	Shuai (1992)
10	Shandong (1989)
11	Zmiewski (1989)
12	Zhang (1992)
R = repeat from same author	

LIVER QI STAGNATION	Reference (refer to legend for source)												Total
	1	2	3	4	5	6	7	8	9	10	11	12	Total
Depression	X		X	R	X		X	X	X	X	X	X	9
Hypochondriac / chest distension	X	R	X	R	X	R	X		X	X	X	X	8
Painful breasts	X	R	X	R	X	R	X		X	X		X	7
Abdominal/Epigastric distension	X	R		X	X	R	X		X			X	6
Sighing	X		X		X		X		X			X	6
Irregular Periods	X	R	X	R	X				X	X		X	6
Irritable		X			X		X	X	X	X			6
Plum Stone	X	R			X				X			X	4
Move Lump Hypo/abdo				X					X			X	3
Poor appetite / Anorexia	X						X			X			3
Goiter					X				X			X	3
Moodiness	X							X					2
Sour regurgitation	X									X			2
Belching	X					R	X						2
Diarrhea	X						X						2
Dysmenorrhea / PMT	X	R	X	R									2
Pulsating epigastrium	X												1
Hiccup	X												1
Melancholy	X	R											1
Nausea	X												1
Vomiting	X												1
Borborygmi	X												1
Diff Swallow	X												1
Flatulence						X							1

LIVER QI STAGNATION	Reference (refer to legend for source)								
	1	2	3	4	5	7	8	9	12
Tongue	normal	severe red sides				thin white coat	purplish red spots		thin white coat
Pulse	wiry	wiry	string	string	wiry	string	wiry	string	taut

LIVER ATTACKING STOMACH	Reference (refer to legend for source)						
Reference	2	3	5	6	7	10	Total
Hypochondriac / chest distension	X	X	X		X	X	5
Acid regurgitation	X	X	X	X	R	X	5
Abdominal/Epigastric distension	X	X		X	R	X	4
Irritability	X		X		X	X	4
Nausea	X		X	X		X	4
Vomiting	X		X	X		X	4
Sighing	X	X	X				3
Poor appetite / anorexia					X	X	2
Belching	X	X					2
Hiccup	X		X				2
Depression			X				1
Diarrhea				X	R		1
Frustration / Anger		X					1
Tiredness	limbs						1
Pulse	red sides	white coat					
Tongue	wiry	string	wiry		string	string	

244

Reference	Reference (refer to legend for source)								Total
LIVER ATTACKING SPLEEN	**1**	**2**	**3**	**5**	**6**	**7**	**10**	**11**	**Total**
Depression			X	X			X		3
Alt loose / firm stool	**X**	R					X		2
Diarrhea			X	X	**X**	R	X		4
Hypochondriac / chest distension			X	X		X	X		4
Abdominal/Epigastric distension	**X**	R	X	X	X		X		5
Irritability	X	R	X	X		X	X		5
Frustration / Anger			X						1
Flatulence	X	R					X		2
Tiredness	X	R	limbs						2
Sighing			X	X			X		3
Poor appetite / anorexia				X			X		2
Borborygmi							X		1
Nausea					X			X	2
Vomiting					X			X	2
Acid regurgitation			X		X			X	3
Headache						X			1
Bitter taste						X			1
Bloating						X			1
Pain in anus				X					1
Pulse	red sides / pale centre	red side					white coat		
Tongue	rhs-weak, lhs-wiry	rhs-weak, lhs-wiry	string	wiry		string	string		

LIVER QI STAGNATION WITH HEAT	
Reference	2
Depression	X
Irritability	X
Moodiness	X
Melancholy	X
Hypochondriac/epigastric distention	X
Chest Oppression	X
PMT	X
Irregular periods	X
Breast distention (PM)	X
Red Cheekbones	X
Muscular Weakness	X
Cramps	X
With/brittle nails	X
Dry Hair	X
Tongue	normal / rootless coat
Pulse	floating empty

APPENDIX C: Liver Qi Stagnation Check List

CHECKLIST OF SIGNS & SYMPTOMS

LIVER QI STAGNATION

- ☐ Hypochondriac stuffiness;
- ☐ Hypochondriac distension;
- ☐ Chest stuffiness;
- ☐ Chest distension;
- ☐ Sighing;
- ☐ Irregular menstruation;
- ☐ Painful breasts (before period);
- ☐ Irritability;
- ☐ Abdominal and/or epigastric distension;
- ☐ Frustration;
- ☐ Plumb stone throat or difficulty swallowing;
- ☐ Wiry pulse;

-

- ☐ Tiredness (but has energy when required);
- ☐ Dysmenorrhea;
- ☐ Pre-menstrual tension. ie headaches, irritability, abdominal pain;
- ☐ Constipation when stools are not dry;
- ☐ Speaking with a shouting voice;
- ☐ Headache aggravated by fatigue which stops with rest;
- ☐ Pain in arms and shoulders when walking;
- ☐ Borborygmi with a feeling of abdominal distension without loose stools;
- ☐ Flatulence;
- ☐ Green face;
- ☐ Pain experienced on palpation of Liver Mu point (Liver 14 - Qimen).

WOOD INVADING EARTH

☐Abdominal distension;

☐Irritability;

-

☐Alternating constipation and diarrhea;

☐Abdominal distension and pain;

☐Tiredness;

☐Greenish colour around the mouth;

☐Fullness of epigastrium or bloating;

☐Sour regurgitation;

☐Belching;

☐Nausea;

☐Vomiting;

☐Sour vomiting;

☐Sour taste in mouth.

LIVER QI STAGNATION WITH HEAT

☐Yellow tongue coat;

☐Bitter taste in mouth;

☐Desire to drink cold drinks;

☐Goats stools;

☐Headaches.

☐Blood shot eyes

APPENDIX D: Acupoint manualisation

Legend for Acupoint Manualisation Tables	
Table Reference number	Source
1	Maciocia (1989)
2	Maciocia (2004)
3	Deadman(1998)
4	Wiseman (1996)
5	Wiseman (1998)
6	Ross (1995)
7	Shuai (1992)
8	Zmiewski (1989)
R = repeat from same author	

LIVER QI STAGNATION									
Reference	1	2	3	4	5	6	7	8	Total
Liver 3 (Taichong)	X	R	X	X	R	X	X	X	6
Liver 14 (Qimen)	X	R	X	X	R	X	X	X	6
Pericardium 6 (Neiguan)	X	R			X	X	X	X	5
Urinary Bladder 18 (Ganshu)			X				X	X	3
Liver 1 (Dadun)					X	X		X	3
Gall Bladder 34 (Yanglingquan)	X	R		X	R			X	3
Spleen 6 (Sanyinjiao)					X		X		2
San Jiao 6 (Zhigou)	X	R		X					2
Conception Vessel 6 (Qihai)					X	X			2
Liver 2 (Xingjian)				X			X		2
Liver 13 (Zhangmen)	X	R	X						2
Large Intestine 4				X				X	2

(Hegu)								
Stomach 36 (Zusanli)				X				1
Conception Vessel 17 (Shanzhong)						X		1
Pericardium 1 (Tianchi)						X		1
Large Intestine 11 (Quchi)				X				1
Gall Bladder 41 (Zulinqi)			X					1

LIVER ATTACKING SPLEEN							
Reference	1	2	3	4	5	7	Total
Liver 3 (Taichong)	X	R	X	X	R	X	4
Liver 13 (Zhangmen)	X	R	X	X	R		3
Stomach 36 (Zusanli)	X	R			X	X	3
Pericardium 6 (Neiguan)		X		X	R	X	3
Liver 2 (Xingjian)				X		X	2
Liver 14 (Qimen)	X	R			X		2
Gall Bladder 34 (Yanglingquan)	X	R			X		2
Urinary Bladder 20 (Pishu)				X	R	X	2
Spleen 6 (Sanyinjiao)	X	R		X	R		2
Stomach 25 (Tianshu)		X				X	2
Conception Vessel 12 (Zhongwan)	X	R		X	R		2
Urinary Bladder 18 (Ganshu)					X		1
Urinary Bladder 21 (Weishu)				X			1
Spleen 3 (Taibai)					X		1
Spleen 4 (Gongsun)				X			1
Spleen 15 (Daheng)		X					1
Conception Vessel 6	X	R					1

(Qihai)							
San Jiao 6 (Zhigou)		X					1

LIVER ATTACKING STOMACH				
Reference	**2**	**4**	**5**	**Total**
Urinary Bladder 21 (Weishu)	X	X	R	2
Stomach 36 (Zusanli)	X		X	2
Liver 14 (Qimen)	X		X	2
Liver 13 (Zhangmen)	X	X		2
Gall Bladder 34 (Yanglingquan)	X		X	2
Urinary Bladder 20 (Pishu)		X		1
Urinary Bladder 18 (Ganshu)			X	1
Stomach 34 (Liangqiu)	X			1
Stomach 25 (Tainshu)			X	1
Stomach 21 (Chengman)	X			1
Stomach19 (Burong)	X			1
Spleen 6 (Sanyinjiao)		X	R	1
Spleen 4 (Gongsun)		X		1
Conception Vessel 12 (Zhongwan)		X	R	1
Conception Vessel 10 (Xiawan)	X			1
Pericardium 6 (Neiguan)		X	R	1
Liver 3 (Taichong)		X	R	1
Liver 2 (Xingjian)		X		1

LIVER QI STAGNATION WITH HEAT	
Reference	2
Liver 8 (Ququan)	X
Spleen 6 (Sanyinjiao)	X
Stomach 36 (Zusanli)	X
Conception Vessel 4 (Guanyuan)	X
Kidney 3 (Taixi)	X
Kidney 6 (Zhaohai)	X
Yu Yao (Extra)	X

APPENDIX E: Guided CM Diagnosis

- **HOT / COLD**   Feeling of fever or chills:

    Perception of being hot of cold:

    Cold hands / feet:

- **SWEAT**   Volume of sweat (profuse/mild):

    Time of day:

    Area of body:

- **THIRST**   Preference for hot or cold drinks:

- **RESPIR.**   Rhythm of respiration:

    Presence of phlegm:

    Wheeze / cough / asthma:

- **CARDIO.**   Blood pressure - high/low:

    Chest pain:

    Dizziness / fainting:

    Swollen ankles:

- **PAIN**   Location of pain:

    Muscle pain / joint pain / bone pain:

    Nature of pain (stabbing, moving, dull ache etc):

- **HEAD**   Headache: (location / nature):

    Eyes (dry / floaters / pain):

    Ears (tinnitus / pitch):

    Mouth (taste / teeth grinding):

    Nose / sinus:

- **SKIN**   Rashes:

Itching:

Eczema / psoriasis:

Acne / pimples:

- **GIT**   Bloating:

  Borborymus:

  Indigestion / reflux / vomiting:

  Flatulence:

  Pain / cramping:

- **APPITITE**   Degree of hunger (overly – absence of):

- **ENERGY**   Rating out of 10:

  Would prefer more or less energy:

  Consistency of energy levels:

- **SHEN**   Depression, anxiety, worry, stress etc

  Five phase interpretation (anger/frustration, joy,

  pensiveness, fear, sadness):

- **SLEEP**   Length of sleep:

  Falling asleep easily:

  Waking up in middle of night:

  Waking up early and unable to get back to sleep:

  Waking up refreshed or tired:

- **GYNAE**   Regularity of periods:

  Cycle length and duration of period:

  Colour:

  Volume:

  Characteristics (clots / breast distention etc):

  Dysmenorrhea (before, during or after):

- **URINE** Colour:

  Volume:

  Pain:

  Urgency / time:

- **STOOLS** Frequency:

  Consistency:

  Colour:

  Odor:

- **PULSE**

- **TONGUE** Colour of body:

  Colour of coat:

  Nature of coat (dry, greasy, papillae etc):

  Shape:

References

Allen, J. (2000). "Depression and Acupuncture: A Controlled Clinical Trial." Psychiatric Times **18**(3).

Allen, J., R. Schnyer, et al. (1998). "The Efficacy of Acupuncture in the Treatment of Major Depression in Women." Psychological Science **9**(5): 397-401.

Beck, A., R. Steer, et al. (1987). Beck Depression Inventory - Second Edition. San Antonio, Harcourt Brace & Company

Beyond Blue (2014) http://www.beyondblue.org.au/resources/for-me/men/depression-in-men

Blitzer, L., D. Atchison-Nevel, et al. (2004). "Using Acupuncture to Treat Major Depressive Disorder: A Pilot Investigation." Clinical Acupuncture and Oriental Medicine **4**: 144-7.

Chen, P. (2004). Diagnosis in Traditional Chinese Medicine. Taos, New Mexico, Complementary Medicine Press.

Deadman, P., M. Al-Khafaji, et al. (1998). A Manual of Acupuncture. East Sussex, Journal of Chinese Medicine Publications

Deng, T. (1999). Practical Diagnosis in Chinese Medicine. Edinburgh, Churchill Livingstone.

Derogatis, L. (1977). Symptom Checklist-90-R. Minneapolis, National Computer Systems, Inc.

Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. (2000). Washington, DC, American Psychiatric Association.

Duan, D., Y. Tu, et al. (2009). "Efficacy Evaluation for Depression with Somatic Symptoms Treated by Electroacupuncture Combined with Fluoxetine." Journal of Traditional Chinese Medicine **29**(3): 167-73.

Dunlop, B. and P. Davis (2008). "Combination Treatment with Benzodiazepines and SSRI's for Co-morbidity Anxiety and Depression: A Review." Primary Care Companion to the Journal of Clinical Psychiatry **10**(3): 222-8.

Ernst, E., M. Lee, et al. (2011). "Acupuncture for Depression?: A Systematic Review of Systematic Reviews." Evaluation and Health Professions **34**(4): 403-12.

Fundamentals of Traditional Chinese Medicine. (1992). Beijing, Foreign Languages Press.

"Facts and Figures About Mental Health and Mood Disorders." (2012). from http://www.blackdoginstitute.org.au/docs/Factsandfiguresaboutmentalhealthand mooddisorders.pdf.

Fundamentals of Chinese Medicine (Revised Edition). (1996). Brookline, Massachusetts, Paradigm Publications.

Fu, W., L. Fan, et al. (2009). "Depressive Neurosis Treated by Acupuncture for Regulating the Liver - A report of 176 Cases." Journal of Traditional Chinese Medicine 29(2): 83-6.

Gallagher, S., J. Allen, et al. (2001). "Six-Month Depression Relapse Rates Among Woman Treated with Acupuncture." Complementary Therapies in Medicine 9: 216-8.

HalBreich, U. (2008). "Systematic Reviews of Clinical Trials of Acupuncture as Treatment for Depression: How Systematic and Accurate Are They?" CNS Spectrums 13(4): 293-300.

Hamilton, M. (1967). "Development of a Rating Scale for Primary Depressive Illness." British Journal of Social Psychology 6: 278-296.

Han, C., X. Li, et al. (2004). "Clinical Study on Electro-Acupuncture Treatment For 30 Cases of Mental Depression." Journal of Chinese Medicine 24(3): 172-6.

Leo, R. and J. Ligot (2007). "A Systematic Review of Randomized Controlled Trials of Acupuncture in the Treatment of Depression." Journal of Affective Disorders 97(1-3): 13-22.

Lou, H., T. Jia, et al. (1990). "Electro-Acupuncture in the Treatment of Depressive Psychosis." International Journal of Clinical Acupuncture 1(1): 7-13.

Luo, H., Y. Jia, et al. (1985). "Electro-Acupuncture Versus Amitriptyline in the Treatment of Depressive States." Journal of Traditional Chinese Medicine 5(1): 3-8.

Luo, H., F. Meng, et al. (1998). "Clinical Research on the Therapeutic Benefits of the Electro-Acupuncture Treatment in Patients with Depression." Psychiatry and Clinical Neuroscience 52((Suppl.)): S338-40.

Maciocia, G. (1989). Foundations for Chinese Medicine: A Comprehensive Text for Acupuncturists and Herbalists. Edinburgh, Churchill Livingstone.

Maciocia, G. (2004). Diagnosis in Chinese Medicine: A Comprehensive Guide. Edinburgh, Churchill Livingstone.

MacPherson, H., B. Elliot, et al. (2013). "Acupuncture for Depression: Patterns of Diagnosis and Treatment within a Randomised Controlled Trial." Evidence Based Complementary Alternate Medicine.

MacPherson, H., B. Elliot, et al. (2013) Acupuncture for depression: patterns of diagnosis and treatment within a randomised controlled trial Evidence-Based Complementary and Alternative Medicine **Volume**, DOI:

MacPherson, H., S. Richmond, et al. (2012) Acupuncture, Counseling, and Usual Care for Depression (ACUDep): Study Protocol for a Randomized Controlled Trial. Trials **Volume**, DOI:

MacPherson, H. and S. Schroer (2007). "Acupuncture as a Complex Intervention for Depression: A Consensus Method to Develop a Standardised Treatment Protocol for a Randomised Controlled Trial." Complementary Therapies in Medicine **15**(2): 92-100.

MacPherson, H., L. Thorpe, et al. (2004). "Acupuncture for Depression: First Steps Towards a Clinical Evaluation." Journal of Alternative and Complementary Medicine **10**(6): 1083-91.

MacPherson, H., A. White, et al. (2001). "Standards for Reporting Interventions in Controlled Trials of Acupuncture: the STRICTA Recommendations." Complementary Therapies in Medicine **9**: 246-249.

Manber, R., R. Schnyer, et al. (2004). "Acupuncture: a Promising Treatment for Depression During Pregnancy." Journal of Affective Disorders **83**: 89-95.

Mukaino, Y., J. Park, et al. (2005). "The Effectiveness of Acupuncture for Depression - A Systematic Review of Randomised Controlled Trials." Acupuncture in Medicine **23**(2): 70-6.

Organisation, W. H. (2008). The global burden of disease: 2004 update, World Health Organisation

Parker, G. "The depressions: how to treat."   Retrieved (09/08/06), 2006, from http://www.australiandoctor.com.au/HTT/display.asp?content=D0001.

Qu, S., Y. Huang, et al. (2013). "A 6-Week Randomized Controlled Trial with 4-Week Follow-Up of Acupuncture Combined with Paroxetine in Patients with Major Depressive Disorder." Journal of Psychiatric Research **47**(6): 726-32.

Quah-Smith, J., W. Tang, et al. (2005). "Laser Acupuncture for Mild to Moderate Depression in a Primary Care Setting - A Randomised Controlled Trial." Acupuncture in Medicine **23**(3): 103-11.

Rogers, C. and C. Rogers (2006). Point Location and Point Dynamics Manual. Sydney, Australian College of Acupuncturists Ltd.

Roschke, J., C. Wolf, et al. (2000). "The Benefit from Whole Body Acupuncture in Major Depression." Journal of Affective Disorders **57**: 73-81.

Ross, J. (1995). Acupuncture Point Combinations: The Key to Clinical Success. Edinburgh, Churchill Livingston.

Schnyer, R. and J. Allen (2001). Acupuncture in the Treatment of Depression: A Manual for Practice and Research. Churchill Livingstone.

Schroer, S. and J. Adamson (2011). "Acupuncture for Depression: A Critique of the Evidence Base." CNS Neuroscience and Therapeutics **17**(5): 398-410.

Schroer, S., M. Kanaan, et al. (2012). "Acupuncture for Depression: Exploring Model Validity and the Related Issue of Credibility in the Context of Designing a Pragmatic Trial." CNS Neuroscience and Therapeutics **18**(4): 318-26.

Schroer, S., H. MacPherson, et al. (2009). "Designing an RCT of Acupuncture for Depression - Identifying Appropriate Patient Groups: A Qualitative Study." Family Practice **26**(3): 188-95.

Shao, N. (1989). Diagnostics of Traditional Chinese Medicine. Shandong, Shandong Science and Technology Press.

Smith, C. and P. Hay (2004). "Acupuncture for Depression." The Cochrane Database of Systematic Reviews(3).

Smith, C., P. Hay, et al. (2010). "Acupuncture for Depression." The Cochrane Database of Systematic Reviews **20**(1).

Spielberger, C. (1983). State-Triat Anxiety Inventory for Adults. Florida, Consulting Psychologists Press, Inc.

Statistics, A. B. o. (2009). National Survey of Mental Health and Wellbeing: Summary of Results. Canberra. **4326.0**.

Wang, H., H. Qi, et al. (2008). "Is Acupuncture Beneficial in Depression: A Meta-Analysis of 8 Randomized Controlled Trials?" Journal of Affective Disorders **111**(2-3): 125-34.

Wang, W., X. Lu, et al. (2013). "Effects of Electro-Acupuncture on Personality Traits in Depression: A Randomised Controlled Trial." <u>Chinese Journal of Integrative Medicine</u> **19**(10): 778-82.

Welfare, A. I. o. H. a. (2007). The Burden of Disease and Injury in Australia. Canberra.

Wiseman, N. and F. Ye (1998). <u>A Practical Dictionary of Chinese Medicine (2nd Ed.)</u>. Brookline, Massachusetts, Paradigm Publications.

Wu, J., A. Yeung, et al. (2012). "Acupuncture for Depression: A Review of Clinical Applications." <u>Canadian Journal of Psychiatry</u> **57**(7): 397-405.

Yang, X., X. Liu, et al. (1994). "Clinical Observation on Needling Extrachannel Points in Treating Mental Depression." <u>Journal of Traditional Chinese Medicine</u> **14**(1): 14-18.

Zhang, E., Ed. (1992). <u>Diagnostics of Traditional Chinese Medicine</u>. Shanghai, Publishing House of Shanghai College of Traditional Chinese Medicine.

Zhang, G., Z. Shi, et al. (2007). "Clinical Observation on Treatment of Depression by Electro-Acupuncture Combined with Paroxetine." <u>Chinese Journal of Integrated Medicine</u> **13**(3): 228-30.

Zhang, W., X. Yang, et al. (2009). "Combination of Acupuncture and Fluoxetine for Depression: A Randomized, Double-Blind, Sham-Controlled Trial." <u>Journal of Alternative and Complementary Medicine</u> **15**(8): 837-44.

Zhang, Z., H. Chen, et al. (2010). "The Effectiveness and Safety of Acupuncture Therapy in Depressive Disorders: Systematic Review and Meta-Analysis." <u>Journal of Affective Disorders</u> **124**(1-2): 9-21.

Zhang, Z., R. Ng, et al. (2012). "Dense Cranial Electroacupuncture Stimulation for Major Depressive Disorder - A Single Blind, Randomized, Controlled Study." <u>PLOS ONE</u> **7**(1).

Zmiewski, P. and R. Feit (1989). <u>Acumoxa Therapy - A Reference and Study Guide</u>. Brookline, Massachusetts, Paradigm Publications.