The Premenstrual Syndrome: Characterisation, Diagnosis and Treatment

Sarah Elizabeth Canning

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
Institute of Psychological Sciences

September 2008

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated overleaf. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.
Chapter 6 of this thesis was based on the jointly-authored publication:


The candidate confirms that her contribution was intellectual and that of the co-authors, Dr Louise Dye and Dr Mitch Waterman was primarily editorial.
I am truly grateful to my supervisors Dr Louise Dye, Dr Mitch Waterman and Mr. Nigel Simpson for their guidance and support throughout my PhD. I would like to thank Mitch for inspiring me and Nigel for providing encouragement. A special thank you goes to Louise for her untiring commitment and enthusiasm to the thesis and for providing support and encouragement throughout.

I would like to thank all of the women who gave up large amounts of their time to participate in this research. Without their dedication to a demanding protocol, this thesis would not exist. I would also like to thank Dr Julie Ayres and Biss Hartley at the Rosalind Bolton PMS clinic for the time and hard work that they put into setting up the clinical trial. The Rosalind Bolton Bequest must also be acknowledged for their provision of the funding which allowed me to do this.

This chapter of my life would not have been the same without the support of all my family and friends. Although I am thankful to so many people, a few deserve a special mention. I would like to thank my mum for constantly uplifting my spirits and for always remaining positive, my dad for reminding me that once in a while I should switch off and try to relax and my sister for always being there with an open ear, ready to listen. I am truly grateful that I could share the highs and lows of the last four years with my friend Laura, who started her PhD at the same time as me and who has remained by my side every step of the way. Finally, a huge thank you goes to Dave, for becoming increasingly domesticated, and for always being there with open arms and a smile, ready to make me laugh.
Premenstrual syndrome (PMS) is a prevalent condition. Key symptoms which promote treatment seeking are primarily psychological e.g. depression and aggression. Sufferers are often reluctant to take prescribed medication and often purchase dietary supplements and herbal remedies over the counter for which the evidence base with regards efficacy is limited. The primary aim of this thesis was to examine the effectiveness of St. John’s Wort (SJW) for PMS. Proposals that this herbal remedy could benefit PMS symptoms are based on evidence that SJW increases serotonin levels and suppresses pro-inflammatory cytokine production. Following a systematic review which demonstrated that although calcium and continuous vitamin B6 administration confer some benefit for premenstrual symptoms, the evidence for most dietary supplements and herbal remedies including SJW is conflicting or insufficient, a ten-cycle randomised double-blind, placebo-controlled, crossover trial was conducted. PMS sufferers (NIMH, 1983) were administered 900mg SJW/day (0.18% hypericin; 3.38% hyperforin) for two menstrual cycles (n=34). SJW was found to benefit physical and behavioural PMS symptoms, but did not significantly improve mood or pain symptoms. A comparison of various commonly used analytical strategies performed on the data highlighted the need for a consensus to be reached regarding the way in which researchers assess treatment efficacy. Hormone (FSH, LH, oestradiol, progesterone, prolactin and testosterone) and cytokine (IL-1β, IL-6, IL-8, IFN-γ and TNF-α) levels were assessed in women with and without PMS during the follicular and luteal phases, and were also studied in PMS sufferers taking SJW and placebo treatment. The hormone and cytokine profiles of PMS sufferers during SJW and placebo treatment did not differ. However, PMS sufferers exhibited significantly greater testosterone and cytokine (IL-6, IL-8 and TNF-α) levels than normally cycling women who did not self-report problematic PMS symptoms across the cycle, suggesting that these mechanisms may be involved in the aetiology of the syndrome. To ensure the scientific quality of the clinical trial, certain methodological considerations were explored. PMS is diagnosed in various ways, which has resulted in PMS studies being conducted on heterogeneous samples of women, who are often not analogous to women requiring treatment in clinical practice. This study highlighted the need for researchers to use a diagnostic procedure that identifies PMS sufferers experiencing PMS symptoms at a severity appropriate to address the aim of their study, and that differentiates women with PMS from those with clinical anxiety and depression. Moreover, the DSR (Freeman et al., 1996) was re-factor analysed and a two factor solution was produced, the DSR-20. This new measure was shown to be a more sensitive tool than the original DSR to assess treatment effects in the sample recruited for this research. Collectively these findings could improve future diagnostic and therapeutic strategies.
# TABLE OF CONTENTS

## CHAPTER 1: INTRODUCTION

1. **The normal menstrual cycle**
   1.1 The follicular phase
   1.1.1 The luteal phase
   1.1.2 Ovulation
   1.1.3 Steroid hormone concentrations across the cycle
2. **Menstrual cycle length**
3. **The Premenstrual Syndrome (PMS)**
   3.1 PMS/ PMDD Diagnosis
   3.1.1 Diagnostic Criteria
   3.1.2 Cycle phase designation in studies of PMS
   3.1.3 Measures
   3.1.4 Co-morbidities
   3.1.5 Summary of current PMS diagnostic techniques
4. **PMS symptoms**
   4.1 Anxiety and depression
   4.1.1 Depression, anxiety and PMS sufferers
   4.1.2 Depression, anxiety and ‘normal’ women
   4.1.3 Depression, anxiety, PMS sufferers and controls
   4.1.4 Summary of depression and anxiety in PMS
   4.2 Aggression and Impulsivity
   4.2.1 Impulsivity
   4.2.2 Aggression
   4.2.2.1 Aggression, crime and the menstrual cycle
   4.2.2.2 Questionnaire and laboratory assessment
   4.2.2.2.1 Aggression levels and PMS sufferers
   4.2.2.2.2 Aggression levels and ‘normal’ women
   4.2.2.3 Summary of impulsivity and aggression in PMS
5. **Aetiology**
   5.1 Metabolic
   5.2 Hormonal
   5.3 Neurotransmitters
   5.3.1 The role of testosterone in PMS
   5.3.2 The role of cytokines in PMS
   5.3.2.1 What are cytokines?
   5.3.2.2 Cytokines and depression
   5.3.2.3 Potential involvement of cytokines in PMS
6. **Conclusion**

## CHAPTER 2: AIMS

1. **Treatment**
2. **Characterisation**
   2.1 Psychological profiles
   2.2 Biological profiles
CHAPTER 3: GENERAL METHODOLOGY ................................................ 43

3.1 Measures .......................................................................................... 43
3.1.1 Daily Diary Booklets ........................................................................... 43
3.1.1.1 The Daily Symptom Report ............................................................ 44
3.1.1.2 The State-Trait Anxiety Inventory ......................................................... 45
3.1.1.3 The Beck Depression Inventory ............................................................ 46
3.1.1.4 The Barratt Impulsiveness Scale ........................................................... 47
3.1.1.5 The Aggression Questionnaire .............................................................. 47
3.1.1.6 Counterbalancing and Order effects ....................................................... 48
3.2 Pilot study of the daily diary booklets ....................................................... 49
3.2.1 Introduction ...................................................................................... 49
3.2.2 Procedure ........................................................................................ 49
3.2.3 Results ........................................................................................... 50
3.2.4 Conclusion ....................................................................................... 53
3.3 Menstrual cycle phases ......................................................................... 53
3.4 The screening cycles ............................................................................. 55
3.4.1 Recruitment Methods ........................................................................... 55
3.4.1.1 Clinical Trial .................................................................................. 55
3.4.1.2 Menstrual cycle study ........................................................................ 56
3.4.2 Selection Criteria ................................................................................ 56
3.4.3 Participation Progress ............................................................................ 57
3.4.4 Measures .......................................................................................... 58
3.4.4.1 Recruitment questionnaire ................................................................... 58
3.4.4.2 PMS symptoms ................................................................................ 59
3.4.4.3 Blood samples .............................................................................. 59
3.4.4.4 Emotional Stroop Task ....................................................................... 59
3.4.5 Procedure ......................................................................................... 59
3.5 Participant groups ............................................................................... 62
3.5.1 Co-morbid criteria .............................................................................. 62
3.5.2 PMS criteria ..................................................................................... 64
3.5.2.1 30% increase criterion ........................................................................... 66
3.5.2.2 Modified 30% increase criterion ........................................................... 70
3.5.2.3 Benefits and limitations of the 30% increase and modified 30% increase criterions ................................................................................................... 74
3.5.2.4 Choosing an appropriate criterion ......................................................... 76
3.5.3 Group formation for the analyses presented in this thesis ......................... 78

CHAPTER 4: THE CLUSTERING OF ITEMS ON THE DSR ......................... 80

4.1 Introduction ....................................................................................... 80
4.2 Data Reduction techniques and their key characteristics ......................... 80
4.2.1 Decision-making in PCA ................................................................. 81
4.2.1.1 Factor Extraction .............................................................................. 81
4.2.1.2 Factor Rotation ................................................................................ 83
4.3 Freeman’s Factors ............................................................................... 84
4.4 Reasons for conducting PCAs ............................................................. 85
4.5 Method ............................................................................................. 86
4.5.1 Participants ...................................................................................... 86
4.5.2 Measures ........................................................................................ 87
4.5.3 Procedure ......................................................................................... 87
4.6 Process of Analysis .............................................................................. 87
CHAPTER 5: CHARACTERISING PMS

5.1 Introduction ................................................................. 112
5.1.1 Psychological profiles .............................................. 112
5.1.2 Physiological profiles .................................................. 114
5.2 Method ................................................................. 115
5.2.1 Participants .......................................................... 115
5.2.2 Measures .......................................................... 115
5.2.2.1 Self-report measures ........................................ 115
5.2.2.2 Biological measures ........................................ 116
5.2.2.2.1 Gonadal steroids ........................................ 116
5.2.2.2.2 Cytokine analysis ...................................... 117
5.2.2.2.3 Serotonin .................................................. 117
5.2.3 Procedure .................................................. 118
5.3 Data Processing and Analysis .......................................... 118
5.3.1 Self-report measures .............................................. 118
5.3.2 Biological measures .................................................. 119
5.3.3 Data Screening .................................................. 119
5.3.3.1 Self-report measures ........................................ 119
5.3.3.1.1 Missing data ........................................ 119
5.3.3.1.2 Sample sizes ........................................ 120
5.3.3.1.3 Outliers .................................................. 120
5.3.3.1.4 Normality .................................................. 120
5.3.3.1.5 Sphericity .................................................. 120
5.3.3.2 Biological measures ........................................ 121
5.3.3.2.1 Missing data and protocol violations .......... 121
5.4 Results ............................................................. 122
5.4.1 Psychological symptom profiles for the three diagnostic groups ........................... 122
5.4.1.1.1 BPAQ total scale scores .................................. 122
5.4.1.1.2 The BPAQ subscales ................................... 124
5.4.1.1.2.1 Anger .................................................. 126
5.4.1.1.2.2 Verbal aggression ................................... 127
5.4.1.1.2.3 Hostility .................................................. 127
5.4.1.1.2.4 Physical aggression ................................... 128
5.4.1.2 Impulsivity (BIS-11) ........................................ 128
5.4.1.2.1 BIS 11 total scale scores ................................ 128
5.4.1.2.2 BIS-11 subscales .......................................... 130
5.4.1.2.2.1 Attentional impulsivity ................................ 132
5.4.1.2.2.2 Non-planning impulsivity ................................ 132
5.4.1.2.2.3 Motor impulsivity .................................... 133
5.4.2 Psychological symptom profiles for the four diagnostic groups ......................... 133
5.4.2.1 Aggression .................................................. 133
5.4.2.1.1 BPAQ total scale scores .................................. 133
5.4.2.1.2 The BPAQ subscales ................................... 135
5.4.2.1.2.1 Anger .................................................. 137
5.4.2.1.2.2 Verbal aggression ................................... 137
CHAPTER 6: DIETARY SUPPLEMENTS AND HERBAL REMEDIES FOR PREMENSTRUAL SYNDROME (PMS): A SYSTEMATIC RESEARCH REVIEW OF THE EVIDENCE FOR THEIR EFFICACY

6.1 Introduction
6.2 Method
6.3 Data synthesis for the dietary supplements and herbal remedies for PMS
6.4 Discussion
6.5 Conclusion

CHAPTER 7: EFFICACY OF ST. JOHN'S WORT FOR PMS

7.1 Introduction
7.1.1 St John’s Wort (SJW)
7.1.2 Antidepressant properties of SJW
CHAPTER 8: GENERAL DISCUSSION

8.1 Treatment
8.2 Variations in treatment assessment methods
8.3 The DSR-20 as a tool to assess treatment effects
8.4 The placebo effect in PMS
8.5 The characterisation and diagnosis of PMS
8.6 Aetiology
8.7 PMS diagnostic criteria
8.8 Limitations of the research
8.8.1 Standardisation of herbal products
8.8.2 Demand characteristics
8.8.3 Co-morbidities
8.8.4 Weekly versus daily diary collection
8.8.5 Absence of data on appetite, alcohol and caffeine intake
8.8.6 Multiple statistical testing
APPENDICES ........................................................................................................ 311
Appendix 3.1 Questionnaires comprising the daily diary booklet .................... 311
Appendix 3.2: Preparations serving as exclusion criteria ................................. 318
Appendix 3.3: Menstrual cycle study handbook .............................................. 319
Appendix 3.4: RCT Information Sheet ................................................................. 322
Appendix 3.5: Information sheet for women who were screened by post .......... 325
Appendix 3.6: Consent form for women being screened by post ..................... 327
Appendix 3.7: Menstrual cycle study consent form ........................................ 328
Appendix 3.8: RCT Admission Form ................................................................. 329
Appendix 3.9: RCT Consent Form .................................................................. 330
Appendix 3.10 Cut-offs for clinically significant anxiety ............................... 331
Appendix 8.1 Summary of product characteristics ..................................... 332
Appendix 8.2: RCT Recruitment Questionnaire ............................................ 335
Appendix 8.3 Interview Guide for Safety Assessment .................................... 339
Appendix 8.4 Example symptom profile for women screened by post .......... 340
Appendix 8.5 Batch analysis ......................................................................... 344
Appendix 8.6 Published article .................................................................... 345
LIST OF TABLES

Table 1.1 Studies assessing anxiety and depression levels during the menstrual cycle ...... 15
Table 1.2 Studies assessing impulsivity levels during the menstrual cycle.....................26
Table 1.3 Studies assessing aggression levels during the menstrual cycle......................30
Table 3.1 Means (sds) and Pearson’s moment correlation coefficients for the original and modified questionnaire subscales and total scores (n=36)...................................50
Table 3.2 Pearson’s Moment Correlation Coefficients between the daily and weekly measures..........................................................52
Table 3.3 Inclusion and exclusion criteria for participation.............................................57
Table 3.4 Co-morbid characteristics..............................................................................64
Table 3.5 Diagnostic groups formed from the clinical trial, postal screening and menstrual cycle study when the 30% increase criterion was employed..................... 67
Table 3.6 Diagnostic groups formed from the clinical trial, postal screening and menstrual cycle study when the modified 30% increase criterion was employed.......... 71
Table 4.1 Factors derived from DSR items (Freeman et al., 1996)...............................84
Table 4.2. Summary table of the number of factors produced by the scree plot for each of the 17- and 20-item models.................................................................90
Table 4.3 Summary of factor loadings produced by PA for the 17-item models.............91
Table 4.4. Summary of factor loadings produced by the K1 rule for the 17-item models...92
Table 4.5. Summary of factor loadings produced by PA and the K1 rule for the different diagnostic groups for the 17-item models................................................94
Table 4.6. Summary of factor loadings produced by PA and the K1 rule for the different diagnostic groups for the 20-item models.............................................96
Table 4.7 Summary of factor loadings produced by PA for the 20-item models.............98
Table 4.8. Summary of factor loadings produced by the K1 rule for the 20-item models...100
Table 4.9 Final models with factor loadings and Cronbach’s α coefficients (N=66).........109
Table 5.1 Numbers of women on which the analyses were conducted.........................119
Table 5.2 Multivariate and univariate statistics for the main effects of phase and group, and for the phase by group interaction......................................................125
Table 5.3 Multivariate and univariate statistics for the main effect of phase and group and the phase by group interaction..........................................................130
Table 5.4 Multivariate and univariate statistics for the main effect of phase and group
Table 5.5 Multivariate and univariate statistics for the main effect of phase and group
and the phase by group interaction ................................................................. 135
Table 5.6 Phase effects and luteal and follicular to luteal statistical differences between
the PMS sufferers and the other groups of women, on the BPAQ and BIS-11 (and their
subscales), when women were categorised into both three and four diagnostic groups.... 145
Table 5.7 Univariate statistics for the effects of cycle phase, group, and cycle phase by
group interactions for the steroid hormone analyses ........................................... 147
Table 5.8 Univariate statistics for the effects of cycle phase, group, and the cycle phase
by group interactions for the cytokine analyses ................................................. 149
Table 5.9 Univariate statistics for the effects of cycle phase, group, and cycle phase by
group interactions for the steroid hormone analyses ......................................... 151
Table 5.10 Univariate statistics for the effects of cycle phase, group, and the cycle phase
by group interactions for the cytokine analyses ................................................ 153
Table 6.1 Summary table of study characteristics .............................................. 168
Table 7.1 Inclusion and exclusion criteria for participation ................................... 188
Table 7.2 Characteristics of the clinical trial completers at study entry .................... 200
Table 7.3 Mean (sd) symptom severity during screening (average of study cycles 1-3).... 201
Table 7.4 Baseline symptomatology: Multivariate/ univariate statistics for the main effects
of phase and cycle, and for the cycle by phase interactions ................................. 201
Table 7.5 Multivariate and univariate statistics for the main effect for study phase, and
for the study phase by cycle phase interaction for the DSR .................................. 204
Table 7.6 Multivariate and univariate statistics for the main effect of study phase, and
for the study phase by cycle phase interaction for the DSR-20 ............................. 206
Table 7.7 Multivariate and univariate statistics for the main effect for study phase, and
for the study phase by cycle phase interaction for the BPAQ subscales ................. 210
Table 7.8 Multivariate and univariate statistics for the main effect for study phase, and
for the study phase by cycle phase interaction for the BIS-11 subscales ............... 212
Table 7.9 Summary table of placebo effects .................................................... 213
Table 7.10 The placebo effect during study cycles 6-7: Multivariate and univariate
statistics for treatment group 2 (n=16) .......................................................... 217
Table 7.11 Success of randomisation: Multivariate and univariate statistics for the main
effects of phase and treatment order, and for the phase by treatment order interactions.... 221
Table 7.12 Multivariate and univariate statistics for the main effect of treatment, and for
the treatment by cycle and treatment by cycle phase interactions for the DSR treatment
analysis ............................................................................................................ 224
Table 7.13 Multivariate and univariate statistics for the main effect of treatment and treatment by cycle interaction for the DSR analysis of premenstrual scores................. 227
Table 7.14 Multivariate and univariate statistics for the main effect of treatment and treatment by cycle interaction for the DSR analysis of follicular to luteal change scores.. 228
Table 7.15 Multivariate and univariate statistics for the main effect of treatment for the DSR analysis of follicular to luteal change scores (2nd cycle)...........................................229
Table 7.16 Multivariate and univariate statistics for the main effect of treatment for the DSR analysis of change from baseline scores.................................................... 230
Table 7.17 Multivariate and univariate statistics for the main effect of treatment, and treatment by cycle and treatment by cycle phase interactions, for the DSR-20 treatment analysis of actual scores................................................................. 231
Table 7.18 Multivariate and univariate statistics for the main effect of treatment and treatment by cycle interaction, for the DSR-20 analysis of premenstrual scores........... 232
Table 7.19 Multivariate and univariate statistics for the main effect of treatment and treatment by cycle interaction for the DSR-20 analysis of follicular to luteal change scores.................................................................................................... 233
Table 7.20 Multivariate and univariate statistics for the main effect of treatment for the DSR-20 analysis of follicular to luteal change scores (2nd cycle)...........................................234
Table 7.21 Multivariate and univariate statistics for the main effect of treatment for the DSR-20 analysis of change from baseline scores.................................................... 234
Table 7.22 Summary table of treatment effects................................................... 240
Table 7.23 Univariate statistics for the main effects of treatment, cycle phase, and treatment by cycle phase interactions for the steroid hormone analyses (n=14).............. 243
Table 7.24 Univariate statistics for the main effects of treatment and cycle phase, and treatment by cycle phase interactions for the cytokine analyses (n=12 to n=14)........... 245
LIST OF FIGURES

Figure 3.1 Weekly diary phase designation for a 37 day menstrual cycle..................... 55
Figure 3.2 Participant recruitment and progress at each stage of the screening cycles....... 58
Figure 3.3 The total DSR profiles of the four groups averaged over the three screening cycles.................................................................................................... 68
Figure 3.4 The mood symptom profiles of the four groups averaged over the three screening cycles ........................................................................................ 69
Figure 3.5 The behavioural symptom profiles of the four groups averaged over the three screening cycles ........................................................................................ 69
Figure 3.6 The pain symptom profiles of the four groups averaged over the three screening cycles ........................................................................................ 69
Figure 3.7 The physical symptom profiles of the four groups averaged over the three screening cycle ........................................................................................... 69
Figure 3.8 The total DSR profiles of the four groups averaged over the three screening cycles .................................................................................................... 72
Figure 3.9 The mood symptom profiles of the four groups averaged over the three screening cycles ........................................................................................ 73
Figure 3.10 The behavioural symptom profiles of the four groups averaged over the three screening cycles ........................................................................................ 73
Figure 3.11 The pain symptom profiles of the four groups averaged over the three screening cycles ........................................................................................ 73
Figure 3.12 The physical symptom profiles of the four groups averaged over the three screening cycle ........................................................................................... 73
Figure 5.1 The BPAQ-total profiles of the three groups averaged over the three screening cycles .................................................................................................... 123
Figure 5.2 The anger profiles of the three groups averaged over the three screening cycles................................................................................................................. 126
Figure 5.3 The verbal aggression profiles of the three groups averaged over the three screening cycles ............................................................................................................. 126
Figure 5.4 The hostility profiles of the three groups averaged over the three screening cycles ................................................................................................................. 126
Figure 5.5 The physical aggression profiles of the three groups averaged over the three
screening cycles ........................................................................................ 126
Figure 5.6 The BIS-11 profiles of the three diagnostic groups averaged over the three
screening cycles ........................................................................................ 129
Figure 5.7 The attentional impulsivity profiles of the three diagnostic groups averaged
over the three screening cycles ................................................................... 131
Figure 5.8 The motor impulsivity profiles of the three diagnostic groups averaged over
the three screening cycles ........................................................................... 131
Figure 5.9 The non-planning impulsivity profiles of the three groups averaged across the
three screening cycles .................................................................................. 131
Figure 5.10 The BPAQ-total profiles of the four groups averaged over the three
screening cycles ........................................................................................ 134
Figure 5.11 The anger profiles of the four groups averaged over the three screening
cycles .................................................................................................... 136
Figure 5.12 The verbal aggression profiles of the four groups averaged over the three
screening cycles ........................................................................................ 136
Figure 5.13 The hostility profiles of the four groups averaged over the three screening
cycles .................................................................................................... 136
Figure 5.14 The physical aggression profiles of the four groups averaged over the three
screening cycles ........................................................................................ 136
Figure 5.15 The BIS-11 profiles of the four diagnostic groups averaged over the three
screening cycles ........................................................................................ 139
Figure 5.16 The attentional impulsivity profiles of the four groups across the averaged
screening cycles ........................................................................................ 141
Figure 5.17 The motor impulsivity profiles of the four groups across the averaged
screening cycles ........................................................................................ 141
Figure 5.18 The non-planning impulsivity profiles of the four groups across the averaged
screening cycles ........................................................................................ 141
Figure 5.19 Mean oestradiol concentration .................................................... 146
Figure 5.20 Mean progesterone concentration .................................................. 146
Figure 5.21 Mean testosterone concentration .................................................. 146
Figure 5.22 Mean LH concentration ............................................................... 146
Figure 5.23 Mean FSH concentration ............................................................... 146
Figure 5.24 Mean PRL concentration ............................................................... 146
Figure 5.25 Mean IL-1β concentration ............................................................. 148
Figure 5.26 Mean IL-6 concentration ............................................................... 148
Figure 5.27 Mean IL-8 concentration ............................................................... 148
Figure 5.28 Mean TNF-α concentration ............................................................ 148
Figure 5.29 Mean IFN-γ concentration ............................................................. 148
Figure 5.30 Mean oestradiol concentration ....................................................... 150
Figure 5.31 Mean progesterone concentration .................................................... 150
Figure 5.32 Mean testosterone concentration .................................................... 150
Figure 5.33 Mean LH concentration ................................................................. 150
Figure 5.34 Mean FSH concentration ............................................................... 150
Figure 5.35 Mean PRL concentration ............................................................... 150
Figure 5.36 Mean IL-1β concentration .............................................................. 152
Figure 5.37 Mean IL-6 concentration ............................................................... 152
Figure 5.38 Mean IL-8 concentration ............................................................... 152
Figure 5.39 Mean TNF-α concentration ............................................................ 152
Figure 5.40 Mean IFN-γ concentration ............................................................. 152
Figure 7.1 Participant recruitment and progress at each stage of the clinical trial ........ 189
Figure 7.2 Schematic representation of the study design and measures completed in each study phase ................................................................. 195
Figure 7.3 Mean DSR scores during study cycles 1-5 ............................................ 203
Figure 7.4 Mean DSR subscale scores during study cycles 1-5 ............................... 203
Figure 7.5 Mean DSR-20 subscale scores during study cycles 1-5 ............................ 205
Figure 7.6 Mean STAIS scores during study cycles 1-5 ........................................ 207
Figure 7.7 Mean BDI scores during study cycles 1-5 ............................................. 208
Figure 7.8 Mean BPAQ scores during study cycles 1-5 ........................................ 209
Figure 7.9 Mean BPAQ subscale scores during study cycles 1-5 ............................. 209
Figure 7.10 Mean BIS-11 scores during study cycles 1-5 ..................................... 211
Figure 7.11 Mean BIS-11 subscale scores during study cycles 1-5 ......................... 211
Figure 7.12 Mean DSR scores reported by treatment group two during study cycles 1-7 ... 214
Figure 7.13 Mean DSR-20 scores reported by treatment group two during study cycles 1-7 ................................................................. 214
Figure 7.14 Mean STAIS scores reported by treatment group two during study cycles 1-7 ................................................................. 215
Figure 7.15 Mean BDI scores reported by treatment group two during study cycles 1-7 ... 215
Figure 7.16 Mean BPAQ scores reported by treatment group two during study cycles 1-7 ................................................................. 216
Figure 7.17 Mean BIS-11 scores reported by treatment group two during study cycles 1-7 ................................................................. 216
Figure 7.18 Effect of treatment on total DSR symptoms during each treatment cycle .... 222
Figure 7.19 Effect of treatment on mood symptoms during each treatment cycle............ 223
Figure 7.20 Effect of treatment on behavioural symptoms during each treatment cycle..... 223
Figure 7.21 Effect of treatment on pain symptoms during each treatment cycle.............. 223
Figure 7.22 Effect of treatment on physical symptoms during each treatment cycle........... 223
Figure 7.23 Treatment effect on psychological symptoms during each treatment cycle..... 231
Figure 7.24 Treatment effect on physical symptoms during each treatment cycle............. 231
Figure 7.25 Effect of treatment on total aggression scores during each treatment cycle..... 235
Figure 7.26 Effect of treatment on verbal aggression during each treatment cycle............ 236
Figure 7.27 Effect of treatment on physical aggression during each treatment cycle....... 236
Figure 7.28 Effect of treatment on hostility aggression during each treatment cycle....... 236
Figure 7.29 Effect of treatment on anger during aggression during each treatment cycle... 236
Figure 7.30 Effect of treatment on total impulsivity scores during each treatment cycle... 237
Figure 7.31 Effect of treatment on attentional impulsivity during each treatment cycle..... 238
Figure 7.32 Effect of treatment on motor impulsivity during each treatment cycle......... 238
Figure 7.33 Effect of treatment on non-planning impulsivity during each treatment cycle. 238
Figure 7.34 Effect of treatment on total anxiety scores during each treatment cycle...... 239
Figure 7.35 Effect of treatment on total depression scores during each treatment cycle... 239
Figure 7.36 Mean oestradiol concentration......................................................... 242
Figure 7.37 Mean progesterone concentration...................................................... 242
Figure 7.38 Mean testosterone concentration....................................................... 242
Figure 7.39 Mean LH concentration................................................................ 242
Figure 7.40 Mean FSH concentration............................................................... 242
Figure 7.41 Mean PRL concentration............................................................... 242
Figure 7.42 Mean IL-1β concentration.............................................................. 244
Figure 7.43 Mean IL-6 concentration.............................................................. 244
Figure 7.44 Mean IL-8 concentration.............................................................. 244
Figure 7.45 Mean TNF-α concentration............................................................. 244
Figure 7.46 Mean IFN- γ concentration.............................................................. 244
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-HT</td>
<td>Serotonin</td>
</tr>
<tr>
<td>5-HIAA</td>
<td>5-hydroxyindoleacetic acid</td>
</tr>
<tr>
<td>APA</td>
<td>American Psychiatric Association</td>
</tr>
<tr>
<td>B6</td>
<td>Vitamin B6</td>
</tr>
<tr>
<td>BDI</td>
<td>Beck Depression Inventory</td>
</tr>
<tr>
<td>BIS</td>
<td>Barratt Impulsiveness Scale</td>
</tr>
<tr>
<td>BPAQ</td>
<td>Buss Perry Aggression Questionnaire</td>
</tr>
<tr>
<td>CL</td>
<td>Confidence limit</td>
</tr>
<tr>
<td>CNV</td>
<td>Contingent negative variation</td>
</tr>
<tr>
<td>CRH</td>
<td>Corticotrophin releasing hormone</td>
</tr>
<tr>
<td>DSM</td>
<td>Diagnostic and Statistical Manual</td>
</tr>
<tr>
<td>DRF</td>
<td>Daily Rating Form</td>
</tr>
<tr>
<td>DSR</td>
<td>Daily Symptom Report</td>
</tr>
<tr>
<td>DSR-17</td>
<td>Two factor, 17 item Daily Symptom Report</td>
</tr>
<tr>
<td>DSR-20</td>
<td>Two factor, 20 item Daily Symptom Report</td>
</tr>
<tr>
<td>FA</td>
<td>Factor Analysis</td>
</tr>
<tr>
<td>FSH</td>
<td>Follicle Stimulating Hormone</td>
</tr>
<tr>
<td>GAD</td>
<td>Generalised Anxiety Disorder</td>
</tr>
<tr>
<td>GMP</td>
<td>Good Manufacturing Practice</td>
</tr>
<tr>
<td>GnRH</td>
<td>Gonadotrophin-Releasing Hormone</td>
</tr>
<tr>
<td>HARU</td>
<td>Human Appetite Research Unit, University of Leeds</td>
</tr>
<tr>
<td>HRP</td>
<td>Horse Radish Peroxidase</td>
</tr>
<tr>
<td>HRT</td>
<td>Hormone Replacement Therapy</td>
</tr>
<tr>
<td>ICD</td>
<td>International classification of diseases</td>
</tr>
<tr>
<td>IFN</td>
<td>Interferon</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>LH</td>
<td>Luteinising Hormone</td>
</tr>
<tr>
<td>MD</td>
<td>Major depression</td>
</tr>
<tr>
<td>MDD</td>
<td>Major Depressive Disorder</td>
</tr>
<tr>
<td>MDQ</td>
<td>Menstrual Distress Questionnaire</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>MHRA</td>
<td>Medicines and Healthcare products Regulatory Agency</td>
</tr>
<tr>
<td>NIMH</td>
<td>National Institute of Mental Health</td>
</tr>
<tr>
<td>NT</td>
<td>Neurotransmitter</td>
</tr>
<tr>
<td>OTC</td>
<td>Over the counter</td>
</tr>
<tr>
<td>PAF</td>
<td>Premenstrual Assessment Form</td>
</tr>
<tr>
<td>PA</td>
<td>Parallel Analysis</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Components Analysis</td>
</tr>
<tr>
<td>PI</td>
<td>Principal investigator</td>
</tr>
<tr>
<td>PMDD</td>
<td>Premenstrual Dysphoric Disorder</td>
</tr>
<tr>
<td>PMS</td>
<td>Premenstrual Syndrome</td>
</tr>
<tr>
<td>PRL</td>
<td>Prolactin</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised controlled trial</td>
</tr>
<tr>
<td>sd</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SJW</td>
<td>St. John’s Wort</td>
</tr>
<tr>
<td>SSRI</td>
<td>Selective Serotonin Reuptake Inhibitors</td>
</tr>
<tr>
<td>STAI</td>
<td>State Trait Anxiety Inventory</td>
</tr>
<tr>
<td>STAIS</td>
<td>State Anxiety Scale</td>
</tr>
<tr>
<td>STAIT</td>
<td>Trait Anxiety Scale</td>
</tr>
<tr>
<td>TNF</td>
<td>Tumor necrosis factor</td>
</tr>
<tr>
<td>VAS</td>
<td>Visual Analogue Scale</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
</tbody>
</table>


Chapter 1: Introduction

1.1 The normal menstrual cycle

The menstrual cycle is biphasic (Cutler and Garcier, 1979; Lane and Francis, 2003). It consists of the follicular phase during which follicles compete to release a mature oocyte (egg), and the luteal phase, where the corpus luteum prepares the body for pregnancy (O’Brien, 1987; Sherwood, 2004). These phases are divided by ovulation, when the oocyte is released into the reproductive tract (Keye, 1988). The onset of menstruation marks the end of one cycle and the beginning of the next, although the events of any two cycles overlap (Asso, 1983; Ferin et al., 1993; Lane and Francis, 2003).

The events of the menstrual cycle are controlled by an interaction between the ovarian hormones, oestrogen and progesterone, and the pituitary hormones, luteinising hormone (LH) and follicle stimulating hormone (FSH) (Asso, 1983), which interact through both complex positive and negative feedback mechanisms (Behl, 2001; Cutler and Garcier, 1979). Gonadotrophin-releasing hormone (GnRH) from the hypothalamus controls the release of the pituitary hormones LH and FSH. A brief description of the main events that occur during the menstrual cycle and the hormonal changes that control them is provided below (see Behl, 2001 for a detailed review).

1.1.1 The follicular phase

Females are born with approximately 200,000 primary follicles in their ovaries (Cutler and Garcier, 1979), which contain a primary oocyte surrounded by a single layer of granulosa cells (Sherwood, 2004). In an average woman’s life, approximately 300-400 primary follicles mature and release ova (O’Brien, 1987; Sherwood, 2004), in a process that takes approximately 13 days (Cutler and Garcier, 1979).

Groups of primary follicles develop throughout the menstrual cycle, although only those that do so during the late luteal (Cutler and Garcier, 1979) and follicular (Sherwood, 2004) phases continue their development, increasing in both size and cell number, under the influence of FSH. As these follicles mature, they secrete increasing amounts of oestrogen, which through a negative feedback system, leads to a reduction in FSH and LH levels (Asso, 1983). However, as progesterone is also needed to fully suppress LH secretion,
which remains low during the follicular phase, LH levels slowly increase (Sherwood, 2004). The follicles compete for dominance, resulting in one, or occasionally two, becoming dominant (Graafian follicle). All others undergo atresia.

1.1.2 Ovulation
Oestrogen levels peak when follicle size is maximal. This triggers a surge in LH and FSH levels through a positive feedback system, with levels reaching a maximum around 24-36 hours after the oestrogen peak (Cutler and Garcia, 1979; Shaw, 1978). As progesterone is also secreted by the Graafian follicle, this may facilitate the surge of pituitary hormones through a positive feedback mechanism (Cutler and Garcia, 1979). It may also minimise the length of the LH surge through a negative feedback system (Asso, 1983).

The release of LH further matures the oocyte and weakens the follicle wall, resulting in its release from the dominant follicle around 18-24 hours after the LH peak (Asso, 1983; Keye, 1988). If the egg is not fertilised soon after ovulation, it disintegrates. LH seems to be the pituitary hormone central to ovulation, as the FSH surge does not always coincide with the LH surge and sometimes does not occur at all (Asso, 1983). However, the combination of these pituitary hormones is optimal for the ovulation process (Shaw, 1978).

1.1.3 The luteal phase
The remains of the dominant follicle transform in the ovary into a solid body, the corpus luteum, under the influence of the pituitary hormones. The corpus luteum produces large amounts of progesterone and smaller amounts of oestrogen (Sherwood, 2004), which peak around 6-9 days after the LH surge (Keye, 1988). The rising progesterone levels result in the endometrium preparing for the implantation of an embryo and also reduces the secretion of the pituitary hormones. If implantation does not occur, the corpus luteum degrades, causing a sharp drop in both oestrogen and progesterone levels. This leads to the onset of menstruation and allows FSH and LH levels to rise, enabling a new group of primary follicles to mature (Sherwood, 2004).

1.1.4 Steroid hormone concentrations across the cycle
Levels of the main ovarian oestrogen, oestradiol, peak twice. Oestradiol increases as the follicles mature, with the largest peak occurring during the late follicular phase as the
Graafian follicle reaches its maximum size. Levels decrease at midcycle, when the dominant follicle disintegrates, and rise again as the corpus luteum develops, reaching another, lower peak. Progesterone remains virtually absent in the follicular phase, with levels rising approximately 12 hours before the LH surge. Levels then increase again around 36 hours after the onset of the LH surge and reach a peak around eight days later. LH levels remain low, except when levels surge at the end of the follicular phase for approximately 48 hours. FSH levels also rise at this time, although to a lesser degree. Levels of FSH then rise again just before the onset of menstruation and peak around 24 hours after menstruation begins (Behl, 2001; Ferin et al., 1993; O’Brien, 1987).

1.2 Menstrual cycle length
Menstrual cycle length varies both between and within women (Chiazze et al., 1968; McIntosh et al., 1980; Sherwood, 2004). Most agree that the follicular phase can vary in length (Fukuda et al., 1998; Weinberg et al., 1995). It has been argued that the luteal phase is fixed to 14 +/- 2 days, and that cycle length variability is solely accounted for by a variable follicular phase (Asso, 1983; Cutler and Garcia, 1979; Setchell and Cassidy, 1999). However, others believe that the luteal phase can vary in length (Dukelow, 1975; Shaikh et al., 1978; Stern and McClintock, 1998). Wilcox et al. (2000) estimated the timing of ovulation by assessing changes in the ratio of urinary metabolites of oestrogen and progesterone, and found that the time from ovulation to next menses ranged from seven to 19 days. Moreover, McIntosh et al. (1980) assessed the variation in follicular and luteal phase length by detecting the midcycle LH surge through daily blood sampling. They found that, on average, a change in cycle length of one day resulted in a 0.77 day change in follicular length and 0.23 day change in luteal length.

1.3 The Premenstrual Syndrome (PMS)
The premenstrual syndrome (PMS) refers to a wide range of cyclic and recurrent psychological, physical and behavioural symptoms that occur in the 7-10 days prior to the onset of menstruation and which remit shortly following the commencement of menstrual flow (Fava et al., 1992; Freeman, 2003; Halbreich et al., 2007; Johnson, 2004; Matsumoto et al., 2007; Reed et al., 2008). Although over 200 symptoms have been associated with PMS, the common symptoms that classically characterise the syndrome include the psychological symptoms irritability, mood swings, depression, anxiety and impulsivity.
physical symptoms bloatedness, breast tenderness, aches and headaches, and the
behavioural symptoms poor coordination, sleep disturbances and appetite changes
(Freeman, 2003; Halbreich et al., 2007; Haywood et al., 2002).

PMS severity falls along a continuum (Johnson, 2004), with symptoms ranging from mild
to severe (Borenstein et al., 2007; Johnson, 2004). The American Psychiatric Association
developed specific criteria for premenstrual dysphoric disorder (PMDD), which appear in
the 4th edition of the Diagnostic and Statistical Manual of Mental Disorders (APA, 1994).
PMDD describes a subgroup of women whose symptoms are particularly severe (Connolly,
2001; Freeman, 2003; Halbreich et al., 2007; Haywood et al., 2002; Johnson, 2004;
Pearlstein et al., 2005). PMDD is less prevalent than PMS, and has a greater psychological
impact, interfering with relationships, work and lifestyle (Steiner and Wilkins, 1996;
Steiner et al., 1999; Steiner and Born, 2000), at levels similar to women suffering from
major depression, MD (Freeman, 2003). Although PMS involves less severe symptoms
than PMDD (Ismail and O’Brien, 2001), PMS symptoms can still have a detrimental
impact on women’s lives (Freeman, 2003; Shaw et al., 2003).

Reports of the prevalence of PMS and PMDD have varied due to the different ways in
which the conditions have been diagnosed (Halbreich et al., 1982a; 2007; Moore and
Emanuel, 2001; see also section 1.3.1). However, researchers have proposed that up to 75% of
regularly menstruating women experience some physical and psychological symptoms of
PMS (Burt and Stein, 2002; Haywood et al., 2002; Hylan et al., 1999; Ramcharan et al.,
1992; Steiner and Wilkins, 1996), while approximately 3-8% meet criteria for PMDD
(Bryant and Dye, 2004; Burt and Stein, 2002; Halbreich et al., 2007; Haywood et al., 2002;
Pearlstein et al., 2005; Steiner and Born., 2000).

1.3.1 PMS/PMDD Diagnosis
There are no universally accepted PMS diagnostic criteria that are applied in clinical
practice (Halbreich et al., 2007). There is also wide variation in the diagnosis of PMS in
research settings (Budeiri et al., 1994; Freeman, 2003; Halbreich et al., 2007). A variety of
diagnostic criteria, assessment methods and measures have been used. Some researchers
(e.g. Bond et al., 2003; De Ronchi et al., 2005; Hsiao et al., 2004; Hsu et al., 2007; Reed et
al., 2008) have confined the diagnosis of PMS to women who do not suffer from co-
morbidities, while others do not take co-morbidities into account (e.g. Van der Ploeg, 1987; Watts et al., 1987). Moreover, researchers have differed in the way in which they have defined the cycle phases. Therefore, PMS diagnosis has been based on the assessment of symptoms during cycle phases based on different criteria.

1.3.1.1 Diagnostic Criteria

Some researchers diagnose PMS through the use of the ICD-10 criteria (WHO, 1996) (e.g. Hsiao et al., 2002; Nisar et al., 2008), which lists PMS as a physical condition. These criteria require women to experience one symptom from an extensive list, which includes symptoms such as feelings of bloating and weight gain, breast tenderness, swelling of hands and feet, various aches and pains, poor concentration, sleep disturbance and changes in appetite. Although symptoms are required to occur premenstrually, symptom severity, duration and impact do not need to be established (Freeman, 2003; Landen and Eriksson, 2003; Steiner et al., 2003; Steiner and Wilkins, 1996).

A more commonly used strategy for diagnosing PMS is the application of the guidelines set out at the 1983 U.S National Institute of Mental Health conference (NIMH) (Gallant et al., 1992a; Haywood et al., 2002), which base the diagnosis of PMS on the severity of symptoms and their pattern across the menstrual cycle. These unpublished criteria have been applied in two ways (Gallant et al., 1992a). The 30% increase criterion requires a woman to show a symptom increase of at least 30% from her follicular to luteal phase for at least two menstrual cycles for a PMS diagnosis to be confirmed (Facchinetti et al., 1991; Gallant et al., 1992a; Gise et al., 1990; Hamilton et al., 1984; Haywood et al., 2002; Logue and Moos, 1986; Mortola, 1996; Pearlstein et al., 1997). The modified 30% increase criterion also requires a woman to show at least a 30% increase in her symptoms from the follicular to luteal phase for at least two menstrual cycles for a PMS diagnosis to be confirmed, but in relation to the range of the scale used. This can be achieved by dividing the difference in the symptoms that are reported between the follicular and luteal phases by the possible range of ratings over all cycle days (Evans et al., 1998; 1999; Freeman, 2003; Gallant et al., 1992a; Schmidt et al., 1998; Schnurr, 1989; Smith et al., 2002; Tung-Ping et al., 1997). Both the 30% increase criterion (e.g. Facchinetti et al., 1991; Hicks et al., 2004; Pearlstein et al., 1997; Rapkin, 1987) and the modified 30% increase criterion (e.g. Evans et al., 1998; 1999; Schmidt et al., 1998; Smith et al., 2002; Tung-Ping et al., 1997) have 60
been widely adopted to diagnose PMS. However, researchers often do not make clear how they have operationalised the modified 30% increase criterion, and it is therefore possible that this criterion has not been applied in a consistent manner (e.g. Schmidt et al., 1998; Smith et al., 2002; Tung-Ping et al., 1997).

Although the majority of researchers have diagnosed PMS through the use of the NIMH (1983) criteria (e.g. Evans et al., 1998; 1999; Facchinetti et al., 1991; Hicks et al., 2004; Pearlstein et al., 1997; Rapkin, 1987; Schmidt et al., 1998; Smith et al., 2002; Tung-Ping et al., 1997), other severity of change criteria have been used (Gallant et al., 1992a; Gehlert and Hartlage, 1997). Some researchers have adopted an absolute severity threshold criterion, where symptoms are required to be of a particular severity during the follicular and luteal phases (Anderson et al., 1988). Others have adopted an effect size criterion, where the follicular to luteal change score is divided by the standard deviation of ratings across all cycle days (De Ronchi et al., 2004; Schnurr, 1988). Others have adopted a stricter percentage increase criterion where the symptoms reported between the follicular and luteal phases are required to increase by 50% (Thys-Jacobs et al., 1998) or 75% (Gallant et al., 1992a), while still others use various combinations of the NIMH (1983), absolute severity threshold, effect size and stricter percentage change criteria to make a PMS diagnosis (Atmaca et al., 2003; Graham and Sherwin, 1993; Stevinson and Ernst, 2000).

The DSM-IV criteria for PMDD (APA, 1994) require that women report a premenstrual worsening of 5 or more symptoms, from 11 symptoms, at a severity sufficient to impair functioning (Freeman, 2003; Steiner et al., 1999). The majority of the specified symptoms are emotional or behavioural in nature, and include fatigue, insomnia, feeling out of control and concentration difficulties. Women must experience either depressed mood, tension, affect lability or irritability, and must not meet criteria for another psychiatric disorder, such as anxiety or depression (Freeman, 2003; Halbreich et al., 2007; Steiner et al., 1999). Again, symptoms must be prospectively confirmed through at least two months of daily self-ratings. Severity criteria are usually applied. Conventionally, the NIMH (1983) criteria are adopted (Gallant et al., 1992a). However, again other severity of change criteria are sometimes used (Steiner et al., 1995). Few researchers have used the strict DSM-IV criteria to select their participants (Connolly, 2001; Ismail and O’Brien, 2001). Therefore, when
researchers have not used these criteria, the term PMS will be used in this thesis. When researchers have only selected participants suffering from the severe form of PMS through the application of the DSM-IV criteria, the term PMDD will be used in this thesis (e.g. Atmaca et al., 2003; Evans et al., 1998; Hsu et al., 2007; Reed et al., 2008).

The various severity of change criteria that have been used to diagnose PMS have been applied in different ways (Gallant et al., 1992a). For example, the NIMH (1983) did not specify the number or type of symptoms that should be affected in order to confirm a diagnosis of PMS (Mortola, 1996; Smith et al., 2002). Therefore, researchers have applied the NIMH (1983) criteria to PMS symptoms in a variety of ways. Some have assessed the overall severity of PMS by applying the NIMH (1983) criteria to a total scale score (e.g. Atmaca et al., 2003; Evans et al., 1998; Facchinetti et al., 1991). Others have required that a certain number of items on a measure meet these criteria (e.g. Gise et al., 1990; Graham and Sherwin, 1993; Jermain et al., 1999; Pearlstein et al., 1997). Others have applied these criteria to the PMS symptoms that are generally considered to be the most troublesome for PMS sufferers (e.g. Schnurr, 1989; Smith et al., 2002), whilst still others have identified each woman’s most troublesome symptoms before she commences daily diary completion and then applied the NIMH (1983) criteria to these symptoms (Steiner et al., 1996). Researchers who have used the effect size criterion, the absolute severity threshold criterion and a stricter percentage change criterion to diagnose PMS have also applied them in these various ways (Gallant et al., 1992a).

The severity of change criterion that is chosen, and the way in which it is applied, can influence the number of women who are identified as suffering from PMS. Gallant et al. (1992a) demonstrated that different proportions of women who self-diagnosed with PMS and who met a provisional PMDD diagnosis (DSM-II-R) had their diagnosis confirmed when different severity of change criteria (30% increase, modified 30% increase, effect size, absolute severity threshold, 75% increase) were applied to their daily symptom ratings. The application of the 30% increase criterion resulted in the greatest percentage of women having their PMS diagnosis confirmed, while the 75% increase criterion resulted in the least. These researchers also demonstrated that there was a dramatic decrease in the proportion of women who had their provisional PMS diagnosis confirmed when the same criterion was applied at a conservative level, where five or more symptoms were required.
to meet the criterion, than when it was applied at a liberal level, where only one symptom was required to do so.

As there is currently no consensus regarding which severity of change criterion should be used to make a PMS diagnosis, or regarding the way in which PMS criteria should be applied to PMS symptoms (Gallant et al., 1992a), researchers should choose to use a diagnostic procedure that identifies PMS sufferers who are experiencing PMS symptoms at a severity appropriate to address the aim of their study (Johnson, 2004). For example, researchers performing PMS treatment trials should use a method of diagnosis that identifies PMS sufferers who are experiencing PMS symptoms at a severity that is representative of the women at whom that treatment is aimed. Clinicians usually take a hierarchical approach to PMS treatment (Johnson, 2004; see also section 6.1). Whilst lifestyle changes and the use of over the counter (OTC) drugs are advocated for PMS sufferers with mild to moderate symptoms (Bryant and Dye, 2004; Connolly, 2001; Johnson, 2004), SSRIs are the treatment of choice for women who experience more severe symptoms (Dimmock et al., 2000; Eriksson et al., 2002; Johnson, 2004; Wyatt et al., 2002). Therefore, researchers who are conducting treatment trials to assess the effectiveness of SSRIs for PMS should use strict PMS criteria that identify women who are experiencing severe PMS symptoms (e.g. Freeman et al., 2005; Tung-Ping et al., 1997), whilst researchers who are conducting PMS treatment trials to assess the effectiveness of OTC drugs could use more liberal PMS criteria that identify women who are experiencing mild to moderate PMS symptoms, which such preparations might alleviate (e.g. Bryant et al., 2005; De Souza et al., 2000; Facchinetti et al., 1991; Freeman et al., 2002; Hicks et al., 2004; Walker et al., 1998). Although more meaningful conclusions can be drawn when the symptom severity of research participants is taken into account, this can also make it difficult to draw meaningful comparisons between studies. This is because studies may have used different criteria to diagnose PMS, and applied these criteria in different ways, resulting in study samples comprising PMS sufferers experiencing PMS symptoms at different severities.

1.3.1.2 Cycle phase designation in studies of PMS

To make a PMS diagnosis, it is necessary to compare the premenstrual days (luteal phase) with those after menstruation (follicular phase) (Connolly, 2001). However, researchers
have defined the follicular and luteal phases differently (Atmaca et al., 2003; Collins et al., 1993; Steiner and Wilkins, 1996; Thys-Jacobs et al., 1998). Some use specific cycle days (e.g. Collins et al., 1993), some cycle length adjusted days (e.g. Facchinetti et al., 1991, Hicks et al., 2004), while others have chosen the worst days based on severity of symptoms from a certain period (e.g. Keenan et al., 1992). The consensus reached at the NIMH PMS workshop (1983) was that cycle length adjusted days should be used, and that cycle days 5 to 10 should be taken to represent the follicular phase and cycle days -6 to -1 the luteal phase. Although many studies have applied these guidelines (e.g. Atmaca et al., 2003; Bryant et al., 2005; Chisholm et al., 1990; Freeman et al., 1996; Kendall and Schnurr, 1987), not all studies of PMS have done so, and a variety of criteria are still used in clinical practice.

1.3.1.3 Measures

Over 65 instruments have been identified to diagnose PMS and/or assess treatment outcomes (Budeiri et al., 1994; Haywood et al., 2002). These cover over two hundred symptoms (Halbreich et al., 2007). Some measures are retrospective in nature, where symptoms from the last or usual cycle are rated from memory. The most widely used retrospective measures (Budeiri et al., 1994; Christensen et al., 1989; Haywood et al., 2002) include the Menstrual Distress Questionnaire (MDQ) (Moos, 1968), and the Premenstrual Assessment Form (PAF) (Halbreich et al., 1982b), both of which were empirically derived.

The MDQ requires women to rate their experience of 47 symptoms during the premenstrual, menstrual and intermenstrual phases of their most recent cycle on a six-point scale. The MDQ items were derived through performing a literature review and through conducting interviews and administering questionnaires to women. These items were then administered to over 800 women, who Moos (1968) claimed to be a representative sample of the wives of graduate students, with respect to menstrual cycle symptomatology. A factor analysis performed on this data revealed the MDQ to have eight 'symptom clusters' (e.g. pain, concentration and negative affect). Although the MDQ allows the comparison of symptom severity between the follicular and luteal phases, this measure has been criticized for being complex (Haywood et al., 2002) and for using a socially homogeneous sample in its construction, some of whom were pregnant at the time of completion (Richardson, 1990).
Halbreich et al. (1982b) derived 150 items by consulting their own experience and through reviewing the literature and previous questionnaire items. They administered these items to over 150 women who were not seeking PMS treatment, and reduced the number of items to the 95 items comprising the PAF. This measure requires women to rate each item according to how they felt during their last three premenstrual phases compared to their ‘usual self.’ Although the PAF covers a range of symptoms, it has been criticized for being long and difficult to score (Budeiri et al., 1994; Haywood et al., 2002), for not providing a comparison of premenstrual symptoms with those reported during the follicular phase, and for not informing women of the timing of the premenstrual phase (Haywood et al., 2002).

More generally, the use of retrospective measures has been criticized, as they promote selective recall (Ainscough, 1990), incorrect symptom timing (Connolly, 2001), bias due to stereotypes (Gallant et al., 1992b), cultural expectations and heavy reliance upon memory (Ainscough, 1990), which results in inaccurate (Ainscough, 1990), and often inflated estimates of symptom severity (Christensen et al., 1989; De Souza et al., 2000; Gallant et al., 1992b; Roy-Bryne et al., 1985). Prospective daily self-report instruments, completed for at least two menstrual cycles, are now the accepted means of confirming PMS (Connolly, 2001; Freeman, 2003; Johnson, 2004; Steiner and Wilkins, 1996). These are easy to administer, are less reliant on memory (Haywood et al., 2002), and highlight inter-cycle variability in symptom type and severity (Connolly, 2001). Although the demanding nature of daily reports may bias symptom patterns though non-adherence (Connolly, 2001; Haywood et al., 2002), non-compliance is rare when the value of the charts is explained, since most PMS sufferers have experienced their symptoms for years (Johnson, 2004).

Various prospective measures are available and there is lack of agreement as to which is the most appropriate (Freeman, 2003; Steiner et al., 1999). Item-specific scales are respected because they are easily understood, target specific symptoms and are sensitive to specific treatment effects (Bryant and Dye, 2004; Steiner et al., 1999). A range of daily diaries are available and there is no consensus as to which of these is best (Steiner et al., 2003a). Therefore, researchers should choose a measure that is appropriate for the sample under investigation, by considering the sample upon which the scale was developed, and the way in which the scale presents items to participants. Four widely used daily rating scales, which require women to rate their symptoms each evening, include the Daily Rating Form
(DRF) (Endicott et al., 1986), which requires women to rate 20 items on a six-point scale, the Daily Symptom Report (DSR) (Freeman et al., 1996), which requires women to rate 17 items on a five-point scale, the Calender of Premenstrual Experiences (COPE) (Mortola et al., 1990), which requires women to rate 22 items on a four-point scale, and the Menstrual Distress Questionnaire-Today (MDQ-T) (Moos et al., 1969), which requires women to rate 20 items on a nine-point scale. Although these scales include many similar items, some differences are apparent. For example, the DSR and DRF both include symptoms relating to mood swings, irritability, depression, anxiety, food cravings and bloating. However, the DSR includes the items insomnia, feeling out of control, crying and headaches, while the DRF includes items relating to the avoidance of social acts, loss of sexual interest, increased sleep and decreased work. Moreover, whilst the DRF categorises symptoms into five factors (1) physical discomfort; 2) more alcohol, sex, active; 3) low energy; 4) consumption; 5) dysphoric mood), the DSR categorises items into four (mood; behavioural: pain; physical).

It is difficult to draw meaningful comparisons between PMS studies that have used different daily rating scales, as the format of PMS measures varies, as does the number and labelling of the anchors provided for making ratings. For example, the COPE, DSR, DRF and MDQ-T all require women to provide daily ratings on Likert scales with a different number of anchors, which range from five (DSR) to nine (MDQ-T). Whilst the MDQ-T and the DRF ask women to rate their symptoms from ‘not at all’ to ‘extreme,’ the COPE and the DSR ask women to rate their symptoms from ‘not present’ to ‘severe.’ Moreover, whilst the DRF and the MDQ-T simply provide a label next to each anchor, the COPE and the DSR provide more detailed guidance on how symptoms should be rated, by providing a category description next to each anchor. However these differ between the measures. Whilst the COPE describes its most severe anchor as relating to symptoms that are ‘intolerable and which prevent a woman from performing her normal activities,’ the DSR describes its most severe anchor as relating to symptoms that are ‘overwhelming’ and/or ‘prevent a woman from performing her usual activities.’ These differences in instructions could clearly influence the way in which women rate items on the measure. If a woman subjectively feels that her symptoms are overwhelming but not that they interfere with her usual activities, then the use of a measure like the DSR, or a measure that does not provide a category description, could result in her using the most severe anchor on the measure to
rate her symptoms. However, the use of another measure that provides a stricter category
description (e.g. DRF) may lead her not to do so. As the number of anchors that are
included on a scale and the way in which they are labelled and described all affect the way
in which items are rated (Weng, 2004), it could be considered problematic to compare the
findings from studies that have diagnosed PMS through the use of measures with anchors
that differ in this way.

Many clinicians do not ask women to rate their symptoms prospectively when diagnosing
PMS (Connoly, 2001). Those that do often do so through the administration of a chart with
ratings from each cycle day appearing on the same page, so that the clinician can easily
‘eyeball’ the data (Endicott et al., 1986). Therefore, although symptoms are rated daily, the
pattern is also evident to the women completing the ratings, which may lead to great
introspection, and symptom reporting in line with their expectations and stereotypical
beliefs.

1.3.1.4 Co-morbidities
Approximately 50% of women presenting with PMS are found not to have symptoms
limited to the premenstrual phase when symptoms are recorded prospectively (Johnson,
2004; Pearlstein, 1995; Wyatt et al., 2002). Many meet a psychiatric diagnosis, most
commonly anxiety and depression (Freeman, 2003; Johnson, 2004; Landen and Eriksson,
2003; Pearlstein, 1995). These women believe that they suffer from PMS, as their
continuing mood disorder can become magnified premenstrually, with new symptoms often
emerging at this time (APA, 1994; Halbreich and Endicott, 1985). Therefore, it is important
to confirm that PMS symptoms are not an exacerbation of another disorder (Freeman, 2003;
Frye and Silverman, 2000; Lampe, 2005; Landen and Eriksson, 2003). Although some
identify anxiety and depression, the most common co-morbid conditions, through
psychiatric assessment (De Ronchi et al., 2005; Hsiao et al., 2004; Laessle et al., 1990;
Lahmeyer et al., 1982) or through the use of standardised measures (Freeman, 2003;
Pearlstein et al., 2005), some studies have not considered whether women are suffering
from comorbid conditions (e.g. Van der Ploeg, 1987; Watts et al., 1987).
1.3.1.5 Summary of current PMS diagnostic techniques

Disparate methods have been used to diagnose PMS (Budeiri et al., 1994) and this makes the comparison of PMS studies difficult (Freeman, 2003). It is now widely accepted that symptoms should be prospectively confirmed for at least two menstrual cycles (Connolly, 2001; Freeman, 2003; Johnson, 2004; Steiner and Wilkins, 1996), and that symptoms should be limited to the luteal phase (Freeman, 2003). Although various criteria are widely used, most researchers follow the NIMH (1983) guidelines by confirming PMS in women who demonstrate at least a 30% follicular to luteal increase in symptoms for at least two menstrual cycles. However, these criteria have been defined and applied in two ways, which have been shown to result in different proportions of women meeting a PMS diagnosis (Gallant et al., 1992a). A greater consensus needs to be reached regarding the criteria that should be used to make a PMS diagnosis, the measure to which these criteria should be applied, and the way in which these criteria should be applied to items on that measure. Currently, researchers should use a method of diagnosis that identifies women who experience PMS symptoms at a severity appropriate to their particular study aim (Johnson, 2004). All currently accepted PMS criteria require daily documentation of symptoms for around 60 days, a demanding undertaking. Hence, more parsimonious diagnostic tools which can differentiate women with PMS from those who do not have it, or who have enduring psychological disorders such as anxiety or depression, and which does not by nature of its length or mode of administration encourage stereotypical reporting patterns, or fall prey to the confounding effects of boredom, would be an important contribution to the field, both clinically and as a research tool.

1.4 PMS symptoms

Debate exists as to whether the cyclical symptoms that characterise PMS represent an abnormal symptom profile, or whether they represent a variation of a normal cyclical pattern (Sanders et al., 1983). Although women with PMS suffer from a variety of symptoms, they usually report that their mood and behavioural symptoms cause them the most difficulty (Corney and Stanton, 1991; Freeman, 2003; Halbreich et al., 2007).

1.4.1 Anxiety and depression

Anxiety (Chisholm et al., 1990; Freeman, 2003; Haywood et al., 2002; Landen and Eriksson, 2003; Mortola et al., 1990; Wyatt et al., 1999; Yonkers, 1997) and depression
(Chisholm et al., 1990; Freeman, 2003; Haywood et al., 2002; Landen and Eriksson, 2003; Mortola et al., 1990; Yonkers, 1997) are premenstrual symptoms that cause PMS sufferers particular distress. The term ‘anxiety’ is used to refer to an unpleasant emotional state that is characterised by feelings of tension, apprehension, nervousness and worry, which occurs in response to a particular situation or set of circumstances (state anxiety), and to a person’s intrinsic characteristics which determines their tendency to perceive a situation as dangerous or threatening (trait anxiety) (Spielberger et al., 1983b). The term ‘depression’ refers to various affective, cognitive, behavioural and physiological symptoms (Brown et al., 1995), such as low mood, loss of pleasure, poor coordination and fatigue (Quah-Smith et al., 2005).

Some studies have assessed the anxiety and depression levels reported by PMS sufferers during the follicular and luteal phases, others have assessed the levels reported by ‘normal’ women across the cycle, while yet others have compared the levels reported by PMS sufferers with those reported by ‘control’ women. Although the studies of PMS sufferers have found levels of anxiety and depression to increase between the follicular and luteal phases, the inconsistent findings that have arisen from the studies which have assessed the other groups of women makes it unclear as to whether these cyclical profiles represent an abnormal process, or whether they represent a variation of a normal cyclical pattern. It is illuminating to provide an overview of the research that has been conducted and to discuss differences in the studies methodologies that may have contributed to these inconsistent findings. Ultimately, the aim is to establish a sound methodology by which to address the issue of what profile of anxiety and depression is characteristic of women with PMS in order to facilitate diagnosis and sample specification.

Table 1.1 describes the studies that have been conducted which have assessed depression and anxiety levels during the menstrual cycle, the majority of which have done so through the administration of self-report measures.
### Table 1.1 Studies assessing anxiety and depression levels during the menstrual cycle

<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample</th>
<th>Assessment Method</th>
<th>Outcome</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Awaritefe et al. (1980)</td>
<td>80 Nigerian females (21-30 yrs; average age=25.5) divided into a menstrual (n=40) and non-menstrual (n=40) group</td>
<td>Questionnaire: STAI. The menstrual group were asked to respond how they felt the week before and during menstruation, the non-menstrual group how they felt at that time</td>
<td>Trait anxiety did not significantly differ between groups. Menstrual participants reported significantly higher state anxiety than non-menstrual participants.</td>
<td>Details of recruitment not provided. It was not ascertained whether women met PMS criteria. Women with psychiatric conditions not excluded. STAIT instructions modified even though they should always be administered in their original format (Spielberger. 1983b). A repeated measures design was not employed.</td>
</tr>
<tr>
<td>Collins et al. (1985)</td>
<td>15 “normally cycling women” (21-36 yrs; mean age=29.5)</td>
<td>Questionnaire: STAI (1970); NMAC twice in the menstrual (day 1), follicular (5-6), ovulatory (12-14) and luteal (22-25) phases</td>
<td>State anxiety levels varied significantly across the cycle. Lowest levels were reported during the follicular/ovulatory phases, and highest during menses.</td>
<td>Unclear whether participants knew the study focus. It was not ascertained whether women self-diagnosed with PMS or met PMS criteria. Women with psychiatric disorders not excluded. Counterbalanced; women began during different phases.</td>
</tr>
<tr>
<td>Ivey and Bardwick (1968)</td>
<td>26 female college students (19-22 yrs)</td>
<td>Verbal Sample: Women talked for 5 minutes on “any memorable life experience” during ovulation (BBT) and premenstrually (-3 to -2), twice. Samples scored according to Gottschalk’s Verbal Anxiety Scale</td>
<td>Anxiety scores were significantly greater premenstrually than during ovulation.</td>
<td>Women volunteered for a “study on the menstrual cycle.” It was not ascertained whether women self-diagnosed with PMS or met criteria for PMS or other psychiatric disorders. Not fully counterbalanced; 7 women began the study premenstrually and 19 during ovulation. One of the two experimenters scoring the transcripts was not blind to the participant/phase corresponding to each sample.</td>
</tr>
<tr>
<td>Lahmeyer et al. (1982)</td>
<td>11 women (19-35 yrs; mean age=23) not complaining of premenstrual distress</td>
<td>Questionnaire: STAI daily for 1 cycle, divided into 6 phases (menstrual, follicular, ovulatory, luteal, late luteal and premenstrual)</td>
<td>State anxiety levels were highest premenstrually, but did not vary significantly across the cycle.</td>
<td>Small sample size. Volunteers were recruited for a study on the biochemical aspects of the normal menstrual cycle. No PMS diagnosis made. Women with psychiatric conditions excluded (psychiatrist).</td>
</tr>
<tr>
<td>Reference</td>
<td>Sample</td>
<td>Assessment Method</td>
<td>Outcome</td>
<td>Comments</td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------------------------------------</td>
<td>------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Veith et al.       | “Normally menstruating women” (n=9)         | Questionnaire: STAIS once in the menstrual (days 2-4), follicular (8-10), ovulatory (BBT), luteal (6-8 days after ovulation), and premenstrual (11-13 days after ovulation) phase | State anxiety levels did not vary significantly between cycle phases.      | Small sample size.  
Women volunteered for a study examining the relationship between β-endorphin and pain thresholds across the cycle.  
It was not ascertained whether women self-diagnosed with PMS nor whether they met PMS criteria.  
Women with psychiatric conditions not excluded.  
Counterbalanced; women began during each of the 5 phases. |
| Study of depression only |                                    |                                                                                     |                                                                        |                                                                                                                                          |
| Laessle et al.     | 30 young, healthy, women (24.4 (2.5) yrs)   | Questionnaire: Items ‘global mood’ and ‘depression’ (VAS) rated daily for one cycle, which was divided into 5 phases | A multivariate analysis of variance assessing ‘global mood’ and ‘depression’ together revealed that neither DV varied significantly across the cycle. | Women recruited for a study assessing the effect of dieting on menstrual function. Data was taken from the control cycles.  
No diagnosis of PMS was made.  
Women with major medical and psychiatric disorders excluded (laboratory, medical and psychiatric assessment).  
Menstrual cycle phases verified through endocrine monitoring. |
| Studies of anxiety and depression |                                    |                                                                                     |                                                                        |                                                                                                                                          |
| Golub et al.       | 50 married, parous, Caucasian women (30-45 yrs) | Questionnaire: DACL: STAI twice during the premenstrual (days -4 to -1) and intermenstrual (~day 14) phase | Depression and state anxiety significantly higher in the premenstrual than intermenstrual phase.  
Trait anxiety did not significantly vary between cycle phases.              | Menstrual cycle focus not disclosed.  
No diagnosis of PMS was made.  
Women with psychiatric disorders not excluded. |
| Moos et al.        | 15 young, nulliparous women (24.1 (2.7) yrs) | Questionnaire: NMAC twice on ~cycle days 2, 7, 14, 19 and 24-28                      | Anxiety levels were highest during menses, decreased rapidly (day 7), increased to day 26 and then decreased slightly (day 28).  
Depression levels showed no consistent pattern across the cycle.             | Women were informed that the study focus involved the menstrual cycle.  
Some women did (n=7) and did not (n=8) self-diagnose PMS.  
Women with psychiatric disorders not excluded.  
No inferential statistics performed. |
<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample</th>
<th>Assessment Method</th>
<th>Outcome</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PMS SUFFERERS/ PMS SUFFERERS V ‘CONTROLS’</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study of anxiety only</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watts et al. (1980)</td>
<td>25 PMS sufferers (37 (6.36) yrs); 23 controls (34 (6.12) yrs)</td>
<td>Questionnaire: STAI once</td>
<td>PMS sufferers reported significantly greater trait anxiety than controls. The groups did not report significantly different state anxiety levels.</td>
<td>PMS sufferers volunteered for a PMS drug trial; controls were non-patient contacts of the authors. PMS diagnosis poor: women had to have cyclically recurring symptoms, but it was not stated how this was determined. OC users/ women with psychiatric conditions not excluded. Women completed the STAI anytime during the cycle.</td>
</tr>
<tr>
<td>Study of depression only</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>De Ronchi et al. (2005)</td>
<td>N=128 (32.4 (8.2) yrs); 43 women with PMDD (DSM-III-R); 85 without</td>
<td>Questionnaire: MADRAS once during the follicular phase</td>
<td>PMDD sufferers reported significantly greater depression levels than those without PMDD, as measured by the total score and all individual items.</td>
<td>PMDD sufferers were seeking medical treatment for PMS, but it was unclear how women without PMDD were recruited. PMDD diagnosis prospectively confirmed (MHQ). Premenstrual severity increase assessed by effect size method. Women with psychiatric disorders excluded (SCID-NP, 1992). Women only assessed in follicular phase, but which cycle days was not specified.</td>
</tr>
<tr>
<td><strong>Studies of anxiety and depression</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chisholm et al. (1990)</td>
<td>40 women (mean age=35.2) reporting moderate-extreme premenstrual increases in anxiety and depression</td>
<td>Questionnaire: IPAT and STAI once during the intermenstrual (days 5-10) and premenstrual (-6 to -1) phase</td>
<td>Depression, state and trait anxiety were significantly greater during the premenstrual than the intermenstrual phase. Follicular depression scores were bimodally distributed. Therefore women scoring above and below 25 pts were analysed separately. The low scoring group demonstrated the pattern of results described above, while the high scoring group reported stable levels.</td>
<td>Unclear whether women knew the study focus. Women included if they retrospectively reported moderate-extreme depression &amp; anxiety premenstrual increases (PAF). Women who had consulted a psychologist/ psychiatrist or had been diagnosed with a psychological disorder excluded. Not fully counterbalanced. The initial analysis included women with high follicular depression &amp; anxiety scores. Women scoring low/ high follicular depression were then analysed separately, but women (n=6) with high follicular anxiety (STAI &gt;51.9) not excluded.</td>
</tr>
<tr>
<td>Reference</td>
<td>Sample</td>
<td>Assessment Method</td>
<td>Outcome</td>
<td>Comments</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Haskett et al., 1984 | 42 women with moderate/severe emotional and physical PMS symptoms (22-42 yrs) | Questionnaire: MAACL and STAIS once in the follicular (day 9) and late-luteal (days -7 to -3) phase | Depression and anxiety (MAACL), and state anxiety (STAIS) scores, were significantly higher in the late luteal than follicular phase. | Women informed the study required women with severe PMS.  
PMS diagnosis unclear; women completed the MDQ-T in the follicular and luteal phases for 2-4 cycles, but it was not stated whether scores were used as exclusion criteria.  
Women with psychiatric disorders excluded (follicular assessment using the RDC). |
| Hsiao et al. (2004) | 43 Taiwanese PMDD (DSM-III-R) sufferers (30.79 (7.13) yrs)             | Questionnaire: Psychiatric assessment of depression and anxiety in the follicular (days 7-11) and luteal (-6 to -2) phases using the HAM-D and HAM-A. STA1 also administered daily 1 cycle. | Depression and anxiety levels assessed by the psychiatrist, and state and trait anxiety (STA1), were greater in the luteal than follicular phase. | Women recruited through seeking medical treatment for PMS.  
Sound PMDD diagnosis, confirmed prospectively. A 50% follicular to luteal increase was required (PRISM).  
Women with psychopathology excluded (psychiatrist).  
No inferential statistics performed.  
Only PMDD sufferers were studied. |
| Morse et al. (1988) | PMSS sufferers (n=75); Lo-volunteers (n=13); High-volunteers (n=19)    | Questionnaire: BDI and STAIS twice in the follicular (days 6-8) and luteal (26-28) phases. Analyses were conducted on data from the 2nd cycle. | There were no significant differences in the depression scores reported between groups during the follicular phase, but PMSS sufferers reported significantly greater depression levels than the Hi- and Lo-volunteers during the luteal phase.  
There were no significant differences in the state anxiety levels reported between groups during the follicular or luteal phases. | PMSS sufferers were women seeking PMS treatment and who met PMS criteria. Hi- and Lo-volunteers did not self-diagnose with PMS but did and did not meet PMS criteria respectively.  
Women met PMS criteria if they showed a >30 unit increase between the follicular and luteal phases of 2 cycles (MDQ).  
Women with psychiatric disorders excluded, but how?  
Follicular to luteal changes in anxiety not assessed. |
| Reed et al. (2008)  | 14 women (30 (6.7) yrs) with PMDD (DSM-IV); 15 without (30 (6.1) yrs) | Questionnaire: BDI-II and STAIS twice in the luteal (days -5 to -1) and follicular (6-10) phases | Both groups reported significantly greater depression levels in the luteal than follicular phase. Only PMDD sufferers reported significantly greater state anxiety in the luteal than follicular phase.  
PMDD sufferers reported significantly greater depression and state anxiety than controls in the luteal phase. | All women volunteered for research requiring PMS sufferers.  
PMDD diagnosis prospectively confirmed (DRF). A 30% follicular to luteal symptom increase criteria was adopted.  
Women with psychiatric conditions excluded (SCID).  
Counterbalanced; women began during different phases. |
<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample</th>
<th>Assessment Method</th>
<th>Outcome</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Van der Ploeg (1987)</td>
<td>317 women with PMS; 385 without</td>
<td>Questionnaire: DACL and STAIS once on cycle days 12, 18, 22 and 26</td>
<td>PMS sufferers reported greater levels of depression and state anxiety across the cycle than women without PMS. PMS sufferers reported a luteal increase in depression and state anxiety, while women without PMS reported more stable levels across the cycle.</td>
<td>All women responded to advertisements for women suffering from menstrual tension and reported that most of their symptoms occurred premenstrually or during menses. PMS diagnosis poor; women divided into those probably/probably not suffering from PMS (MDQ-T) but it was not stated how this was done. Women with psychiatric conditions not excluded. No inferential statistics performed.</td>
</tr>
</tbody>
</table>

**KEY:** BBT = Basal body temperature; BDI = Beck Depression Inventory (Beck et al., 1961); DACL = Depression Adjective Check List (Lubin, 1967); DRF = Daily rating form (Endicott et al., 1986); HAM-A = Hamilton rating scale of anxiety (Hamilton, 1959); HAM-D = Hamilton rating scale for depression (Hamilton, 1960); IPAT = Institute of Personality and Ability Testing Depression Scale (Krug and Laughlin, 1976); MAACL = Multiple Affect Adjective Checklist (Zuckerman and Lubin, 1965); MADRAS = Montgomery and Asberg (1979) Depression Rating Scale; MDD = Major depressive disorder; MDQ = Menstrual Distress Questionnaire (Moos, 1968); MDQ-T = Menstrual Distress Questionnaire (Moos et al., 1969); MHQ = Menstrual Health Questionnaire (Warner et al., 1991); NMAC = Mood Adjective Check List (Nowlis, 1965); OC = oral contraceptive; PAF = Premenstrual Assessment Form (Halbreich et al., 1982b); PRISM = Prospective Record of the Impact and Severity of Menstrual Symptomatology (Reid, 1985); RDC (Spitzer et al., 1977); SCID = Structured Clinical Interview for DSM-IV (First et al., 1995); SCID-NP = Structured Clinical Interview for DSM-III-R non-patient version (Spitzer et al., 1983, 1992); STAI = State trait anxiety inventory (Spielberger et al., 1983b); STAIS = state anxiety; STAIT = trait anxiety; STAIS (1970) = State anxiety scale (Spielberger et al., 1970)
1.4.1.1 Depression, anxiety and PMS sufferers

Four studies assessed depression and state anxiety levels during the follicular and luteal phases in women meeting PMS criteria (Chisholm et al., 1990; Hsiao et al., 2004; Reed et al., 2008; Van der Ploeg, 1987). These studies consistently found PMS sufferers to report greater depression and state anxiety levels during the luteal than follicular phase. Hsiao et al. (2004) also assessed trait anxiety levels and found this pattern of symptom reporting. Although Hsiao et al. (2004) and Van der Ploeg (1987) only discussed mean values and did not perform any inferential statistics, the consistency between findings suggests that PMS sufferers experience greater depression and state anxiety premenstrually than during the follicular phase. However, no studies assessed the levels of depression, state or trait anxiety reported by PMS sufferers during any cycle phase other than during the follicular and luteal phase.

1.4.1.2 Depression, anxiety and ‘normal’ women

The studies that assessed depression and anxiety levels across the cycle of ‘normal’ women have produced inconsistent findings (see Table 1.1). Three studies specifically assessed depression levels. Two of these found levels not to vary between cycle phases (Laessle et al., 1990; Moos et al., 1969), while the third found levels to be significantly greater during the luteal than intermenstrual phase (Golub et al., 1976). Two studies assessed trait anxiety, and neither found levels to vary between cycle phases (Awaritefe et al., 1980; Golub et al., 1976). Seven studies assessed state anxiety. Four of these found levels to vary significantly between cycle phases (Awaritefe et al., 1980; Golub et al., 1976; Ivey and Bardwick, 1968; Collins et al., 1985), two did not (Lahmeyer et al., 1982; Veith et al., 1984), while one could not ascertain this information as no inferential statistics were performed (Moos et al., 1969). Three of the four studies which found state anxiety to vary found the greatest levels to occur premenstrually (Awaritefe et al., 1980; Golub et al., 1976; Ivey and Bardwick, 1968). However, these studies did not assess state anxiety during menses. The only study which did (Collins et al., 1985) found levels to be greatest during this phase.

These inconsistent findings may partly result from methodological weaknesses present in the studies. Awaritefe et al. (1980) modified the instructions of the STAIT, even though this is not recommended by the original authors (Spielberger et al., 1983b). One of the two experimenters scoring the verbal samples in Ivey and Bardwick’s (1968) study was not
blind to the participant’s phase. Lahmeyer et al. (1982) had a small sample size (n=11), Laessle et al. (1990) only assessed the individual items ‘depression’ and ‘global mood,’ while Moos et al. (1969) did not perform any inferential statistics. However, more importantly, anxiety and depression levels were assessed through the administration of different measures, and disparate methods were employed to select women for the studies, which possibly resulted in heterogeneous samples of women being studied.

None of the studies determined whether their participants suffered from PMS. As depression and anxiety are core PMS symptoms (Chisholm et al., 1990; Freeman, 2003; Haywood et al., 2002; Landen and Eriksson, 2003; Mortola et al., 1990; Wyatt et al., 1999; Yonkers, 1997), it is likely that women with and without PMS display different profiles of depression and anxiety across the cycle. Moreover, all but one study (Lahmeyer et al., 1982) did not exclude women suffering from psychiatric conditions. As discussed in Section 1.3.1.4, women suffering from depression and anxiety disorders often experience a premenstrual magnification of symptoms (APA, 1994; Halbreich and Endicott, 1985). Therefore, the inconsistencies in findings may partly be explained by the samples that were studied comprising different proportions of women suffering from PMS, anxiety and depression.

Women report significantly more negative psychological and somatic premenstrual symptoms if they know the study focus involves menstrual cycle symptomatology (AuBuchon and Calhoun, 1985). The studies differed in the information that was provided to their participants. Some researchers informed their participants that the study focus involved the menstrual cycle (Ivey and Bardwick, 1968; Moos et al., 1969). Others disclosed the menstrual cycle focus, but not that changes in mood were being assessed (Laessle et al., 1990; Lahmeyer et al., 1982; Veith et al., 1984). One study disguised its aims (Golub et al., 1976), while others did not specify the nature of the information provided to their participants (Avaritefe et al., 1980; Collins et al., 1985). Therefore, the inconsistencies in the findings may partly be explained by the differential information provided to participants resulting in demand characteristics affecting women’s responses more in some studies than others.
Demand characteristics only increase cyclicity of symptom reporting in women who do not consider themselves to have problematic PMS symptoms (Gallant et al., 1992b). Moos et al. (1969) included both women who did and who did not self-diagnose with PMS, while Lahmeyer et al. (1982) only included women who did not consider themselves to suffer from PMS. However, the majority of studies did not ascertain whether women considered themselves to be PMS sufferers (Awaritefe et al., 1980; Collins et al., 1985; Golub et al., 1976; Ivey and Bardwick, 1968; Laessle et al., 1990; Veith et al., 1984). Therefore the inclusion of women self-diagnosing with PMS in some studies and not others may also contribute to the inconsistencies present in their findings.

1.4.1.3 Depression, anxiety, PMS sufferers and controls

Studies which compared depression, trait and state anxiety levels reported by PMS sufferers with those of ‘control’ women also produced inconsistent findings. Although two studies assessing depression found PMS sufferers to report significantly greater levels than controls in the luteal, but not follicular phase (Morse et al., 1988; Reed et al., 2008), another study found PMDD sufferers to report significantly greater levels than controls in the follicular phase (De Ronchi et al., 2005). Two studies assessed trait anxiety. One of these found PMS sufferers to report significantly greater levels than controls during the follicular and luteal phases (Morse et al., 1988), while the other did not find the groups to differ significantly (Watts et al., 1980). However, these researchers only assessed trait anxiety once, at no particular point during the menstrual cycle. Four studies assessed state anxiety. Two of these did not find PMS sufferers to report significantly different levels to controls (Morse et al., 1988; Watts et al., 1980). Although Morse et al. (1988) assessed levels during the follicular and luteal phases, Watts et al. (1980) only assessed levels once, at no particular point during the cycle. Two studies found differences between groups (Reed et al., 2008; Van der Ploeg, 1987). Reed et al. (2008) found PMS sufferers to report significantly greater levels than controls premenstrually, while Van der Ploeg (1987) found this group to report consistently greater levels than controls across the cycle. However, no inferential statistics were performed.

These inconsistent results may partly be explained by the diverse samples of PMS sufferers and controls that were studied. Although the majority of studies ensured that control women did not meet PMS criteria, one study (Watts et al., 1980) did not state whether this
was the case, and it is therefore possible that this study included PMS sufferers in this group. Control groups in some studies comprised women who self-diagnosed with PMS (Reed et al., 2008; Van der Ploeg, 1987), while in others only included women who did not (Morse et al., 1988). This may have contributed to the inconsistent results, as it is reasonable to hypothesize that women who do and who do not believe they have PMS experience and report PMS symptoms differently from each other across the cycle.

Heterogeneous samples of PMS sufferers were studied due to the considerably different PMS diagnostic criteria that were used. Three studies used sound diagnostic methods (De Ronchi et al., 2005; Morse et al., 1988; Reed et al., 2008). Although they all excluded women with co-morbidities and prospectively confirmed symptoms, they used different criteria to determine the premenstrual severity increase required to confirm PMS (see Table 1.1). Other studies used poor diagnostic criteria. Van der Ploeg (1987) simply divided women into those who were ‘probably’ and ‘probably not’ suffering from PMS. Although they administered the MDQ-T, they did not state how this was used to allocate women to groups. Watts et al. (1980) only required women to have ‘cyclically recurring symptoms,’ but did not state how this was ascertained. Moreover, Van der Ploeg (1987) and Watts et al. (1980) did not exclude women with co-morbidities. Therefore some of the women in their research may not have been suffering from PMS, but from a cyclical exacerbation of anxiety or depression (Freeman, 2003; Frye and Silverman, 2000; Lampe, 2005; Landen and Eriksson, 2003).

Moreover, the studies assessed anxiety and depression levels through the use of different assessment methods. For example, some studies assessed anxiety levels through the use of Spielberger et al.’s (1983b) STAI (e.g. Awaritefe et al., 1980; Collins et al., 1985; Golub et al., 1976; Lahmeyer et al., 1982; Veith et al., 1984), which differentiates anxiety from depression well (Cox et al., 1993; Endler et al., 1991; McWilliams et al., 2001), while another study (Hsiao et al., 2004) did so through the use of Hamilton’s (1959) rating scale of anxiety, which includes items that are applicable to both anxiety and depression (Snaith and Taylor, 1985). Moreover, one study assessed depression levels through the use of Hamilton’s (1960) rating scale of depression, while another (Morse et al., 1988) did so through the use of the BDI (Beck et al., 1961). These measures assess different aspects of depression. While the BDI mostly assesses cognitive and behavioural symptoms of
depression, the HAM-D assesses somatic and behavioural symptoms (Brown et al., 1995). Therefore, the inconsistencies in the findings may partly be explained by the measures that were used tapping different aspects of anxiety and depression.

1.4.1.4 Summary of depression and anxiety in PMS

Although PMS sufferers appear to experience an increase in depression and anxiety between the follicular and luteal phases, the inconsistent findings from the studies which assessed levels of anxiety and depression in other groups makes it unclear as to whether these cyclical changes represent an abnormal profile, or whether they represent a variation of a normal cyclical pattern. To address this issue, further research is needed to prospectively assess anxiety and depression levels across the cycle of women with and without PMS during a greater number of cycle phases, using validated measures that are sensitive to the aspects of anxiety and depression that women with PMS find problematic. PMS sufferers should be identified through the use of sound diagnostic criteria, prospective confirmation of symptoms, and exclusion of women with psychiatric conditions. These women should be compared with a ‘control’ group who do not self-diagnose with PMS, and do not meet PMS criteria or criteria for anxiety and/ or depression. Research assessing the depression and anxiety profiles reported by women with co-morbidities, and by women who do and do not self-diagnose with PMS could shed light on whether the inclusion of these women in some studies and not others may have contributed to the inconsistencies present in their findings.

1.4.2 Aggression and Impulsivity

PMS sufferers often report that impulsivity (Elliott, 2002; Halbreich, 2003; Halbreich et al., 2007; Hallman et al., 1987; Hsu et al., 2007; Matsumato et al., 2007) and aggression (Endicott et al., 1999; Hylan et al., 1999; Landen and Eriksson, 2003; Lurie and Borenstein, 1990; Warner and Bancroft, 1990) are among their most distressing symptoms. Moreover, women maintain that these symptoms premenstrually are among the symptoms which lead them to seek treatment (Hartlage and Arduino, 2002). However, the variability of aggressive and impulsive symptoms across the menstrual cycle has been the subject of little research, and these symptoms are often not included on PMS diagnostic measures (e.g. Freeman et al., 1996; Moos, 1968). This section will provide an overview of the research that has been conducted and will outline a suitable methodology by which to determine
whether women with and without PMS display different impulsivity and aggression profiles across the cycle. If this is found to be the case, then the inclusion of items relating to these symptoms on PMS scales could aid the diagnosis of PMS and the assessment of treatment efficacy.

1.4.2.1 Impulsivity

Researchers have defined impulsivity in various ways. For example, Dickman (1990) considers the term impulsivity to refer to the tendency to act with less deliberation than most people of equal ability before taking action, while Eysenck et al. (1984) considers the same term to characterise people who act quickly without being mindful of the risk involved. Although most agree that impulsivity is a multidimensional construct (Dawe and Loxton, 2004; Miller et al., 2004), researchers who have constructed measures to assess impulsivity levels have not agreed on the number and content of the dimensions that should be included (e.g. Dickman, 1990; Patton et al., 1995). Therefore, difficulty arises when attempting to measure impulsivity in research settings, as it is unclear whether different measures tap the same construct (Miller et al., 2004). As PMS sufferers report that impulsivity causes them distress, which people in everyday life consider to refer to acting rashly, without forethought (Dawe and Loxton, 2004), the Barratt Impulsiveness Scale, BIS-11 (Patton et al., 1995, see also section 3.1.1.4), may reflect an appropriate method by which to assess impulsivity levels in this population, as this measure includes a motor impulsivity subscale, which refers to acting without prior thought.

Few studies have been conducted which have assessed impulsivity levels during the menstrual cycle (see Table 1.2).
<table>
<thead>
<tr>
<th>Ref</th>
<th>Sample</th>
<th>Assessment Method</th>
<th>Outcome</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hallman et al.</td>
<td>34 controls (28 (7) yrs); 40 PMS (36 (7) yrs); i. irritability &amp;or</td>
<td>Questionnaire: EPQ, IVE and KSP once</td>
<td>No significant differences between groups on the EPQ, IVE or the impulsiveness scale of the KSP.</td>
<td>PMS sufferers were seeking treatment; controls did not self-diagnose with PMS.</td>
</tr>
<tr>
<td></td>
<td>hostility; ii. depression as main symptom</td>
<td></td>
<td></td>
<td>PMS diagnosis (physician) not prospectively confirmed Women with ‘psychiatric disease’ excluded, but how? Impulsivity measures administered at any time during cycle.</td>
</tr>
<tr>
<td>Howard et al.</td>
<td>7 females without PMS (22-31 yrs)</td>
<td>Laboratory: Go/ No Go Avoidance Task once in the menstrual, early mid-cycle,</td>
<td>CNV differentiation lowest premenstrually, highest at mid-cycle, intermediate at menses. CNV differentiation significantly lower premenstrually than in all other phases combined.</td>
<td>Details of recruitment not provided.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>late mid-cycle &amp; luteal phase</td>
<td></td>
<td>Women were “not suffering from PMS” but no confirmation was made.</td>
</tr>
<tr>
<td>Howard et al.</td>
<td>30 women with PMDD (DSM-III-R) divided into a high (33.6 (6.8) yrs) and</td>
<td>Laboratory: Go/ No Go Avoidance Task once in the menstrual (days 0-5), late</td>
<td>Non-significant main effect of group, phase or phase by group interaction for CNV differentiation. Groups did not significantly differ in trait impulsiveness (MMPI). Groups significantly differ in state impulsiveness premenstrually (PAF); highly impulsive PMS group reported the highest levels and controls the lowest. Low impulsive PMS and control groups reported stable impulsivity across the cycle, the high impulsive group a premenstrual increase (DSR).</td>
<td>PMDD sufferers responded to advertisements requiring women with premenstrual problems, but it was unclear how controls were recruited. PMDD diagnosis prospectively confirmed (DRF) although it was not clear what criteria were used. PMDD sufferers divided into high/low impulsive groups on the basis of their PAF and DSR scores. Women with major psychiatric illness were excluded (psychiatric interview during the premenstrual phase). Counterbalanced; women began during different phases.</td>
</tr>
<tr>
<td></td>
<td>low (34.4 (6.8) yrs) impulsive group; 24 controls without PMDD (30.5 (9.1) yrs)</td>
<td>late follicular (3 days preceding/including mid-cycle), and late luteal (-5 to 0) phase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsu et al.</td>
<td>51 PMDD (; 39 MDD (MINI); 52 without PMDD/MDD)</td>
<td>Questionnaire: TPQ (Chinese version) once in the luteal phase (days 23-28)</td>
<td>PMDD sufferers and controls did not differ significantly in their levels of novelty seeking.</td>
<td>Details of recruitment not provided.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PMDD diagnosis (DSM-IV) prospectively confirmed for PMDD but not controls group.</td>
</tr>
<tr>
<td>Keenan et al.</td>
<td>14 women with PMS (SCID-NP (1983): 10 without</td>
<td>Laboratory: Stroop Colour Interference Task twice in follicular (days 7-9) and</td>
<td>Both groups made significantly more errors and demonstrated significantly faster RTs in the luteal than follicular phase. Non-significant difference between groups in error rate/ RT patterns across the cycle.</td>
<td>Menstrual cycle focus not disclosed.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>luteal (-4 to -2) phases</td>
<td></td>
<td>PM diagnosis was prospectively confirmed (PRISM). Women with psychiatric disorders only excluded if psychotropic medication was required.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Counterbalanced; women began in different phases.</td>
</tr>
</tbody>
</table>

**Key:** DRF = Daily rating form (Endicott et al., 1986); EPQ = Eysenck Personality Questionnaire (Eysenck et al., 1985a); IVE = Eysenck’s IVE inventory (Eysenck and Eysenck, 1978); KSP = Karolinska Scales of Personality (Schalling and Edman, 1993); MDD = Major depressive disorder; MMPI = Minnesota Multiphasic Personality Inventory (Hathaway and McKinley, 1940); PAF = Premenstrual Assessment Form (Halbreich et al., 1982b); RT = reaction time; TPQ = Tridimensional Personality Questionnaire (Chen et al., 2002)
Only two studies compared impulsivity levels self-reported by women with and without PMS and found levels not to significantly differ between groups (Hallman et al., 1987; Hsu et al., 2007). However, both studies only assessed impulsivity levels once and Hallman et al. (1987) did so at no particular point during the cycle. Moreover their diagnosis of PMS was poor; symptoms of PMS sufferers were not prospectively confirmed, while controls were not examined to confirm that their symptom reports did not meet the diagnostic criteria for PMS. Hsu et al. (2007) also did not prospectively confirm the absence of symptoms in their control group.

Three studies assessed impulsivity levels across the cycle using proxy task based measures. Keenan et al. (1992) assessed impulsivity levels in women with and without PMS during the follicular and luteal phases using the Stroop Colour-Interference Task (Stroop, 1935). Participants are shown single colour words printed in contrasting colours and are asked to name the colours in which the words are printed. This task measures the difficulty participants have in changing perceptual set to conform to the changing demands (naming the colour) and to inhibit attention to the competing stimuli (reading the word). Both women with and without PMS made significantly more errors but had significantly faster reaction times during the luteal than follicular phase. This suggests that both women with and without PMS demonstrate a degree of disinhibition and impulsivity premenstrually.

Howard et al. (1988) assessed impulsivity levels in women without PMS using the Go/No Go avoidance task during various cycle phases. This task requires participants to press a button to the second of two successive tones in the ‘Go’ condition, and refrain from pressing to the second of the tones in the ‘No Go’ condition. Electrodes are attached to the scalp to measure the Contingent Negative Variation (CNV), an electric event indicative of a state of readiness. Highly impulsive patients have been shown to demonstrate a smaller difference between the CNVs elicited in the ‘Go’ and ‘No Go’ conditions (Brown et al., 1989). Howard et al. (1988) found that women in their study demonstrated significantly lower CNV differentiation in the premenstrual week than in all of the other weeks combined. However, their sample size was extremely small (n=7) and these findings were not replicated in women with or without PMS when a larger sample size was used (Howard et al., 1994). As they found that a sub-group of their PMDD participants self-reported significantly increased impulsivity levels premenstrually, but did not show increased
impulsivity using this task, they concluded that this may not be the most suitable method of assessing impulsivity levels in women.

The measures that have been used to assess impulsivity levels during the menstrual cycle call into question whether these studies have actually measured the aspects of the multidimensional construct of impulsivity that women with PMS find distressing. For example, Hsu et al. (2007) assessed novelty seeking, which refers to the tendency to actively respond to novel stimuli in pursuit of rewards and avoidance of punishment (Hansenne et al., 1997). Although novelty seeking taps an impulsivity factor, and highly correlates with impulsiveness (Weyers et al., 1995), this measure does not assess the tendency to act rashly, without forethought. Nor do the proxy task based measures that have been used, which simply assess participants’ reaction times (Keenen et al., 1992) and state of readiness (Howard et al., 1988; 1994) to cognitive tasks.

1.4.2.2 Aggression
Aggression has been defined in various ways (Barratt et al., 1999; Caprara et al., 1985; Parrott and Giancola, 2007). However, most agree that this term should be used to refer to behaviours that are carried out with the intention of harming another person, who is motivated to avoid that act (Anderson and Bushman, 2002; Baron and Richardson, 1994; Parrott and Giancola, 2007). Aggressive behaviours can be physical or verbal in nature (Anderson and Bushman, 2002; Archer, 2004; Buss and Perry, 1992; Parrott and Giancola, 2007). Although the terms ‘anger’ and ‘hostility’ are often used interchangeably with ‘aggression,’ although related, should be considered conceptually distinct (Parrott and Giancola, 2007). Anger refers to an emotional state comprising of feelings of varying intensity, from mild irritation to intense fury (Spielberger et al., 1983a). Hostility refers to a persistent negative evaluation of the environment and others (Buss, 1961). Hostility and anger motivate and prepare an individual for an aggressive act (Spielberger, 1988; Buss and Perry, 1992 respectively).

1.4.2.2.1 Aggression, crime and the menstrual cycle
The relationship between crime and aggression and the menstrual cycle is widely cited (e.g. Dalton, 1961; D’Orban and Dalton, 1980) as evidence that aggression levels increase during menses and the luteal phase (e.g. Dougherty et al., 1997b; Ritter, 2003; Singh et al.,
2004). Dalton (1961) asked newly convicted female prisoners when their last period began and estimated what menstrual cycle phase they were in when they committed their crime. She reported that 49% of the 156 women interviewed, committed their crime in the paramenstrum (4 days before menses to 4 days after) in comparison to the 29% chance level. However, a major criticism of such studies is that women do not accurately recall cycle dates from memory (Matsumoto et al., 1962) and researchers are often not explicit about how they calculated which phase women were in when they committed their crime, and whether they accounted for cycle length variability (Horney, 1978). Also, other explanations could account for these findings, such as the stress of entering prison inducing early menstrual onset (Horney, 1978), since psychological stress, including exam stress (Dalton, 1968) and surgery (Matsumoto et al., 1968), have been linked with precipitating menstruation.

Other studies have tracked female inmates and reported that they engage in the majority of aggressive acts during menses or premenstrually (e.g. Ellis and Austin, 1971; Mortan et al., 1953). Ellis and Austin (1971) asked female inmates to report information about their overt aggressive acts daily along with dates of menstruation. They reported that 41% of aggressive acts occurred during the premenstrual and menstrual phases. However, prospective monitoring of cycle dates can lead to biases in symptom reporting (AuBuchon and Calhoun, 1985). Women who believe that their behaviour is influenced by menstrual cycle phase have been found to exaggerate behavioural changes (Gallant et al., 1992b). Also, studies of women who live in close proximity may be influenced by the women experiencing menstrual synchrony (McClintock, 1971). This may lead to inflated aggression estimates, as aggressive acts instigated by a few women affected by the menstrual cycle may result in retaliation from others who are in the same cycle phase from coincidence or pheromone influence (Horney, 1978).

1.4.2.2.2 Questionnaire and laboratory assessment

Few studies have assessed aggression levels across the menstrual cycle. The research that has been conducted has assessed levels of aggression using a variety of self-report measures and laboratory assessment techniques (see Table 1.3).
<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample</th>
<th>Assessment Method</th>
<th>Outcome</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>‘Normal’ women</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collins et al. (1985)</td>
<td>15 “normally cycling women” (21-36 yrs; mean age=29.5)</td>
<td>Questionnaire: STAIS (1970) and NMAC twice during the menstrual (day 1), follicular (5-6), ovulatory (12-14) and luteal (22-25) phases</td>
<td>Aggression levels did not significantly vary across the cycle (NMAC). Anger levels significantly varied between cycle phases (STAIS). The lowest scores were reported in the follicular and ovulatory phases, the highest during menses.</td>
<td>Details of recruitment not provided. It was not ascertained whether women self-diagnosed with PMS or met PMS criteria. Women with psychiatric disorders not excluded. Counterbalanced; women began during different phases.</td>
</tr>
<tr>
<td>Moos et al. (1969)</td>
<td>15 young, nulliparous women (24.1 (2.7) yrs)</td>
<td>Questionnaire: NMAC twice on ~ cycle days 2, 7, 14, 19 and 24-28</td>
<td>Aggression levels were highest during the menstrual (day 2) and post-ovulatory (day 14) phases.</td>
<td>Data from women self-diagnosing (n=7) and not self-diagnosing (n=8) with PMS were analysed together. Women with psychiatric disorders not excluded. No inferential statistics were performed.</td>
</tr>
<tr>
<td>Ritter (2003)</td>
<td>29 female (21.6 (2.45) yrs) and 22 male (22.3 (2.11) yrs) undergraduates</td>
<td>Questionnaire: BPAQ once in the menstrual (cycle days 2-4) and mid-luteal (~10 to -5) phase</td>
<td>Significantly greater physical aggression reported during the menstrual than the midfollicular phase. This trend was found for verbal aggression. Men reported significantly greater physical aggression than women in the midluteal but not menstrual phase, and showed this trend for verbal aggression.</td>
<td>Details of recruitment not provided. No PMS diagnosis was made. Women tested outside target days excluded from the analyses. Aggression levels not assessed during the luteal phase.</td>
</tr>
<tr>
<td><strong>PMS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond et al. (2003)</td>
<td>24 PMDD sufferers (34.5 (5.7) yrs); 18 controls without PMDD (31.7 (6.4) yrs)</td>
<td>Questionnaire: CTS once during the follicular and luteal phase</td>
<td>PMDD sufferers used significantly more physical and verbal aggression during the luteal than follicular phase. Controls showed the same, but non-significant pattern. There was a trend for PMDD sufferers to use more verbal aggression than controls during the luteal, but not follicular phase.</td>
<td>Recruitment methods advertised for healthy volunteers; women suffering from severe premenstrual symptoms. PMDD diagnosis made through the SCID and prospectively confirmed, although unclear with which daily rating scale. Women with psychiatric disorders excluded (SCID). Not clear which cycle days comprised the follicular phase.</td>
</tr>
<tr>
<td>Reference</td>
<td>Sample</td>
<td>Assessment Method</td>
<td>Outcome</td>
<td>Comments</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Dougherty et al.  | 11 women with low (24.1 (7.4) yrs) and 11 with high (28.1 (8.6) yrs)  | Laboratory: PSAP once during the menstrual (days 1-3), midfollicular (7-9), ovulatory (12-16) and premenstrual (-4 to -1) phase | Non-significant effect of cycle phase on aggressive responding when groups were analysed together and separately. High symptom group responded more aggressively across the cycle than the low symptom group. | Recruitment methods advertised for “paid behavioural research study volunteers.”
Women retrospectively allocated to groups on the basis of their menstrual or premenstrual MDQ score.
Women with psychiatric disorders excluded (psychiatrist).
Women tested outside target days excluded from the analyses. |
| Haskett et al., 1984 | 42 women (22-42 yrs)                                                   | Questionnaire: MAACL once in the follicular (day 9) and late-luteal (days -7 to -3) phase | Mean hostility score significantly greater during the late-luteal than follicular phase.          | Women volunteered for a study for severe PMS sufferers and reported moderate/ severe emotional and physical symptoms.
Method of PMS diagnosis unclear.
Women with psychiatric disorders excluded (RDC). |
| Van der Ploeg     | PMS (n=317) and non-PMS (n=385) participants                          | Questionnaire: The Dutch adaptation of the MDQ (Van der Ploeg, 1983) and STAXI once on cycle days 12, 18, 22 and 26 | Women with PMS reported greater levels of aggression across the cycle than women without PMS.
Women with PMS reported a luteal increase in aggression, while women without PMS reported more stable levels across the cycle. | All participants volunteered for a study requiring women with menstrual tension symptoms and reported that most of their symptoms occurred premenstrually or during menses.
The researchers divided women into those probably/probably not suffering from PMS using the MDQ but did not state the method by which this was done.
No inferential statistics were performed. |
| Van Goozen et al. | 20 PMS reporters (mean age 29.4 yrs); 18 intermediate reporters (29.2 yrs) and 20 non-reporters (28.1 yrs) | Laboratory: Anger induction technique (Van Goozen et al., 1994) once in the luteal phase (days -4 to -2) for the PMS reporters/ non-reporters and once in the follicular phase (5-9) for the intermediate reporters | Women tested premenstrually (PMS reporters and non-reporters) responded significantly more angrily than those tested in the follicular phase (intermediate reporters). This pattern was stronger for the PMS reporters. | Volunteers recruited for research assessing the relationship between hormones and mood during the menstrual cycle i.e. unaware of true study purpose.
Women were allocated to groups on the basis of their perception of premenstrual symptom severity.
PMS diagnosis not confirmed.
Different women performed the task during the follicular and luteal phases. |

**Key:** BPAQ = The Aggression Questionnaire (Buss and Perry, 1992); CTS = Conflict Tactics Scale (Straus, 1979); MAACL = Multiple Affect Adjective Checklist (Zuckerman and Lubin, 1965); MDQ = Menstrual Distress Questionnaire (Moos, 1968); MAC = Mood Adjective Check List (Nowlis, 1965); PSAP = The Point Subtraction Aggression Paradigm (Cherek, 1981); RDC (Spitzer et al., 1977); SCID = Structured Clinical Interview for DSM-IV (First et al., 1995); STAI = State Anxiety Scale (Spielberger et al., 1970); STAXI = State-Trait Anger Inventory (Spielberger et al., 1983a)
1.4.2.2.1 Aggression levels and PMS sufferers

Two studies compared aggression levels of PMS sufferers with those of ‘control’ women during various cycle phases (Dougherty et al., 1998; Van der Ploeg, 1987), and both found PMS sufferers to demonstrate greater aggression levels across the cycle. Although Dougherty et al. (1998) found aggressive responding to remain stable between cycle phases in women reporting low and high levels of symptoms in the menstrual and/or luteal phases, Van der Ploeg (1987) found PMS sufferers reported an increase in anger premenstrually, but did not perform any inferential statistics. Three studies assessed aggression levels during the follicular and luteal phases (Bond et al., 2003; Haskett et al., 1984; Van Goozen et al., 1996). PMS sufferers were found to report significantly greater levels of hostility (Haskett et al., 1984), physical and verbal aggression (Bond et al., 2003) during the luteal than follicular phase. Women without PMDD were also found to demonstrate a luteal increase in physical and verbal aggression (Bond et al., 2003), although their differences between cycle phases did not reach statistical significance. Van Goozen et al. (1996) found that women who did and who did not self-diagnose with PMS responded significantly more angrily to an anger induction technique during the luteal phase than women who reported an intermediate level of premenstrual symptoms who were tested in the follicular phase. This pattern of results was stronger for the women who self-diagnosed with severe PMS symptoms.

1.4.2.2.2 Aggression levels and ‘normal’ women

Two studies assessed aggression levels in ‘normal’ women at various points during the cycle. Collins et al. (1985) did not find aggression levels to vary significantly between cycle phases. Although Moos et al. (1969) found levels to be highest during the menstrual and post-ovulatory phases, he did not perform any inferential statistics to determine whether levels differed significantly between phases. Two studies assessed anger. Collins et al. (1985) found anger levels to vary significantly between cycle phases. The greatest levels were reported during menses, the lowest during the ovulatory and follicular phases. Although Ritter (2003) did not find anger levels to vary significantly between cycle phases, anger was only assessed during the menstrual and midluteal phases. However, they found that levels of physical aggression were significantly higher during menses, with a similar trend for verbal aggression. Although these studies suggest that ‘normal’ women experience variable levels of aggression across the cycle, with the highest levels during
menses, these studies did not verify whether their participants met criteria for PMS. Therefore, the results from these studies may have been influenced by the inclusion of PMS sufferers, who self-report aggression to be among their most distressing symptoms (Endicott et al., 1999; Hylan et al., 1999; Landen and Eriksson, 2003; Lurie and Borenstein, 1990; Warner and Bancroft, 1990).

Therefore, most women seem to demonstrate some variability in aggression levels across the cycle, and PMS sufferers appear to demonstrate a greater fluctuation of symptoms. However, it is difficult to compare the findings from the studies, as various assessment methods were used. For example, Bond et al. (2003) measured aggression levels through the use of the CTS (Straus, 1979), which assesses the conflict resolution methods that are used to resolve disagreements in the family. In contrast, Dougherty et al. (1998) assessed aggression levels through the use of the PSAP, which operationalises aggression as the removal of points (money) from a fictitious opponent, in response to the loss of money (Cherek, 1981). This may be considered analogous to the real life scenario of taking someone’s car keys to prevent them from driving their car (Parrott and Giancola, 2007), and thus, clearly assesses a different form of aggressive behaviour than that assessed by the CTS. Moreover, the measures that were used in some of the studies call into question whether these studies actually measured aggressive behaviour at all. For example, some studies assessed anger (e.g. Collins et al., 1985; Van der Ploeg, 1987) and hostility (e.g. Haskett et al., 1984), which although are related to aggression (Buss and Perry, 1992; Spielberger et al., 1988), should be considered conceptually distinct (Parrott and Giancola, 2007). More studies are needed to assess aggression levels across the cycle of women with and without PMS using instruments that successfully tap aggressive behaviour, operationally defined as behaviours that are carried out with the intention of harming another person who is motivated to avoid that act (see section 1.4.2.2). Studies assessing anger and hostility could shed light on whether the inclusion of items depicting these related constructs on PMS scales could aid the differentiation of women with and without PMS.

1.4.2.3 Summary of impulsivity and aggression in PMS

Few studies have assessed impulsivity and aggression levels across the menstrual cycle of women with and without PMS. The studies that have been conducted have assessed these
symptoms through the use of disparate self-report measures and laboratory techniques, many of which were probably not successful in tapping these constructs. The two studies that assessed impulsivity levels in women with and without PMS using self-report measures could not establish whether these groups displayed different profiles across the cycle, as they only assessed levels once. Although three studies assessed impulsivity levels during a greater number of cycle phases through the use of objective tests, two of these did so through the use of the Go/No Go Avoidance Task (Howard et al., 1988; 1994), which is not a sensitive method by which to assess impulsivity levels in women (Howard et al., 1994). The results from the aggression studies suggest that most women demonstrate some variability in aggression levels across the cycle, and that PMS sufferers demonstrate a greater fluctuation in symptoms. However, few studies were conducted and some of these used the terms ‘anger’ and ‘hostility’ interchangeably with ‘aggression,’ which refer to related, but separate constructs (see section 1.4.2.2). Moreover, no study assessed aggression levels during menses, the cycle phase during which ‘normal’ women were found to report their highest levels. In order to determine whether women with and without PMS display different impulsivity and aggression profiles across the cycle, more studies are needed to assess these symptoms during a greater number of cycle phases, using sensitive and validated self-report measures. These studies should compare PMS sufferers, who have had their symptoms prospectively confirmed, with ‘control’ women who do not self-diagnose with PMS. Women with co-morbidities should be excluded from both groups, as PMS should not be confirmed in women suffering from psychiatric conditions. If these groups are shown to display different impulsivity and aggression profiles across the cycle, then the diagnosis of PMS and the assessment of treatment efficacy could be aided by the inclusion of items relating to these constructs on PMS measures.

1.5 Aetiology

Although the aetiology of PMS is not clearly understood (for a review see Halbreich, 2003), a multifactorial aetiology is likely (Keye and Trunnell, 1986; Lurie and Borenstein, 1990; Milewicz and Jedrzejuk, 2006). Various physiological mechanisms are postulated to be involved. Metabolic, hormonal and neurotransmitter actions have been considered to be causative factors in PMS (Lurie and Borenstein, 1990).
1.5.1 Metabolic
Metabolic processes hypothesised to explain the aetiology of PMS include disturbed prostaglandin (PG) metabolism (Horrobin, 1983), calcium dysregulation (Shamberger, 2003; Thys-Jacobs and Alvir, 1995), poor magnesium status (Abraham, 1982, Sherwood et al., 1986), abnormal fatty acid metabolism (Brush et al., 1984) and an increased sensitivity to (Horrobin, 1983) and elevated levels of (Carroll and Steiner, 1978; Halbreich, 1976) prolactin (PRL).

1.5.2 Hormonal
As PMS symptoms coincide with the hormonal fluctuations of the menstrual cycle, these have been proposed to result from an oestrogen excess and progesterone deficit (Bäckström and Carstensen, 1974; Mortola, 1998). However, oestrogen and progesterone levels do not differ between women with and without PMS (Connolly, 2001; Dennerstein et al., 1993; Rubinow, 1992), but ovarian hormones may play a role in the pathophysiology of PMS, as ovulation suppression ameliorates PMS symptoms (Eriksson et al., 2002; Wyatt et al., 2004; Yonkers et al., 2008).

1.5.3 Neurotransmitters
Women with PMS appear more sensitive to normal cyclical fluctuations of steroid hormones, which influence neurotransmitter function in the central nervous system (Eriksson et al., 2002; Rapkin, 1992; Reid and Yen, 1981; Steiner et al., 1997a). Increasing evidence suggests that serotonin plays an important role (Eriksson et al., 2002; Halbreich, 1997; Rapkin, 1992; Steiner, 1997). Many PMS symptoms have been linked to serotonergic dysfunction, including aggression (Bond et al., 2003; Cleare and Bond, 1995; Ho et al., 2001), depression (Hirschfeld, 2000; Steinberg et al., 1999; Veeninga and Westenberg, 1991; Zanoli, 2004), anxiety (Den Boer and Westenberg, 1988; Kahn et al., 1988; Veeninga and Westenberg, 1991) and irritability (Russo et al., 2005). Women with PMS have lower serotonin function premenstrually (Ashby et al., 1988; Rapkin et al., 1987; Rasgon et al., 2000; Taylor et al., 1984). Moreover, randomised controlled trials (RCTs) have shown that selective serotonin reuptake inhibitors (SSRIs), which facilitate serotonergic transmission, successfully reduce physical and psychological PMS symptoms (Dimmock et al., 2000; Eriksson et al., 2002; Steiner et al., 2003b; Szegedi et al., 2005; Wyatt et al., 2002; Yonkers et al., 2008). Fluoxetine (Hunter et al., 2002; Miner et al., 2002; Pearlstein et al., 1997;
Rapkin, 2003; Steiner et al., 1997b), citalopram (Wikander et al., 1998), paroxetine (Eriksson et al., 1995), sertraline (Freeman et al., 1999; Yonkers et al., 1997) and clomipramine (Sundblad et al., 1993) have all been shown to be effective PMS treatments. These beneficial effects occur rapidly (Dimmock et al., 2000; Landen and Eriksson, 2003; Wyatt et al., 2002; Yonkers et al., 2008), with efficacy demonstrated during the first luteal phase of treatment (Halbreich et al., 2002; Pearlstein et al., 1997; Steiner et al., 1995; Wikander et al., 1998). Moreover, intermittent SSRI treatment has been shown to effectively reduce PMS symptoms (Halbreich et al., 2002; Steiner et al., 1997b; Wikander, 1998; Wyatt et al., 2002). Therefore, SSRIs probably reduce PMS symptoms by acting upon different receptor sites than those involved in the beneficial SSRI effects for depression, which take approximately four to eight weeks to become apparent (Sundblad et al., 1993; Wikander et al., 1998; Wyatt et al., 2002).

The serotonergic system has been proposed to interact with other physiological mechanisms (Steiner et al., 2003b), including testosterone production (Eriksson et al., 1994; Steiner et al., 2003b; Virkkunen et al., 1994) and the cytokine network (Gemma et al., 1997; 2003; Hardin-Pouzet et al., 1996; Linthorst et al., 1994; Pousset et al., 1996; Silverman et al., 1989). Hence, fluctuations in the activity of these systems may result in symptom production. However, the exact mechanisms by which these fluctuations may lead to the production of symptoms is not known and beyond the scope of this thesis.

1.5.3.1 The role of testosterone in PMS

There is an inverse relationship between serotonin and testosterone production (Eriksson et al., 1994; Steiner et al., 2003b). Increased androgen levels have been associated with increased levels of anger and aggressive behaviour in women (Dabbs and Hargrove, 1997; Harris et al., 1996; Von der Pahlen et al., 2002), which are common and distressing PMS symptoms (Endicott et al., 1999; Hylan et al., 1999; Landen and Eriksson, 2003; Lurie and Borenstein, 1990; Warner and Bancroft, 1990). Moreover, androgens have been tentatively linked with mood and impulsive behaviour (Steiner et al., 2003b). These observations have led to the proposal that testosterone may be involved in the production of PMS symptoms (Rapkin, 2003), particularly irritability and impulsivity (Dougherty et al., 1997a; Sunblad et al., 1994). The administration of spironolactone, an androgen antagonist, reduced PMS symptoms in PMS sufferers who had high luteal testosterone levels (Burnet et al., 1991).
Although in some studies women with PMS show elevated testosterone levels (Eriksson et al., 1992; 1994), others do not (Bäckström and Aakvaag, 1981; Dougherty et al., 1997a; Watts et al., 1985). It is therefore difficult to conclude confidently that altered testosterone levels are present in women with PMS (Steiner et al., 2003b).

1.5.3.2 The role of cytokines in PMS

Recently the involvement of cytokines in PMS symptom formation has been reported. It is useful to discuss some background into cytokines in order to appreciate potential mechanisms by which these proteins could influence the expression of PMS symptoms.

1.5.3.2.1 What are cytokines?

Cytokines are a diverse group of proteins which act as signals between cells to regulate response to injury and infection. Cytokines have similar properties to the endocrine hormones. Some cytokines, such as IL-1, IL-6, IFN-γ and TNF-α, enhance immune response and are therefore proinflammatory, while others, such as IL-4, IL-10 and IL-13, are anti-inflammatory (O’Brien et al., 2004). Still other cytokines, such as IL-8, have both proinflammatory and anti-inflammatory roles (Kronfel and Remick, 2000). As cytokines have been associated with neurochemical and neuroendocrine activity in the brain, their potential involvement in psychiatric disorders has been highlighted (for a review, see Kronfol and Remick, 2000).

1.5.3.2.2 Cytokines and depression

Increased production of proinflammatory cytokines may play a role in the pathophysiology of major depression (MD) (Dantzer et al., 1999; Maes et al., 1995; Maes, 2001; Smith, 1991). MD has been associated with increased IL-6 production (Schlatter et al., 2001), and IL-6 levels have been found to correlate with symptom severity (Maes et al., 1995). Moreover, the administration of certain cytokines, including IFN-γ (Fent and Zbinden, 1987), IFN-α (Bonaccorso et al., 2002; MacDonald et al., 1987; Niiranen et al., 1988) and TNF-α (Kelley et al., 2003), has been shown to induce depressive symptoms, often at a severity sufficient for patients to meet DSM criteria for MD (Niiranen et al., 1988; Smith, 1991). It has even been proposed that depression is caused by an excessive secretion of particular cytokines, including IL-1, IFN-α and TNF-α (Smith, 1991). Smith (1991)
suggested that this could explain the higher prevalence of depression in the female population, as oestrogens promote cytokine production.

Cytokines and neurotransmitters (NTs) interact (Bonaccorso et al., 2002; Gemma et al., 2003; O'Brien et al., 2004). For example, IL-1 activates the serotonergic system (Gemma et al., 1997; Linthorst et al., 1994). Serotonin can stimulate IL-1 release (Silverman et al., 1989), and regulates IL-6 and TNF-α (Pouset et al., 1996). As MD has been associated with activation of cytokine networks and disturbances in serotonin metabolism, depression may be explained by an interaction of cytokines with the serotonergic system (Bonaccorso et al., 2002; Menkes and MacDonald, 2000). Animal models of depression (Van Gool et al., 2003) and human studies (Maes, 2001; Sluzewska et al., 1995) have shown that antidepressants, including SSRIs, normalize cytokine levels through their negative immunoregulatory effects, when they are administered for six to eight weeks (Maes, 2001; Sluzewska et al., 1995).

1.5.3.2.3 Potential involvement of cytokines in PMS
Cytokine release varies over the menstrual cycle (Gorai et al., 1997; Vrieze et al., 2003). Studies have shown that women in general have increased serum levels of IL-1 (Cannon and Dinarello, 1985; Polan et al., 1990) and IL-6 (Konecna et al., 2000) during the luteal phase. These cytokines have been associated with symptoms that are characteristic of PMS, including depression (Maes et al., 1995; Schlatter et al., 2001), anxiety (Connor et al., 1998), fatigue (Kronfel and Remick, 2000), headaches (Martelletti et al., 1993) and sleep disturbances (Kapas and Krueger, 1992; Vgontzas et al., 2005). Endotoxin administration, which causes the production and secretion of TNF-α, IL-1 and IL-6, has been shown to induce anxiety and depression (Reichenberg et al., 2001). Therefore, researchers have proposed that the up-regulation of cytokines may contribute to the pathophysiology of PMS (Konecna et al., 2000). Serotonin has been suggested to be involved in the aetiology of PMS (see section 1.5.3). Cytokines may interact with the serotonergic system to produce PMS symptoms, in a manner similar to that suggested for the aetiology of depression (Bonaccorso et al., 2002; Menkes and MacDonald, 2000). SSRIs, which normalize cytokine levels (Maes, 2001; Sluzewska et al., 1995; Van Gool et al., 2003), are an effective PMS treatment (Dimmock et al., 2000; Eriksson et al., 2002; Yonkers et al.,
To the author’s knowledge, no studies have yet examined cytokine levels in PMS sufferers.

1.6 Conclusion

While there is no universally accepted definition, PMS diagnosis has been operationalised as a 30% increase in symptoms in the luteal phase with a reduction in the follicular phase, severe enough to interfere with life. This operationalisation will be used in the studies presented in this thesis. As is evident from the previous discussion, the aetiology of PMS is as yet unclear but serotonergic involvement is generally accepted as a key aetiological factor. The role and interaction of other factors such as hormones and cytokines with 5-HT, symptom timing and severity merits further investigation. Core symptoms of PMS are anxiety and depression and it is important to distinguish cyclical reports from a manifestation of clinical or trait anxiety and depression which may be exacerbated premenstrually. These comorbidities require careful screening of potential volunteers using appropriate measures of clinical anxiety and depression. Symptoms of aggression and impulsivity cause PMS sufferers distress but have not been subject to much research. Further research is warranted to determine whether the inclusion of these symptoms on PMS scales could aid the diagnosis of PMS and the assessment of treatment efficacy.
Chapter 2: Aims

2.1 Treatment

Premenstrual symptoms affect almost all women of reproductive age. When symptoms are perceived to be sufficiently severe, the woman may be suffering from PMS or even PMDD (see section 1.3). Because there is no single clear and accepted aetiology, a large number of possible treatments are prescribed and purchased. The most commonly prescribed medications are SSRIs, based on the hypothesised role of serotonin (see section 1.5.3), which can be taken solely in the luteal phase. However, women with PMS are often reluctant to take SSRIs, partly due to the side effects produced (see section 6.1). Therefore, many women turn to dietary supplements and herbal remedies to treat their symptoms. Although various dietary supplements and herbal remedies have been proposed to alleviate PMS symptoms and are widely consumed, rigorous scientific studies to test their efficacy are lacking (see section 6.1). Therefore, this thesis aims to advance understanding regarding the effectiveness of the dietary supplements and herbal remedies that are commonly used by PMS sufferers. This aim is addressed by a) conducting a systematic research review of dietary supplements and herbal remedies commonly used for PMS (Chapter 6), and by b) performing a ten cycle double-blind RCT to determine whether St John’s Wort (SJW) is more beneficial than placebo supplements in relieving symptoms of PMS (Chapter 7). In order to throw light on the possible mechanism of action of SJW, serotonin (5-HT, 5-HIAA), hormone (oestrogen, progesterone, LH, FSH, PRL, testosterone) and pro-inflammatory cytokine concentrations (IL-1β, IL-6, IL-8, IFN-γ, and TNF-α) are examined in PMS sufferers on SJW and placebo treatment (Chapter 7).

To ensure the scientific quality of the RCT, it is necessary to firstly explore certain methodological considerations. These relate to the diagnosis and measurement of PMS. Although various PMS diagnostic criteria are currently used (Budeiri et al., 1994; Halbreich et al., 2007), the majority of researchers follow the NIMH (1983) guidelines. However, these unpublished criteria have been defined in two ways, and both have been widely adopted (30% increase criterion; modified 30% increase criterion) (see section 1.3.1.1). The 30% increase and modified 30% increase criteria result in different proportions of women being classified as PMS sufferers (Gallant et al., 1992a). As there is no consensus regarding the most appropriate PMS diagnostic criteria to use (Gallant et al., 1992a), researchers
should choose to use a criterion that identifies PMS sufferers who are experiencing PMS symptoms at a severity appropriate to address the aim of their study (Johnson, 2004; see also section 1.3.1.1). The primary objective of this research is to determine whether the widely available OTC herbal remedy, SJW, effectively relieves PMS symptoms. As it is usually PMS sufferers who are experiencing mild to moderate PMS symptoms who buy OTC preparations, and who are advised to try this form of treatment (see section 3.5.2), this research examines whether the 30% increase criterion or the modified 30% increase criterion best identifies PMS sufferers experiencing PMS symptoms at this level of severity (see section 3.5.2).

While there are a number of possible instruments with which to capture and track women’s symptoms during the menstrual cycle, these raise certain methodological, content and statistical concerns. For example, the well used American measure, the DSR (Freeman et al., 1996), does not include the items anger, aggression or impulsivity, which are distressing PMS symptoms that propel PMS sufferers to seek treatment (see section 1.4.2). Moreover, the factor structure of this measure was produced through a Principal Components Analysis (PCA) which produces components that are only applicable to the sample under analysis (Field, 2005). Hence this research explores the factor analysis of this measure with the aim of determining whether Freeman et al.’s factor structure can be generalised to the sample recruited for this research. Moreover, this research explores the factor structure during different cycle phases, in other diagnostic groups, and when the additional items anger, aggression and impulsiveness are added to the scale, to determine whether the DSR is a valid tool for assessing PMS symptoms in phases other than the luteal phase, in samples other than women meeting criteria for PMS, and when additional symptoms are added to the scale (Chapter 4).

2.2 Characterisation

2.2.1 Psychological profiles

Women with PMS often state that their most disturbing symptoms comprise aggression and impulsivity (see section 1.4.2). Moreover, women maintain that these symptoms premenstrually are the symptoms that motivate them to seek treatment (Bancroft et al., 1993; Hartlage and Arduino, 2002). However, research investigating the variability of
aggressive and impulsive symptoms across the menstrual cycle is lacking. Therefore, this research aims to further understanding of the PMS symptoms of aggression and impulsivity by a) examining whether women suffering from PMS demonstrate variable profiles of aggression and impulsivity across the menstrual cycle, and by b) determining whether cyclical profiles of aggression and impulsivity represent an abnormal symptom profile, or are variations of a normal cyclical pattern (Chapter 5).

2.2.2 Biological profiles

Although the aetiology of PMS is not clearly understood (Halbreich, 2003), it has been proposed to be multifactoral (Milewicz and Jedrzejuk, 2006). Increasing evidence suggests that serotonin is involved (Bryant and Dye, 2004; Halbreich, 1997; Steiner et al, 1997a). As the serotonin system interacts with testosterone (Eriksson et al., 1994; Steiner et al., 2003b; Virkkunen et al., 1994) and pro-inflammatory cytokine production (Gemma et al., 2003; 2007; Hardin-Pouzet et al., 1996; Linthorst et al., 1994; Pousset et al., 1996; Silverman et al., 1989), researchers have proposed that increased levels of testosterone (Dougherty et al., 1997a; Rapkin, 2003; Sunblad et al., 1994) and the up-regulation of pro-inflammatory cytokine levels (Konecna et al., 2000) may be involved in the pathophysiology of PMS. However, research which has prospectively examined the activity of these systems in PMS sufferers is scarce. Moreover, research which has investigated the relationship between these systems and the serotonergic system is lacking. Therefore, this research aims to determine whether serotonin (5-HT, 5-HIAA), hormone (oestrogen, progesterone, LH, FSH, PRL, testosterone) and pro-inflammatory cytokine concentrations (IL-1β, IL-6, IL-8, IFN-γ, and TNF-α) are associated with PMS symptoms (Chapter 5).

Taken together, the research presented in this thesis aims to further understanding of a possible treatment (SJW) for the Premenstrual Syndrome (PMS) and to improve the characterisation of the condition.
Chapter 3: General Methodology

This chapter will describe each of the measures used in the studies presented within this thesis and will discuss how they were developed or adapted from existing measures. It will then discuss the pilot study that was performed to assess the appropriateness of modifying general questionnaires for weekly administration in this way. The recruitment methods that were employed and the screening method that was used for all of the studies presented in this thesis will then be described. Finally, the selection process of the PMS diagnostic method that was considered appropriate to address the aims of this research will be discussed.

3.1 Measures

3.1.1 Daily Diary Booklets

The daily booklets were constructed as a tool to assess PMS symptoms across the cycle of women meeting PMS criteria (chapters 4, 5, 6 and 8) and across the cycle of women who did not consider themselves to be PMS sufferers (chapters 4, 5 and 6). The booklets were also used to assess the presence and severity of PMS, anxiety and depression in all women who participated in each of these studies.

Some PMS symptoms were measured daily, while others were measured on a weekly basis. The Daily Symptom Report (DSR) (Freeman et al., 1996), was used to provide a daily measure of core PMS symptoms. Some symptoms also contained on the DSR as single items were measured in greater depth through the administration of a weekly measure. As aggression, impulsivity, anxiety and depression cause women with PMS great distress and are often the symptoms that cause them to seek treatment (see sections 1.4.1 and 1.4.2), these symptoms were also assessed weekly through the administration of the Buss Perry Aggression Questionnaire (BPAQ) (Buss and Perry, 1992), the Barratt Impulsiveness Scale (BIS-11) (Patton et al., 1995), the State Trait Anxiety Inventory (STAI) (Spielberger, 1983b) and the Beck Depression Inventory (BDI) (Beck et al., 1961). Anxiety and depression were also assessed in order to identify whether women were suffering from clinical levels of anxiety and depression (see section 3.5.1). Each daily diary booklet contained one week of ratings i.e. seven copies of the DSR and one copy of the BDI, STAI, BPAQ and BIS-11. See Appendix 3.1 for a complete set of these measures.
3.1.1.1 The Daily Symptom Report (DSR) (Freeman et al., 1996)

The DSR consists of 17 items depicting common PMS symptoms, which load onto four factors (Freeman et al., 1996), describing mood, behavioural, pain and physical symptoms (see Table 4.1 for more detail on item loadings). The DSR requires women to rate their feelings and experiences daily, on five point Likert scales (0 = not present at all, to 4 = severe) for each item.

This measure was chosen because it is a brief, single page self-report questionnaire that assesses feelings and experiences prospectively and has been widely used (e.g. Atmaca et al., 2003; Bryant et al., 2005; Puder et al., 2006). Recordings from each day are provided on a separate page, reducing the likelihood of patterns in symptom reporting being spotted. Freeman et al. (1996) reported the scale to have an overall Cronbach’s alpha coefficient of 0.92, indicating high internal consistency of the 17 items. Construct validity was also demonstrated by the finding that all item scores were significantly higher in the PMS group than the non-clinical comparison group, which comprised women who reported that they did not have PMS (Freeman et al., 1996). Although Freeman et al. (1996) showed that there were moderate correlations between the DSR total and factor scores and those on the Premenstrual Assessment Form (PAF) (Halbreich et al., 1982b), a 95-item self-report scale designed to measure premenstrual mood, behaviour and physical changes (see section 1.3.1.3), little information is available about the relationships between PMS scales (Freeman, 2003).

Three symptoms were added to the original 17 DSR items; aggression, anger and impulsiveness. Aggression (Endicott et al., 1999; Hylan et al., 1999; Landen and Eriksson, 2003; Lurie and Borenstein, 1990; Warner and Bancroft, 1990) and impulsivity (Elliott, 2002; Halbreich, 2003; Halbreich et al., 2007; Hallman et al., 1987; Hsu et al., 2007; Matsumato et al., 2007) cause women with PMS great distress (Endicott et al., 1999; Hylan et al., 1999; Landen and Eriksson, 2003; Warner and Bancroft, 1990), and are often the symptoms that cause women to seek treatment (Hartlage and Arduino, 2002). Therefore, these items were added to the DSR to determine whether their inclusion could aid PMS diagnosis (chapter 6) and assessment of treatment efficacy (chapter 8).
3.1.1.2 The State-Trait Anxiety Inventory (Form Y) (Spielberger, 1983b)

Spielberger et al. (1983b) made the distinction between transitory anxiety states, which refer to unpleasant emotional states characterised by feelings of tension, apprehension, nervousness and worry and more stable anxiety traits, which refer to individual differences in anxiety proneness. A person’s trait anxiety refers to their tendency to perceive a situation as dangerous or threatening. This in turn influences the level of elevation in their state anxiety reaction to that situation. The STAI assesses state anxiety (STAIS) through 20 statements that evaluate how a person feels ‘right now, at this moment,’ on 4-point Likert scales (1=not at all, 4=very much so) for each item. Half the items refer to non-anxious experience (e.g. relaxed) and are reversed scored. Trait anxiety (STAIT) is assessed through 20 items that evaluate how a person ‘generally feels’ on 4-point Likert scales (1=almost never, 4=almost always) for each item. Nine items are reversed scored. High scores on the STAIS and STAIT depict high anxiety levels.

The STAI was chosen instead of other anxiety measures, such as the Hamilton Rating Scale of Anxiety, HAM-A (Hamilton, 1959) because it separates and measures state and trait anxiety. It is widely used (Sheehan-Dare et al., 1990) and has been used previously to assess changes in anxiety levels over the menstrual cycle of women who do (Chisholm et al., 1990; Haskett et al., 1984; Kuczmierczyk et al., 1992; Morse et al., 1984; Reed et al., 2008; Van der Ploeg, 1987) and do not (Golub et al., 1976; Lahmeyer et al., 1982; Veith et al., 1984) meet PMS criteria. The STAI is also a valid and reliable measure (Barnes et al., 2002; Spielberger, 1983b). For example reliability has been demonstrated through measures of internal consistency, where Spielberger (1983b) found uniformly high alpha coefficients for both the STAIS and STAIT scales in samples of working adults, students and military recruits, with median alpha coefficients of 0.93 (STAIS) and 0.90 (STAIT). The STAI also effectively differentiates between anxiety and depression (Cox et al., 1993; Endler et al., 1991; McWilliams et al., 2001; Spielberger, 1983b), which is important for the research presented in this thesis. Anxiety and depression are separate conditions (APA, 1994), although are often found co-morbidly (Maser and Cloninger, 1990). Both anxiety and depression levels have been shown to vary over the menstrual cycle (see section 1.4.1). It was desirable to assess the variability of anxiety and depression levels separately over the menstrual cycle of the women participating in the RCT presented in Chapter 7. The differentiation of anxiety and depression was also central for diagnostic purposes (see
section 3.5.1). Both clinical anxiety and depression can become magnified premenstrually. This can make women believe they have PMS (Freeman, 2003). However, women fail to meet PMS criteria if they suffer from either of these conditions continuously (APA, 1994). One aim of the research presented in this thesis was to determine how many women who believed they suffered from PMS failed to meet PMS criteria due to the presence of anxiety and depression, as separate entities and in combination.

3.1.1.3 The Beck Depression Inventory, BDI (Beck et al., 1961)

The BDI consists of 21 items designed to measure the presence and severity of affective, cognitive, motivational, vegetative and psychomotor manifestations of depression. Each item consists of a graded series of 4 to 5 self-evaluative statements, which are ranked to reflect the severity of the symptom. Respondents tick the statement that best describes how they feel. The severity of each item is scored by giving each statement a numerical value from 0-3. A total score is obtained by summating the values from each item. High scores on this measure signify high depression levels.

This measure was chosen because it is widely used (Richter et al., 1999; Schotte et al., 1997; Weeks and Heimberg, 2005), easy to complete, sensitive to change over time (Ormel et al., 1989; Richter et al., 1999), discriminates well between depressed and anxious patients (Beck et al., 1988; Cox et al., 1993; Endler et al., 1991; McWilliams et al., 2001; Weeks and Heimberg, 2005), has previously been used to assess changes in depression levels over the cycle of women meeting PMS criteria (Kuczmierczyk et al., 1992; Man et al., 1999; Morse et al., 1988) and is a valid and reliable measure (Beck et al., 1988; Richter et al., 1999; Schotte et al., 1997; Weeks and Heimberg, 2005). For example, Beck et al. (1988) conducted a meta-analysis of studies that had assessed the psychometric properties of the BDI from the years 1961-1986. They demonstrated the measure to have high internal consistency in both psychiatric (mean alpha coefficient of 0.86) and non-psychiatric (0.81) samples. These findings were confirmed by a meta-analysis performed by Richter et al. (1999). Beck et al. (1988) also found the BDI to correlate highly with clinical ratings of depression in psychiatric (mean correlation of 0.72) and non-psychiatric (0.60) samples. Further evidence of concurrent validity has been demonstrated by high correlations between the BDI and other self-rating scales for depression in both psychiatric and non-psychiatric samples (Beck et al., 1988; Richter et al., 1999).
3.1.1.4 The Barratt Impulsiveness Scale, BIS-11 (Patton et al., 1995)

The BIS-11 is a 30-item self-report inventory that measures three components of general impulsivity, including motor impulsivity (11 items), which refers to acting without prior thought, attentional impulsivity (8 items) which refers to making quick decisions, and nonplanning impulsivity (11 items) which refers to a lack of concern for the future (Barratt, 1994) by means of 4-point Likert scales (1=rarely/never, 4=almost always/always) for each item. 11 items are reversed scored to reduce response bias. High scores denote high impulsivity levels. Reliability of the BIS-11 has been demonstrated in various samples, including psychiatric inpatients, male inmates and undergraduates (Patton et al., 1995). The overall Cronbach’s alpha coefficient has been found to be high in undergraduate samples, with values of 0.79 and 0.82 having been reported (Archer and Webb, 2006; Patton et al., 1995). Barratt (1994) also reported the alpha coefficients for the individual subscales to be high in an undergraduate sample, ranging from 0.50-0.72. Miller et al. (2004) replicated these findings in a more diverse sample of 245 adults (mean age 42.7 years, sd 15.4), reporting alpha coefficients ranging from 0.61-0.72. Construct validity of the BIS-11 has been demonstrated by Miller et al. (2004), who found moderate positive correlations between the BIS-11 subscales and other self-report measures of impulsivity, including Dickman’s (1990) Impulsivity Inventory and Eysenck et al.’s (1985b) Impulsiveness Questionnaire.

3.1.1.5 The Aggression Questionnaire, BPAQ (Buss and Perry, 1992)

This 29-item measure assesses self-reported aggressive feelings and behaviours by means of five point Likert scales (1=extremely uncharacteristic of me, 5= extremely characteristic of me) for each item. Two items are reversed scored. High BPAQ scores indicate low aggression levels. Through the use of factor analysis, Buss and Perry (1992) demonstrated that this scale comprises four subscales, categorized as physical aggression (9 items), verbal aggression (5 items), anger (7 items) and hostility (8 items). Buss and Perry (1992) reported internal consistency coefficients for the subscales ranging from 0.72 to 0.85 in college students. High alpha coefficients have also been reported for female undergraduates (Harris, 1997) and for slightly older women (mean age 28.3 years sd 5.8) who were experiencing regular menstrual cycles and who were not taking hormonal contraception (Von der Pahlen et al., 2008). Researchers have also demonstrated the BPAQ to have good test-retest reliability. Buss and Perry (1992) found nine-week test–retest
coefficients ranging from 0.72-0.80 in a college sample, while Harris (1997) found moderately high (r=0.47 to 0.88) seven-week test-retest coefficients in female undergraduates. Construct validity of the BPAQ has been demonstrated by many researchers, who have found the BPAQ to be related to other self-report measures of trait aggression, including Gladue's (1991) Aggression Inventory (Archer et al., 1995a), to measures of act-based aggressive behaviour (Archer and Webb, 2006) and to behavioural indicators of aggressive behaviour, including fighting behaviour (Archer et al., 1995b) and salivary testosterone in male and female university students (Harris et al., 1996). The BPAQ has previously been used to assess changes in aggression levels over the menstrual cycle (Ritter, 2003).

3.1.1.6 Counterbalancing and order effects

Data from a previous study in the Human Appetite Research Unit (HARU), University of Leeds, suggested that the order of items on the DSR might influence reporting behaviour, such that many women completing the DSR gave the same score for similar items that followed each other e.g. nervous tension and anxiety, swelling, cramps and aches etc. Therefore, three alternative orders of the DSR were created by randomising the items. The analysis of these "order" effects is described in Chapter 4. The seven DSR checklists in each daily diary booklet were presented in the same item order.

The four weekly measures were counterbalanced using a Latin square design, which produced four orders of the weekly questionnaires. The STAI was counted as one measure, with the STAIS always being administered before the STAIT, as recommended by Spielberger (1983b). Therefore 16 orders of the daily diary booklet were created. For all studies within this thesis, participants were given packs of daily diary booklets, with diary orders being administered in a semi random order, such that each woman received approximately the same number of each diary order. To prevent women from looking at previously completed diaries and spotting patterns in symptom reporting, they were asked to return each diary as soon after its completion as possible.
3.2 Pilot study of the daily diary booklets

3.2.1 Introduction

The purposes of the pilot study were to examine whether the original questionnaires had concurrent validity when administered on a weekly basis, and to confirm that the average daily measures of the symptoms anxiety, depression, impulsiveness and aggression correlated with their retrospective measures at the end of the week.

The daily diary booklets described above were designed to provide a daily measure of premenstrual symptoms, and weekly measures of depression, anxiety, impulsivity and aggression levels, using validated instruments. To administer these measures on a weekly basis, it was necessary to modify the instructions of some of the original questionnaires. The BDI (Beck et al., 1961) was not modified since it directs respondents to answer the questions in relation to the last week. The instructions for the STAIT ask respondents to think about how they ‘generally feel.’ This was not modified, as Spielberger (1983b) maintains that the measure should always be administered in the original format. The STAIS instructions were altered from asking respondents to answer the questions in relation to how they feel ‘right now, at this moment,’ to how they ‘felt over the last week.’ Spielberger (1983b) stated that the STAIS instructions can be modified in this manner to measure state anxiety for any time interval. The instructions on the BIS-11 (Patton et al., 1995) and the BPAQ (Buss and Perry, 1992) were also reworded to refer to the previous week, although the authors do not comment on whether these questionnaires should or should not be modified in this way.

3.2.2 Procedure

Ethical approval for the pilot study was obtained from the Research Ethics Committee, Institute of Psychological Sciences, University of Leeds. 101 women were recruited from around the campus of the University of Leeds via posters and email distributions which offered course credits. When a woman contacted the researcher expressing an interest in the study, they were sent an information sheet outlining the details of the study and a briefing meeting was arranged. At this meeting, the researcher answered any questions the participants had and informed consent was obtained. All participants were provided with one daily diary booklet with the original instructions of the weekly questionnaires and one booklet with the modified instructions. Participants were instructed to begin the first diary
as soon as possible, by completing one copy of the DSR (see section 3.1.1.1) each evening and the weekly questionnaires at the end of the week. They were then asked to leave a 3 day gap before beginning their second diary. The date was recorded on each copy of the DSR. The two daily booklets were administered in counterbalanced order, with 50 women being given the booklet with the original instructions followed by the modified booklet, and 51 women being given the modified booklet followed by the booklet containing the original instructions. The daily diary booklets were clearly labelled to indicate the order in which they were to be completed.

3.2.3 Results

The following results are based on the 38 women who completed the study. Table 3.1 displays the means (sds) for each measure using both the original and modified instructions, calculated from the participants’ ratings in the daily diary booklets. Pearson’s Moment Correlation Coefficients between the original and modified measures are also included.

Table 3.1. Means (sds) and Pearson’s moment correlation coefficients for the original and modified questionnaire subscales and total scores (n=36)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Original Mean (sd)</th>
<th>Modified Mean (sd)</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAIS*</td>
<td>38.03 (10.83)</td>
<td>44.39 (12.53)</td>
<td>0.18</td>
</tr>
<tr>
<td>STAIT</td>
<td>41.95 (10.08)</td>
<td>41.78 (10.13)</td>
<td>0.79**</td>
</tr>
<tr>
<td>BDI</td>
<td>7.79 (5.34)</td>
<td>7.59 (6.08)</td>
<td>0.58**</td>
</tr>
<tr>
<td>BPAQ-verbal</td>
<td>17.26 (4.00)</td>
<td>18.11 (4.22)</td>
<td>0.67**</td>
</tr>
<tr>
<td>BPAQ-physical*</td>
<td>36.61 (7.81)</td>
<td>38.74 (5.75)</td>
<td>0.65**</td>
</tr>
<tr>
<td>BPAQ-anger</td>
<td>25.61 (5.25)</td>
<td>25.05 (5.32)</td>
<td>0.77**</td>
</tr>
<tr>
<td>BPAQ-hostility</td>
<td>30.11 (5.84)</td>
<td>30.95 (6.29)</td>
<td>0.69**</td>
</tr>
<tr>
<td>BPAQ-total</td>
<td>109.58 (17.51)</td>
<td>112.89 (16.84)</td>
<td>0.78**</td>
</tr>
<tr>
<td>BIS-attentional</td>
<td>15.62 (3.23)</td>
<td>16.23 (3.50)</td>
<td>0.83**</td>
</tr>
<tr>
<td>BIS-motor</td>
<td>22.00 (4.59)</td>
<td>21.51 (4.22)</td>
<td>0.59**</td>
</tr>
<tr>
<td>BIS-nonplanning*</td>
<td>26.08 (4.98)</td>
<td>27.89 (5.10)</td>
<td>0.77**</td>
</tr>
<tr>
<td>BIS-total</td>
<td>63.73 (11.02)</td>
<td>65.69 (11.15)</td>
<td>0.83**</td>
</tr>
</tbody>
</table>

* p<0.05, ** p<0.001 NB. The asterisks next to the variable names represent significant differences between the means of the original and modified questionnaires.
The total questionnaire scores appeared very similar between the original and modified versions, with the exception of the STAIS where the modified version (44.39) gave a significantly higher mean score than the original (38.09) \((t(35)=-2.58, p<0.05)\). The mean scores on the subscales of the BPAQ and BIS-11 also appeared very similar between the original and modified versions, with the exception of the BPAQ physical subscale \((t(37)=-2.20, p<0.05)\) and the BIS-11 nonplanning subscale \((t(34)=-2.85, p<0.05)\), where the modified version produced significantly higher mean scores.

Table 3.1 also shows the Pearson's Moment Correlation Coefficients between the original and modified weekly questionnaires. As the standard deviations were found to be large, the correlations were also explored using Spearman's ranked data, as this coefficient relies on ranked data (Field, 2005), and its calculation does not take into account standard deviation. These findings were consistent with the Pearson's Moment Correlation Coefficients shown in Table 3.1.

The total scores for the BDI and the STAIT from one week to the next were significantly positively correlated, which was expected, as there was no change in the wording of their instructions. When comparing the test scores on the originally worded BIS-11 and BPAQ with the test scores on their modified versions, significant, highly positive correlations were produced for their total scores and subscales. As the original BPAQ and BIS-11 have been shown to be both reliable and valid (e.g. Buss and Perry, 1992; Patton et al., 1995), this would suggest that these modified questionnaires have concurrent validity and that they are valid measures of assessing the constructs of impulsivity and aggression over the course of a week. The correlation between the original and modified STAIS was found to be non-significant \((r=0.18, p>0.05)\). These results could be explained by the fact that state anxiety is a transitory state, dependent upon the situations that a person is subjected to (Spielberger, 1983b). Therefore, the low correlation could be explained by the participants being subject to different situations between the two weeks, and not due to the differences in instructions.

To fulfill the second aim of the pilot study, it was then necessary to determine whether the daily ratings of the items 'anxiety' 'depression' 'aggression' and 'impulsivity,' which were either original DSR items, or items that had been added to the scale (see section 3.1.1.1), correlated with the measures designed to gauge these constructs at the end of the week.
This was performed for the original and modified booklets by producing Pearson’s Moment Correlation Coefficients between the average DSR ratings of ‘anxiety,’ ‘depression,’ ‘impulsivity’ and ‘aggression’ and the STAI, BDI, BIS-11 and BPAQ scores. These correlation coefficients are displayed in Table 3.2.

Table 3.2 Pearson’s Moment Correlation Coefficients between the daily and weekly measures (n=36)

<table>
<thead>
<tr>
<th></th>
<th>Anxiety</th>
<th>Depression</th>
<th>Aggression</th>
<th>Impulsiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAILY</td>
<td>O</td>
<td>M</td>
<td>O</td>
<td>M</td>
</tr>
<tr>
<td>STAIS</td>
<td>0.28</td>
<td>0.66**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STAIT</td>
<td>0.09</td>
<td>0.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDI</td>
<td>0.48**</td>
<td>0.56**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPAQ-verbal</td>
<td>0.04</td>
<td>-0.40*</td>
<td>-0.52**</td>
<td>-0.54**</td>
</tr>
<tr>
<td>BPAQ-physical</td>
<td>-0.15</td>
<td>-0.46**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPAQ-anger</td>
<td>-0.24</td>
<td>-0.50**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPAQ-hostility</td>
<td>-0.36*</td>
<td>-0.58**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPAQ-total</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIS-attentional</td>
<td>0.46**</td>
<td>0.56**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIS-motor</td>
<td>0.32</td>
<td>0.52**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIS-nonplanning</td>
<td>0.23</td>
<td>0.43*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIS-total</td>
<td>0.35*</td>
<td>0.56**</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key: O = original M = modified * p<0.05 ** p<0.001

N.B. As discussed in 3.1.1.5, high BPAQ scores indicate low aggression levels

As table 3.2 shows, a highly significant, moderate correlation was found between the average DSR rating of ‘depression’ and the BDI score, in both the original (r = 0.48, p<0.001) and modified (r = 0.56, p<0.01) booklets. When looking at the correlations between the average DSR rating of ‘aggression’ and the BPAQ subscale and total scores, greater, and more significant correlations were found when the instructions had been modified to the last week. This would suggest that the modified instructions provide a more accurate picture of a woman’s aggression levels over the course of a week. The same pattern was found for the BIS-11.

When assessing the average DSR ‘anxiety’ ratings in relation to the STAIS, the original instructions were not well correlated with their average daily anxiety score (r = 0.28, p>0.05). However, the modified instructions gave a highly significant correlation (r=0.66, p<0.001). The instructions for the STAIT remained in their original form in both booklets.
The average daily rating of ‘anxiety’ did not correlate well with the STAIT in either booklet ($r = 0.09$, $p > 0.05$ and $r = 0.28$, $p > 0.05$). Trait anxiety refers to relatively stable individual differences in the way individuals perceive stressful situations and influences their state anxiety as a reaction to that situation. Therefore, one might expect state anxiety to correlate with a person’s rating of their ‘anxiety’ levels that week, as this is a measure of a person’s anxiety state. However, for trait anxiety to be correlated with women’s ‘anxiety’ ratings, the previous week would have to be representative of their normal situation.

### 3.2.4 Conclusion

The pilot study aimed to test the validity of the modified BIS-11, BPAQ and STAIS. The results suggest that the modified BIS-11 and BPAQ can be considered to be valid measures to assess impulsivity and aggression levels over the course of a week, as highly significant correlations were produced between the original and modified versions. The original measures are validated measures that have been shown to be valid and reliable (Archer and Webb, 2006; Archer and Haigh, 1997; Buss and Perry, 1992; Harris, 1997; Miller et al., 2004; Patton et al., 1995; Von der Pahlen et al., 2008). The correlations between the BPAQ and BIS-11 and women’s average daily ratings of ‘impulsivity’ and ‘aggression’ suggest that the modified measures reflect levels of ‘impulsivity’ and ‘aggression’ over the previous week more closely than the original versions. Although the original and modified STAIS measures were not found to be correlated, the modified STAIS was found to be significantly correlated with women’s average daily ratings of anxiety, suggesting that the modified STAIS is a valid tool for assessing women’s anxiety levels over the course of a week.

### 3.3 Menstrual cycle phases

For the research presented in this thesis, it was essential that symptoms were rated during each menstrual cycle phase, and that other measures were obtained during specific phases. The average cycle length is approximately 28-29 days (Chiazze et al., 1968; McIntosh et al., 1980). However, cycle length varies between and within women (McIntosh et al., 1980; Sherwood, 2004). There remains disagreement as to how the menstrual cycle should best be divided (Atmaca et al., 2003; Collins et al., 1993; Steiner and Wilkins, 1996; Thys-Jacobs et al., 1998; see also section 1.3.1.2).
For the research presented in this thesis, where a cycle was taken to mean the first day of one bleed to the first day of the next (WHO, 1966), cycle days 1-4 were averaged and identified the ‘bleeding phase’, days 5-10 the follicular phase and days -1 to -6 the luteal phase (NIMH, 1983). As these criteria were applied to cycles of different lengths, the number of days between the end of the follicular phase and the beginning of the luteal phase varied between and within women. This phase therefore consisted of the average of the remaining cycle days and was referred to as ‘rest (of cycle).’ These criteria were those agreed upon at the NIMH PMS workshop (Blume, 1983) and have been widely adopted (e.g. Atmaca et al., 2003; Bryant et al., 2005; Kendall and Schnurr, 1987). They are also the criteria suggested by Freeman et al. (1996) for use with her instrument.

Women began a daily diary booklet on the first day of a menstrual cycle. They began a new booklet each week, until they began their next period. On this day they began a new daily diary booklet and completed the weekly questionnaires from the previous booklet. Therefore, the weekly questionnaires (STAI, BDI, BPAQ, BIS-11) were always completed on cycle days 7, 14 and day 1 of the next menstrual cycle. As the weekly questionnaires asked women to report how they had been feeling over the previous week, those that were completed on day 7 were taken to represent the bleeding phase, those on day 14 the follicular phase and those on cycle day 1 of the next cycle the luteal phase. This left the time period between cycle days 15 to -1 of the next cycle. As cycles differ in length, the number of diaries completed during this phase varied.

To best represent the ‘rest’ phase, if there were two completed diaries, the one immediately before the ‘luteal’ diary was excluded, as this usually overlapped with days in the luteal diary, due to women being asked to move straight to the weekly questionnaire on the first day of a period. In the rare case where three diaries were completed (i.e. in a cycle greater than 35 days) the middle diary was used, as this was most equally distanced from the follicular and luteal phases.

For example, for a 37 day cycle, the weekly questionnaires completed on cycle day 7 (covering days 1-7) would represent the bleeding phase, day 14 (days 8-14) the follicular phase, day 1 of the next cycle (days -6 to -1) the luteal phase and day 28 (days 22-28) the ‘rest’ phase, as shown in Figure 3.1.
This method meant that the weekly questionnaires representing the bleeding and follicular phases did not completely correspond to those phases outlined above. However, it was thought that asking women to complete weekly questionnaires on cycle days 4 and 10 and to then complete them weekly until the end of the cycle, when they should complete another, would have caused greater confusion and poor adherence.

3.4 The screening cycles

All data collected for the analyses conducted in this thesis arose from the recruitment of two groups of women. Firstly, women who thought that they suffered from PMS were recruited to take part in the clinical trial detailed in Chapter 8. The first part of this study involved women completing three screening cycles, where eligibility to enter the intervention stage was assessed. As it was of interest to compare the symptom profiles, biological profiles and cognitive performance of these women with those of women not meeting PMS criteria, a second group of women who believed that they experienced ‘normal’ menstrual cycles were recruited to take part in the screening cycles (menstrual cycle study). This section will describe the recruitment methods that were used to recruit these two groups of women, the numbers of women that were recruited and completed the screening cycles, and the screening method that was employed.

3.4.1 Recruitment Methods

3.4.1.1 Clinical Trial

Volunteers were recruited for the clinical trial through various strategies, all of which made it clear that the study required women who thought that they may suffer from PMS and
directed attention to the study website (www.psyc.leeds.ac.uk/q/pms), where the online recruitment questionnaire could be found. Posters were sent to all doctors’ surgeries, leisure centres, community centres and libraries in the Leeds area, as well as posted throughout the campus of the University of Leeds. Recruitment advertisement emails were sent to every departmental secretary in the University to distribute to staff and students and to the university email distribution lists. Letters were posted to 35 past participants of a similar study at the HARU, who had indicated their willingness to participate in future studies. Short articles were published in the newsletters of the National Association of Premenstrual Syndrome (NAPS) and the University of Leeds (The Reporter), and also appeared online on their websites. Study flyers were placed in the conference packs given out at the annual NAPS conference and were distributed at an alternative health day in Leeds. A longer article describing the Premenstrual Syndrome, its diagnosis and treatment, before detailing this study was published in the Yorkshire Post. The Principal Investigator (PI) (Sarah Canning) and lead PhD supervisor (Dr Louise Dye) appeared on BBC Radio Leeds, while the PI worked with the producers of BBC Look North to produce a short programme about PMS before advertising the study.

3.4.1.2 Menstrual cycle study

Volunteers were recruited for the menstrual cycle study through various methods, all of which made clear that the focus of the study was to look at the subtle changes in feelings and behaviours that may occur across the menstrual cycles of women who experience ‘normal’ cycles, and directed attention to the study website (www.psyc.leeds.ac.uk/yourcycle). Posters were sent to 50 community centres in the Leeds area and were placed across the campus of the University of Leeds. Recruitment advertisement emails were sent to every departmental secretary in the University of Leeds and Leeds Metropolitan University to distribute to staff and students and to distribution lists of the University of Leeds interest networks. A study advertisement was placed on the website of the staff newsletter (The Reporter). These adverts were placed at a different time to the PMS recruitment adverts.

3.4.2 Selection criteria

Table 3.3 displays the inclusion and exclusion criteria that were applied for the clinical trial and menstrual cycle study.
Table 3.3 Inclusion and exclusion criteria for participation

<table>
<thead>
<tr>
<th>INCLUSION CRITERIA</th>
<th>EXCLUSION CRITERIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applicable to both studies:</td>
<td>Applicable to both studies:</td>
</tr>
<tr>
<td>- From the Leeds area, or able to visit regularly</td>
<td>- Pregnant or breast-feeding</td>
</tr>
<tr>
<td>- Aged 18-45 years</td>
<td>- Using hormonal preparations,</td>
</tr>
<tr>
<td>- Regular menstrual cycles (25-35 days)</td>
<td>including oral contraception, the mirena coil or HRT</td>
</tr>
<tr>
<td></td>
<td>- Currently diagnosed with, or feels</td>
</tr>
<tr>
<td></td>
<td>that may be suffering from depression and/or anxiety</td>
</tr>
<tr>
<td>Applicable to clinical trial only:</td>
<td>Applicable to clinical trial only:</td>
</tr>
<tr>
<td>- Of good health, assessed by the clinic doctor</td>
<td>- Taking prescribed or over the counter medication for PMS</td>
</tr>
<tr>
<td></td>
<td>- Photosensitive</td>
</tr>
<tr>
<td></td>
<td>- Taking any of the preparations listed in Appendix 3.6</td>
</tr>
<tr>
<td>Applicable to menstrual cycle study only:</td>
<td>Applicable to menstrual cycle study only:</td>
</tr>
<tr>
<td>- n/a</td>
<td>- Having been diagnosed with, or feels</td>
</tr>
<tr>
<td></td>
<td>that may be suffering from PMS</td>
</tr>
</tbody>
</table>

3.4.3 Participation Progress

In total, 156 women emailed the PI to show an interest in the menstrual cycle study and were sent the study handbook (Appendix 3.7). Of these, 133 replied. 63 did not meet the selection criteria (Table 3.3). Therefore, 70 women began this study and attended their first study visit. 53 completed all three screening cycles.

371 women completed the online recruitment questionnaire for the clinical trial. Of these, 270 were not suitable to begin screening for the clinical trial, or lived too far away. These women were offered the opportunity to take part in postal screening and 84 women were sent daily diary booklets. Of these, 39 completed the full three screening cycles. 101 women attended the briefing meeting about the clinical trial, of which 91 began screening. 50 clinical trial participants completed the three screening cycles. A record of the progress of all participants is shown in Figure 3.2
Figure 3.2 Participant recruitment and progress at each stage of the screening cycles

3.4.4 Measures

3.4.4.1 Recruitment questionnaire: All participants completed a recruitment questionnaire at the beginning of the study to gather information such as age, psychiatric history, use of hormonal preparations and menstrual cycle length. The recruitment questionnaire acted as an initial screening tool to assess whether participants met the inclusion/exclusion criteria.
3.4.4.2 PMS symptoms: The daily diary booklets (see section 3.1.1), containing seven copies of the DSR (Freeman et al., 1996) and one copy of the BDI (Beck et al., 1961), STAI (Spielberger, 1983b), BPAQ (Buss and Perry, 1992) and BIS-11 (Patton et al., 1995) were given to participants throughout the screening cycles to monitor PMS symptoms and levels of depression, anxiety, aggression and impulsivity across the cycle. They were also the primary tool for the diagnosis of PMS (see section 3.3). As clinical trial participants were asked to complete these diaries for an additional seven menstrual cycles, the trait measure of anxiety, STAIT, was only administered to these participants during the first screening cycle to reduce study burden. Spielberger (1983b) recommended that the state measure of anxiety, STAIS, should be given each time a measure is needed to assess change in anxiety over time.

3.4.4.3 Blood samples: Two blood samples (each 18mls) were taken during the screening cycles to assess follicular and luteal hormone (FSH, LH, oestradiol, progesterone, prolactin and testosterone), serotonin (5-HT and 5-HIAA) and cytokine (IL-1β, IL-6, IL-8, IFN-γ and TNF-α) levels. Blood samples were collected in EDTA tubes and were immediately centrifuged for 10 minutes at 4200rpm. Plasma was removed and aliquoted into 8 Eppendorfs, which were stored at -80° centigrade in the Institute of Psychological Sciences, until they were transferred to St James’s University Hospital, where the samples were thawed and analysed.

3.4.4.4 Emotional Stroop Task: In order to investigate the salience to stimuli denoting PMS symptoms in women with PMS, in comparison to women not meeting PMS criteria, participants were asked to perform an emotional Stroop task. The results that were produced from the administration of this cognitive task will be reported at a later date.

3.4.5 Procedure

Ethical approval for the screening cycles was gained from the Institute of Psychological Sciences, University of Leeds, for the participants who volunteered for the clinical trial and for those who volunteered for the menstrual cycle study. A number of recruitment methods were undertaken to make women aware that the study was taking place (see sections 3.6.1.1 and 3.6.1.2), all of which directed attention to the relevant study website and gave the contact details of the PI.
If a woman contacted the PI to show an interest in the menstrual cycle study, they were sent the study handbook (Appendix 3.7), which described the study in detail and listed the inclusion/exclusion criteria. Women were asked to read the study handbook and to contact the PI to arrange their first study visit if they were still interested in taking part, met the inclusion criteria and did not meet the exclusion criteria.

If a woman contacted the PI regarding the clinical trial, they were asked to complete the recruitment questionnaire online, or were sent a copy by post if they did not have internet access. Women who met the selection criteria (Table 3.3) were contacted to arrange their first study visit and sent the information sheet detailing the clinical trial (Appendix 3.8). If the recruitment questionnaire indicated that women did not live in the Leeds area, did not meet the inclusion criteria or met the exclusion criteria, they were contacted, informed that they could not take part in the clinical trial and were asked if they would like to take part in the screening cycles by post. If they did, they were sent an information sheet outlining what would be involved (Appendix 3.9) and a consent form (Appendix 3.10). On return of the consent form, participants were sent their first diary pack and asked to begin diary completion on the first day of their next period and to continue doing so for three full menstrual cycles. Once the participants had completed their symptom recording, the PI analysed their data and sent them a detailed symptom profile. These participants were not required to attend any study visits.

Women participating in the menstrual cycle study and the screening cycles of the clinical trial were asked to attend the study visits detailed below:

**Study visit 1 (Briefing meeting)**: During this visit, women participating in the menstrual cycle study were asked if they had read the study handbook and had any questions about the study. If women were happy to continue, they were asked to complete the recruitment questionnaire and were taken through the inclusion/exclusion criteria (Table 3.3) by the PI. Informed consent was then taken (Appendix 3.11).

Clinical trial participants were given the opportunity to ask any questions that had arisen from reading the information sheet (Appendix 3.8). A brief medical history was then taken from these participants, and their blood pressure recorded. If the study doctor (Dr Julie
Ayres) approved, the participant was recruited, the admission form was completed (Appendix 3.12) and informed consent was taken (Appendix 3.13).

All participants were familiarised with the emotional Stroop task and given their first set of diaries to take away. Women were asked to begin completing these on the first day of their next period and to continue doing so for three full menstrual cycles, returning a diary at the end of each week in a Freepost envelope provided. Throughout the study, two diaries were sent to participants biweekly with Freepost envelopes for their return, with a letter reminding them of what they should be doing at that time and dates of forthcoming visits.

When participants had returned diaries from one full menstrual cycle, they were contacted by the PI to arrange their second study visit during the follicular phase (cycle days 5-10) of their third menstrual cycle. This was predicted from estimating the start date of their third period using information they had provided during the briefing meeting and counting forwards. Participants were asked to contact the PI on the first day of their third cycle, to check that the follicular visit had been booked in during the correct menstrual cycle phase. If not, this appointment was re-arranged.

**Study visit 2:** During the second study visit, participants performed the emotional Stroop task and had a blood sample taken. Their third visit was arranged during their luteal phase (cycle days -6 to -1) of the same menstrual cycle, which was estimated by the PI, who counted backwards from the predicted start date of their fourth period.

**Study visit 3:** Participants performed the emotional Stroop task and had a blood sample taken. Clinical trial participants were informed whether they were eligible to continue to the intervention phase (see section 8.2.11.1). Women participating in the menstrual cycle study were weighed and measured. All women were asked to continue completing their diaries until the first day of their next period. As clinical trial participants would then continue to take part in a demanding protocol, they were not asked to attend any more visits during the screening cycles. However, in order to minimise missing data from the luteal phase in the menstrual cycle study, an additional appointment was booked five days after the third study visit. If women began their period before this meeting was due, they were asked to contact
the PI to cancel this appointment, as this meant that visit 3 had taken place during the luteal phase of the third screening cycle (cycle days -6 to -1).

**Study visit 4 (menstrual cycle study only):** If participants had not begun their period five days after study visit 3, they returned to the HARU to perform the emotional Stroop task and to have a blood sample taken. They were asked to contact the PI on the first day of their next period. If the PI had not heard from the participant within five days of study visit 4, the participant was contacted and asked to return to the HARU to perform the emotional Stroop task and to have another blood sample taken. After participants from the menstrual cycle study had returned their diaries from the three full menstrual cycles, they were sent a detailed symptom profile in the post and were given an honorarium (£60) to compensate them for the time and effort that they had invested in the study.

### 3.5 Participant groups

#### 3.5.1 Co-morbid criteria

Women who volunteered for the clinical trial (clinical trial participants and women undergoing postal screening) were labelled *PMS reporters*, as they responded to recruitment methods that advertised for women who believed that they suffered from PMS. Women who volunteered for the menstrual cycle study were labelled *PMS non-reporters*, as they responded to recruitment methods that advertised for women who experienced ‘normal’ menstrual cycles. Moreover, women who had been diagnosed with PMS, or felt that they may suffer from this condition, were excluded from taking part in the menstrual cycle study (see Table 3.3).

The PI analysed the diaries completed by the participants who took part in the clinical trial, postal screening and menstrual cycle study, to determine whether PMS criteria were met. As the diagnosis of PMS requires the absence of clinical anxiety and depression (APA, 1994), it was firstly necessary to exclude women who met criteria for these conditions.

To determine whether a woman was suffering from clinically relevant depression, their follicular phase BDI scores were examined, since these are not influenced by premenstrual symptoms and should reflect the woman’s baseline depression levels (Kuczmierczyk et al., 1992). The clinical threshold for ‘caseness for depression’ is a score of 12 on the BDI, with
scores of 12-20 indicating mild depression, 20-30 moderate depression and over 30 severe depression (Quah-Smith et al., 2005; Nielson and Williams, 1980). Therefore, if a woman had an average follicular phase BDI score of 12 or above, she was considered to be suffering from depression.

In order to determine whether a woman was suffering from anxiety, her follicular phase STAIT scores were examined. Although the STAIT is widely used, there is no accepted cut-off to indicate clinical anxiety (Sheehan-Dare et al., 1990). However, normative data has been produced. Spielberger et al. (1983b) reported the norm for working females between 19-39 years to be 36.15 (sd 9.53). Fisher and Durham (1999) reviewed six RCTs of psychological therapy for generalized anxiety disorder (GAD). From combining all available pre-treatment STAIT scores (N=404) a mean score of 57.00 (sd 9.45) was produced.

Researchers have adopted different techniques based on normative data to determine a cut-off to represent significant anxiety levels. Some researchers have determined cut-off values based on whether a person’s score lies closer to the mean of the functional population than it does to the dysfunctional population (Fisher and Durham, 1999; Jacobson et al., 1984; Jacobson and Truax, 1991). If these criteria were adopted for the studies presented in this thesis, a cut-off of 46.28 would have to be applied (Appendix 3.12). Other researchers have adopted a cut-off based on the upper 90% probability limit (1.645 sd above the mean) for the general adult population (Sheehan-Dare et al., 1990). If these criteria were applied to the norms derived by Spielberger et al., scores greater than 51 would represent significant anxiety (Appendix 3.12). These criteria were chosen to determine whether a woman was suffering from significant levels of anxiety in the research presented in this thesis for pragmatic reasons. Therefore, if a woman had an average follicular phase STAIT score of 51 or above, she was considered to be suffering from anxiety.

Table 3.4 displays the characteristics of the women who were identified as suffering from anxiety and/or depression (co-morbid), the vast majority of whom self-diagnosed with PMS, and met criteria for both anxiety and depression.
Table 3.4 Co-morbid characteristics

<table>
<thead>
<tr>
<th></th>
<th>Anxiety only</th>
<th>Depression only</th>
<th>Anxiety &amp; depression</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMS reporter</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>25</td>
</tr>
<tr>
<td>PMS non-reporter</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>

3.5.2 PMS criteria

It was then necessary to determine whether women met PMS criteria. As discussed in Section 1.3.1.1, various PMS diagnostic criteria are currently used, and these have been shown to result in different proportions of women being classified as PMS sufferers (Gallant et al., 1992a). As there is no consensus regarding the most appropriate PMS diagnostic criteria to use (Gallant et al., 1992a), researchers should choose to use a criterion that identifies PMS sufferers who are experiencing PMS symptoms at a severity appropriate to address the aim of their study (Johnson, 2004; see also section 1.3.1.1). The primary objective of this research was to determine whether St John’s Wort (hypericum perforatum), a widely available OTC herbal remedy, was more beneficial than placebo treatment in relieving PMS symptoms (Chapter 2). Women who are experiencing mild to moderate PMS symptoms are often reluctant to take prescribed medication (Bendich, 2000), or even to visit their doctor (Domoney et al., 2003), and frequently buy OTC preparations for their symptoms (Bendich, 2000; Domoney et al., 2003; Eriksson et al., 2002; Freeman et al., 2002; Fugh-Berman and Kronenberg, 2003; Girman et al., 2003). Even if women with mild to moderate PMS symptoms do decide to visit their doctor, they are most often advised to firstly make lifestyle changes and to try OTC preparations to relieve their symptoms (Bryant and Dye, 2004; Connolly, 2001; Johnson, 2004; see also section 1.3.1.1). If these treatment strategies prove ineffective, or if symptoms are more severe, then prescribed medications, such as SSRIs, are administered (Dimmock et al., 2000; Johnson, 2004; Wyatt et al., 2002; see also section 1.3.1.1). Therefore, as it is usually PMS sufferers who are experiencing mild to moderate PMS symptoms who buy OTC preparations, and who are advised to try this form of treatment, this research aimed to identify PMS sufferers who experienced PMS symptoms at this level of severity, using appropriate diagnostic criteria.
Although various severity of change criteria are widely employed as a means to diagnose PMS (e.g. Anderson et al., 1988; Atmaca et al., 2003; De Ronchi et al., 2005; Gallant et al., 1992a; Gehlert and Hartlage, 1997; Graham and Sherwin, 1993; Schnurr, 1988; Stevinson and Ernst, 2000; Thys-Jacobs et al., 1998; see also section 1.3.1.1), the majority of researchers follow the NIMH (1983) guidelines. However, these unpublished criteria have been defined in two ways, and both have been widely adopted (see section 1.3.1.1). As described in Section 1.3.1.1, some researchers follow the 30% increase criterion, by confirming PMS in women who demonstrate a follicular to luteal symptom increase of at least 30% in at least two menstrual cycles (e.g. Facchinetti et al., 1991; Hicks et al., 2004; Pearlstein et al., 1997; Rapkin, 1987). Others follow the modified 30% increase criterion, by confirming PMS in women who demonstrate at least a 30% follicular to luteal symptom increase in at least two menstrual cycles, in relation to the range of the scale being used to make the diagnosis (e.g. Evans et al., 1998; 1999; Schmidt et al., 1998; Smith et al., 2002; Tung-Ping et al., 1997). Given these two approaches, the NIMH (1983) criteria could be operationalised either through the application of the 30% increase criterion or through the application of the modified 30% increase criterion. It was considered appropriate to examine which criterion best identified PMS sufferers who were experiencing PMS symptoms at a severity appropriate to the aims of this research (i.e. at a mild to moderate severity). This was ascertained by firstly examining the proportion of women who self-diagnosed with PMS who met each criterion. Then, the symptom profiles of the groups that were formed according to each criterion were examined. Finally, the premenstrual symptom severity of the PMS sufferers that were identified through each of these diagnostic methods was compared with the premenstrual symptom severity of groups of PMS sufferers previously classified as experiencing PMS at different severities.

The 30% increase criterion and the modified 30% increase criterion both require a woman to demonstrate at least a 30% follicular to luteal symptom increase for at least two out of three menstrual cycles for a PMS diagnosis to be confirmed (see section 1.3.1.1). Therefore, a diagnosis could not be made for women who completed booklets for one menstrual cycle only, or for those who completed diaries for two menstrual cycles, but showed a premenstrual increase in symptoms of less than 30% in one of those cycles. 153 women completed diaries for long enough for a PMS diagnosis to be made. Sections
3.5.2.1 and 3.5.2.2 describe the diagnostic groups that were formed when each of these
diagnostic methods were employed.

3.5.2.1 30% increase criterion

The NIMH (1983) did not specify the number or type of symptoms that should be affected
in order for a PMS diagnosis to be confirmed (Mortola, 1996; Smith et al., 2002). Researchers have therefore applied these criteria to PMS symptoms in different ways (see
section 1.3.1.1). However, most researchers take the conservative approach of applying the
NIMH criteria to the total scale score of the measure that they are using to make their PMS
diagnosis (e.g. Atmaca et al., 2003; Bryant et al., 2005; De Souza et al., 2000; Facchinetti et
al., 1991; Walker et al., 1998). Therefore, a woman was diagnosed with PMS using the
30% increase criterion in this research if she demonstrated at least a 30% increase in her
DSR total scale score between her follicular and luteal phase, during at least two out of her
three screening cycles. The formula below was used to calculate the follicular to luteal
symptom increase for each woman. The luteal score was obtained by averaging the DSR
total scale scores from cycle days -6 to -1. The follicular score was obtained by averaging
the DSR total scale scores from cycle days 5 to 10.

\[
\text{Luteal score} - \frac{\text{Follicular score}}{\text{Luteal score}} \times 100
\]

Table 3.5 displays the diagnostic groups that were formed through the application of the
30% increase criterion.
Table 3.5 Diagnostic groups formed from the clinical trial, postal screening and menstrual cycle study when the 30% increase criterion was employed

<table>
<thead>
<tr>
<th>PMS reporter</th>
<th>Symptom Increase</th>
<th>Co-morbidity</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMS sufferers</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>Unconfirmed PMS reporters</td>
<td>✓</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PMS non-reporters</td>
<td>X</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>Controls</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Co-morbids
- Symptom increase
- No symptom increase

The application of the 30% increase criterion resulted in the majority of PMS reporters meeting PMS criteria (n=68). These women were labelled *PMS sufferers*. Only five women who self-diagnosed with PMS were found not to meet the 30% increase criterion. These women were labelled *unconfirmed PMS reporters*. Approximately one third of women who did not report PMS did not meet the 30% increase criterion, or criteria for anxiety or depression. These women were labelled *controls*. However, the majority of PMS non-reporters met the 30% increase criterion (n=33). These women were labelled *PMS non-reporters*. Figure 3.3 displays the DSR symptom profiles that were reported by the PMS sufferer, PMS non-reporter, control and co-morbid groups across the average of the three screening cycles (i.e. mean of screening cycles 1 to 3). The symptom profile of the unconfirmed PMS reporters is not displayed in this figure, as this group only comprised five women.
The total DSR symptom profiles displayed in Figure 3.3 demonstrate that the PMS sufferers showed greater variability in DSR symptom reporting across the menstrual cycle than all other groups, and demonstrated a larger follicular to luteal increase in symptoms. Although the co-morbids generally reported symptoms at a greater severity across the cycle than all other groups, the PMS sufferers reported symptoms at a similar severity to this group premenstrually. The controls appeared to report symptoms at a slightly greater severity than the PMS non-reporters during the follicular phase. However, the control and PMS non-reporter groups appeared to report symptoms at a similar severity to each other during all other cycle phases.

Figures 3.4 to 3.7 display the mood, behavioural, pain and physical DSR symptom profiles of the four diagnostic groups across the averaged screening cycles. As these four DSR subscales are made up from different numbers of items, each subscale was divided by the number of items it comprised so that meaningful comparisons could be made between the subscales.
The mood symptom profiles that were reported across the cycle by each group (Figure 3.4) appear similar to their behavioural symptom profiles (Figure 3.5). As was apparent for the total DSR symptom profiles, the PMS sufferers showed the greatest variation in mood and behavioural symptom reporting across the cycle, and demonstrated a larger follicular to luteal symptom increase than all other groups. Again, the co-morbid group appeared to
report mood and behavioural symptoms at a greater severity than all other groups across the cycle, although the PMS sufferers appeared to report mood and behavioural symptoms at a similar severity to the co-morbid group premenstrually. Although the controls appeared to report mood and behavioural symptoms at a slightly greater severity than the PMS non-reporters during the bleed, follicular and rest menstrual cycle phases, the control and PMS non-reporter groups seemed to report mood and behavioural symptoms at an almost identical severity premenstrually. The four groups appeared to report similar, variable patterns of pain symptoms across the cycle (Figure 3.6). Although the same pattern was apparent for physical symptoms, slightly greater group separation was apparent premenstrually (Figure 3.7). Whilst the PMS non-reporter and control groups both reported physical symptoms at a lower severity than the PMS sufferer and co-morbid groups during the luteal phase, slightly greater separation was apparent between the control and PMS non-reporter groups for physical symptoms during the luteal phase than was found to be the case for the other three DSR subscales.

3.5.2.2 Modified 30% increase criterion
To be diagnosed with PMS through the application of the modified 30% increase criterion, a woman had to show an average symptom increase of at least 1.5 points (i.e. 30% increase on a 5-point scale) from the follicular (days 5-10) to the luteal phase (days -6 to -1), in at least two out of the three screening cycles. This firstly involved the calculation of an average follicular and luteal score. The follicular score was obtained by averaging the DSR total scale scores from cycle days 5 to 10. The luteal score was obtained by averaging the DSR total scale scores from cycle days -6 to -1. The follicular and luteal scores were then divided by the number of items included on the DSR (17 items). To meet the 30% increase criterion, a woman had to show an increase of at least 1.5 points between these scores.

Table 3.6 displays the diagnostic groups that were formed through the application of the modified 30% increase criterion.
Table 3.6 Diagnostic groups formed from the clinical trial, postal screening and menstrual cycle study when the modified 30% increase criterion was employed

<table>
<thead>
<tr>
<th></th>
<th>PMS reporter</th>
<th>Symptom Increase</th>
<th>Co-morbidity</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMS sufferers</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
<td>7</td>
</tr>
<tr>
<td>Unconfirmed PMS reporters</td>
<td>✓</td>
<td>X</td>
<td>X</td>
<td>66</td>
</tr>
<tr>
<td>Controls</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>50</td>
</tr>
<tr>
<td>Co-morbid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Symptom increase</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>2</td>
</tr>
<tr>
<td>- No symptom increase</td>
<td></td>
<td>X</td>
<td>✓</td>
<td>28</td>
</tr>
</tbody>
</table>

When the modified 30% increase criterion was adopted, only seven women who self-diagnosed with PMS had this diagnosis confirmed (PMS sufferers). This resulted in 66 women who self-diagnosed with PMS being classified as not suffering from the syndrome (unconfirmed PMS reporters). The 50 women who reported not to suffer from PMS were found not to meet PMS criteria (controls). This meant that the use of the modified 30% increase criterion did not result in any women who did not self-diagnose with PMS meeting PMS criteria (PMS non-reporters).

Figure 3.8 displays the DSR symptom profiles that were reported by the PMS sufferer, unconfirmed PMS reporter, control and co-morbid groups across the averaged screening cycles (i.e. mean of screening cycles 1 to 3).
The total DSR symptom profiles displayed in Figure 3.8 suggest that the PMS sufferers showed much greater variability in DSR symptom reporting across the cycle than all other groups, and demonstrated a much larger follicular to luteal increase in symptoms. Moreover, this group of women seemed to demonstrate premenstrual symptoms at a much greater severity than the women who were identified as PMS sufferers through the application of the 30% increase criterion (see Figure 3.3). The unconfirmed PMS reporters and controls appeared to report symptoms at an almost identical severity during the bleed, follicular and rest menstrual cycle phases. However, the unconfirmed PMS reporters appeared to demonstrate a greater increase in symptoms premenstrually. The co-morbid reporters reported symptoms at a greater severity than the unconfirmed PMS reporter and control groups across the menstrual cycle. Although they also reported symptoms at a greater severity than the PMS sufferers during the follicular and rest menstrual cycle phases, they reported symptoms at a lower severity than the PMS sufferers during the bleed and luteal phases.
Figures 3.9 to 3.12 display the mood, behavioural, pain and physical DSR symptom profiles of the four diagnostic groups across the averaged screening cycles.

The PMS sufferers appeared to demonstrate much greater variability in mood, behavioural, pain and physical symptom reporting across the cycle than all other groups, and showed a larger follicular to luteal increase in these symptoms (Figures 3.9 to 3.12). Again, this
group of seven women appeared to report all types of PMS symptom at a much greater severity premenstrually than the women who were identified as PMS sufferers through the use of the 30% increase criterion (see Figures 3.4 to 3.7). As was apparent for the total DSR symptom profiles, the unconfirmed PMS reporter and control groups appeared to report mood, behavioural, pain and physical symptoms at a similar severity to each other during the bleed, follicular and rest menstrual cycle phases. However, the unconfirmed PMS reporters appeared to report all types of PMS symptom at a greater severity than the control group premenstrually, with this pattern being most evident for mood and behavioural symptoms (Figures 3.9 and 3.10). The co-morbid group generally reported symptoms at a greater severity than the unconfirmed PMS reporter and control groups across the cycle for all types of PMS symptom. However, the unconfirmed PMS reporters appeared to report mood and physical PMS symptoms at a similar level to the co-morbid group premenstrually (Figures 3.9 and 3.12).

3.5.2.3 Benefits and limitations of the 30% increase and modified 30% increase criteria
Appropriate PMS criteria should differentiate women who claim to experience problematic premenstrual symptoms from women who do not report experiencing premenstrual complaints (Gallant et al., 1992a). As has previously been found, neither the use of the 30% increase criterion nor the use of the modified 30% increase criterion successfully managed to separate these groups of women (Gallant et al., 1992a).

The application of the 30% increase criterion seemed to successfully identify the vast majority of women who self-reported problematic PMS symptoms, as the use of this criterion resulted in the confirmation of a PMS diagnosis in 68 of the 73 women who self-diagnosed with PMS (see Table 3.5). However, the use of this criterion did not seem to successfully differentiate this group of women from women who did not self-report premenstrual complaints, as 33 of the 50 women who did not self-diagnose with PMS were also found to meet the 30% increase criterion (PMS non-reporters). In fact, only 17 women who did not self-diagnose with PMS did not meet this criterion (controls). Previous research has also found that a large proportion of women who do not self-report distressing premenstrual symptoms meet a PMS diagnosis when the 30% increase criterion is used (Gallant et al., 1992a; Morse et al., 1988). This group of women (PMS non-reporters) appeared to report premenstrual symptoms at an almost identical severity to the control
group (see Figures 3.4 to 3.7). This, together with the fact that the PMS non-reporters did not consider themselves to experience premenstrual complaints (Gallant et al., 1992a), suggests that women belonging to this group should not meet a diagnosis for PMS, but should instead be considered to experience ‘normal cyclicity.’ The main difference in the symptom profiles that were reported by the PMS non-reporter and control groups was that the PMS non-reporters appeared to report symptoms at a slightly lower severity than the control group during the follicular phase. As the 30% increase criterion bases its diagnosis on a woman’s follicular to luteal symptom increase, the use of this criterion seemed to sub-divide ‘control’ women into two groups by producing a positive PMS diagnosis in women who showed a low symptom severity during the follicular phase (PMS non-reporters) and by producing a negative PMS diagnosis in women who reported a slightly higher symptom severity during this phase. As such, a diagnosis of PMS was confirmed in a group of women who should not have met PMS criteria (false positives).

The modified 30% increase criterion did not result in the production of any ‘false positives,’ as no women who did not self-diagnose with PMS were found to meet PMS criteria (see Table 3.6). However, this criterion did not successfully differentiate women who did and who did not consider themselves to suffer from problematic PMS symptoms, as only seven of the 73 women who self-diagnosed with PMS had this diagnosis confirmed. This resulted in 66 women who self-diagnosed with PMS being considered not to suffer from the syndrome (unconfirmed PMS reporters). Therefore, the application of the modified 30% increase criterion to the total scale score of the DSR in this research seemed to only identify women at the severe end of the PMS spectrum (see section 3.5.2.2). This is not surprising given that women were required to demonstrate an average symptom increase of at least 1.5 DSR points (see section 3.5.2.2). Therefore, for women to meet this criterion, they were required to either experience all 17 symptoms included on the DSR at a mild to moderate severity, or to experience fewer symptoms at a greater severity. PMS sufferers do not usually experience all PMS symptoms (Steiner and Wilkins, 1996). The minimum number of DSR items that a woman could experience and meet the modified 30% increase criterion was six. These symptoms would have to be rated using the highest anchor on the scale in the luteal phase, and the lowest anchor on the scale in the follicular phase. The highest anchor requires symptoms to be ‘overwhelming and/or prevent a woman from carrying out her usual activities.’ Therefore, the use of this criterion did not result in a
PMS diagnosis being confirmed in women who experienced less than six DSR symptoms, women who experienced approximately six DSR symptoms but who did not rate each of these at the highest possible severity during the luteal phase with complete amelioration in the follicular phase, or in women who experienced a greater number of symptoms at a mild to moderate severity. If researchers do not only wish to select women who suffer from severe PMS symptoms, then the application of the modified 30% increase criterion to a total scale score could produce ‘false negatives,’ such that women who experience PMS symptoms at a mild to moderate severity are considered not to suffer from the condition (Figures 3.9 to 3.12).

3.5.2.4 Choosing an appropriate criterion

This section aimed to establish whether the application of the 30% increase criterion or the application of the modified 30% increase criterion would identify PMS sufferers who were experiencing premenstrual symptoms at an appropriate severity level to examine the efficacy of the OTC preparation SJW for PMS i.e. at a mild to moderate severity (see section 3.5.2). The PMS sufferers who were identified through the use of the modified 30% increase criterion (see Figure 3.3) were found to experience more severe premenstrual symptoms than the PMS sufferers who were identified through the use of the 30% increase criterion (Figure 3.8), and demonstrated a much larger follicular to luteal increase in symptoms. This is not surprising given that the 30% increase criterion only required women to demonstrate a 30% or greater increase in their total DSR score between the follicular and luteal phases to meet a PMS diagnosis, whilst the modified 30% increase criterion required women to demonstrate an average symptom increase of at least 1.5 DSR points between these phases (see section 3.5.2.3). Therefore a much greater proportion of women who self-diagnosed with PMS met a PMS diagnosis when the 30% increase criterion was applied to the total DSR score (see Table 3.5) than when the modified 30% increase criterion was utilised (see Table 3.6). It is clear that the use of the 30% increase criterion identified PMS sufferers who were experiencing milder PMS symptoms than those who were identified through the use of the modified 30% increase criterion. However, in order to determine whether the 30% increase criterion or the modified 30% increase criterion best identified PMS sufferers who were experiencing premenstrual symptoms at a mild to moderate severity, as was required for the RCT, it was necessary to compare the premenstrual symptom severity of the PMS sufferers who were identified
through each of these diagnostic methods with the premenstrual symptom severity of groups of PMS sufferers identified in previous research.

Researchers who have conducted PMS treatment trials have applied various diagnostic criteria to the DSR in order to select PMS sufferers experiencing PMS symptoms at a particular severity (Atmaca et al., 2003; Bryant et al., 2005; Freeman et al., 1995; Freeman et al., 2002; Freeman et al., 2005). The researchers who developed the DSR (Freeman et al., 1996) applied different diagnostic criteria to this measure when they required PMS sufferers with mild (e.g. Bryant et al., 2005; Freeman et al., 2002), severe (e.g. Freeman et al., 1995), and extremely severe (e.g. Freeman et al., 2005) PMS symptoms, depending on the treatment that they were evaluating. The women who were identified as PMS sufferers through the use of the 30% increase criterion in this research were found to report premenstrual symptoms at an almost identical severity (mean premenstrual DSR score=14.70, sd=8.36) to the symptom severity previously reported by women suffering from mild PMS in a dietary treatment study (mean=13.40, sd=9.70) (Bryant et al., 2005). Ratings were at a much lower severity to the symptom severity previously reported by women experiencing severe PMS in a treatment trial of oral progesterone and alprazolam (mean=23.00, sd=12.5) (Freeman et al., 1995) and by women with PMDD in a treatment trial of the SSRI escitalopram (mean=30.17, sd=9.67) (Freeman et al., 2005). In contrast, the seven women who were identified as PMS sufferers through the application of the modified 30% increase criterion (mean=30.45, sd=3.76) reported premenstrual symptoms at a greater severity than the premenstrual symptom severity reported by the PMDD sufferers in Freeman et al.’s (2005) RCT of escitalopram (SSRI).

The 30% increase criterion was chosen as the PMS diagnostic method to use in this research for several reasons. The primary aim of the research presented in this thesis was to determine whether the OTC preparation, SJW, was more beneficial than placebo supplementation for the treatment of PMS. The 30% increase criterion identified the majority of women who self-diagnosed with PMS (see Table 3.5), and as such identified women representative of the women likely to buy this form of treatment. As the modified 30% increase criterion only identified a minority of these women (see Table 3.6), the use of this criterion would fail to identify the majority of the women likely to try an OTC preparation to relieve symptoms they attribute to the premenstrual phase. Moreover, the
30% increase criterion identified women who were experiencing PMS symptoms at a mild severity, a severity representative of the women who buy OTC preparations to treat their symptoms (Bendich, 2000; Domoney et al., 2003; Eriksson et al., 2002; Freeman et al., 2002; Fugh-Berman and Kronenberg, 2003; Girman et al., 2003), and who are advised to try this form of treatment by their clinicians (Connolly, 2001; Johnson, 2004; see also section 1.3.1.1). The modified 30% increase criterion only identified PMS sufferers who reported extremely severe PMS symptoms, a group of women who would be less likely to buy OTC preparations to treat their symptoms, and more likely to receive prescribed medication (Dimmock et al., 2000; Eriksson et al., 2002; Johnson, 2004; Wyatt et al., 2002). Furthermore, many researchers who have previously conducted double-blind, placebo-controlled RCTs to assess the effectiveness of various OTC preparations for PMS symptoms have also used the 30% increase criterion to select their participants (e.g. Bryant et al., 2005; De Souza et al., 2000; Facchinetti et al., 1991; Freeman et al., 2002; Hicks et al., 2004; Pearlstein et al., 1997; Sayegh et al., 1995; Walker et al., 1998). The researchers who developed the DSR (Freeman et al., 1996) used the 30% increase criterion to diagnose PMS in their treatment trial of the carbohydrate-rich beverage Escape. In contrast, the modified 30% increase criterion is usually used when researchers wish to select PMS sufferers who are experiencing severe symptoms, for example when the effectiveness of treatments such as alprazolam (e.g. Evans et al., 1998) and SSRIs (e.g. Tung-Ping et al., 1997) are being assessed.

3.5.3 Group formation for the analyses presented in this thesis

The analyses presented in this thesis were conducted on the groups of women that were identified through the application of the 30% increase criterion; PMS sufferers; PMS non-reporters; controls; co-morbids (see Table 3.5). Although some women self-diagnosed with PMS but did not meet PMS criteria, or criteria for anxiety and depression (unconfirmed PMS reporters), this group was extremely small (n=5), and was therefore excluded from all subsequent analyses.

Co-morbids who also demonstrated a symptom increase were analysed separately from those who did not in the Principal Components Analyses that are reported in Chapter 4, in order to determine whether these groups reported PMS symptoms in the same way in relation to one another. However, these groups were combined for the analyses presented in
Chapter 5, where the main aim was to determine whether the impulsivity and aggression profiles reported by PMS sufferers differed to those reported by women suffering from anxiety and/or depression.

As discussed in Section 3.5.2.3, the use of the 30% increase criterion seemed to sub-divide normally cycling women into two (controls and PMS non-reporters). Therefore, the PMS non-reporter and control groups were combined into one group of normally cycling women for the analyses presented in Chapter 5, in order to provide a comparison of the impulsivity and aggression profiles that were reported by PMS sufferers with those reported by women experiencing 'normal' cyclicity. However, some researchers attempt to disguise the purpose of their research and allocate women to diagnostic groups after they have completed symptom reports in order to reduce the impact of expectation effects or increased introspection (e.g. Christensen and Oei, 1988; Dougherty et al., 1998). This can actually result in women who do not report distressing PMS symptoms, and who should be included in control groups in PMS studies (PMS non-reporters), actually being included in PMS groups (Gallant et al., 1992; Morse et al., 1988; see also section 3.5.2.3). Therefore, the analyses presented in Chapter 5 were also conducted with the PMS non-reporters and controls as separate groups, in order to assess the impact of retrospective group allocation.

The implications that the use of the 30% increase criterion may have had, versus the use of other diagnostic criteria, such as the modified 30% increase criterion, on the analyses that were performed in this thesis, will be returned to in the final chapter.
4.1 Introduction

Freeman et al.'s DSR (1996) was used in all of the studies presented in this thesis to diagnose and measure PMS symptoms (see sections 3.1.1.1 and 3.5.2). In the development of this measure, Freeman et al. (1996) used Principal Components Analysis (PCA) with Varimax rotation to reduce the items on the DSR into overarching components. This chapter will describe available data reduction techniques, the components derived by Freeman et al., and the issues that should be considered when performing a PCA. The PCAs performed in this research and the results of these are also considered in detail.

4.2 Data Reduction techniques and their key characteristics

When data are collected via a questionnaire, a large amount of information is obtained. It is important to reduce the data set to a manageable size and to understand the structure between the items being measured (Fabrigar et al., 1999; Field, 2005; Hayton et al., 2004; Zwick and Velicer, 1986). Such summarising of the observed variables aims to derive a smaller set of variables (Zwick and Velicer, 1986), whilst retaining as much information from the original variables as possible (Fabrigar et al., 1999). PCA and factor analysis (FA) are commonly used data reduction techniques that condense questionnaire data in different ways (Tabachnick and Fidell, 2007a).

FA estimates which factors are involved by producing a mathematical model based on the correlations among the measured variables (Fabrigar et al., 1999). PCA analyses variance, not correlations, and extracts the smallest number of components to explain the maximum variance (Tabachnick and Fidell, 2007a). PCA establishes which linear components exist within the data set and determines how particular variables contribute to them (Field, 2005). PCA is the method of choice when the main aim is to condense a data set. It empirically summarizes a specific data set, assuming that the sample being analysed is the population of interest (Field, 2005). Therefore, the findings produced relate to a specific sample. Generalizability to the whole population can be achieved by confirming the obtained factor structure in other samples (Field, 2005). Analyses performed using FA can be more easily generalised to the population, as this method assumes that the participants being analysed have been randomly selected from the population of interest. While these
theoretical differences exist between PCA and FA, they often produce very similar results (Arrindell and van der Ende, 1985; Gorsuch, 1997; Velicer and Jackson, 1990). Although PCA technically produces components rather than factors, these terms are used interchangeably by researchers in the field and while these differences in meaning are acknowledged in this thesis, both terms (factors and components) will be used, in line with the research being discussed i.e. in accordance with how other researchers have used these terms.

4.2.1 Decision-making in PCA
When performing a PCA, the researcher should make important decisions about the way in which factors are extracted and rotated based on their hypotheses and knowledge of the research field (Fabrigar et al., 1999; Gorsuch, 1997; Zwick and Velicer, 1986). However, many researchers simply use the default settings on the available statistical package, and these may not be the most appropriate techniques for their research aim (Fabrigar et al., 1999; Gorsuch, 1997).

4.2.1.1 Factor Extraction
There are various methods available to determine how many components should be retained, and these have been shown to produce very different results (Hayton et al., 2004; Zwick and Velicer, 1986). Under- and over-extraction of components can lead to considerable errors in subsequent results (Hayton et al., 2004; Zwick and Velicer, 1986). If too few components are retained, important information can be lost through ignoring a factor or combining it with another (Zwick and Velicer, 1986). Specifying too many factors can result in solutions that are difficult to interpret and replicate (Zwick and Velicer, 1986), focus attention on minor components at the expense of major components, and create components which only comprise of one high loading item (Comrey, 1978). Therefore, determining how many components should be retained is one of the most critical decisions that must be made when conducting a PCA (Fabrigar et al., 1999; Gorsuch, 1997; Hayton et al., 2004; Zwick and Velicer, 1986).

The "eigenvalue greater than one" rule, K1, (Kaiser, 1960) is a measure of the total variance explained by each extracted component based on the use of standardised variables in a PCA. The rationale for retaining components with eigenvalues greater than one is that
they represent a sizeable amount of variation (Field, 2005) and provide more summarizing power than the original variables (Zwick and Velicer, 1986).

This method has been suggested to be arbitrary, differentiating between factors with eigenvalues of 0.99 and 1.01 (Fabrigar et al., 1999), and has been shown to lead to severe overestimation (Fabrigar et al., 1999; McWilliams et al., 2001; Zwick and Velicer, 1986). Also, this rule is often used to determine the exact number of factors which should be retained, when really it demonstrates the maximum number that should be retained (Gorsuch, 1997). Even though this method has been heavily criticized, with some advocating that it should no longer be used (Zwick and Velicer, 1986), it is still widely employed and is the default method of factor extraction on many statistical packages, including SPSS and SAS (Gorsuch, 1997; Hayton et al., 2004).

The scree test (Cattell, 1966) determines how many factors should be retained by plotting descending eigenvalues. Its' rationale is that a few major factors account for most of the variance, and so only these should be retained. To find these, the researcher must examine the scree plot and determine the point at which the difference in descending eigenvalues becomes small. This is known as the 'elbow.' Factors lying above this point are retained. Although some have found this method to be relatively accurate (Zwick and Velicer, 1986), it has been criticized for being subjective and ambiguous (Fabrigar et al., 1999; Hayton et al., 2004), and is particularly difficult to conduct when eigenvalues form a gradual slope, producing no clear 'elbow' (Zwick and Velicer, 1986).

Parallel Analysis (PA) (Horn, 1965) is based on the rationale that components from a real data set that has a valid underlying factor structure should have larger eigenvalues than components derived from a random set of uncorrelated variables, based on the same sample size and number of variables. To conduct a PA, 50 random correlation matrices of random variables based on the sample size and number of variables in the data set must firstly be created. The average of these eigenvalues and/or the 95th percentile of eigenvalues must then be compared to the real data. The factors that have eigenvalues greater than the eigenvalues from the random data are retained (Hayton et al., 2004). If the average/ 95th percentile eigenvalues produce a different result, many argue that the 95th percentile
structure should be used, as this reduces PA’s slight tendency to overestimate (Glorfeld, 1995; Harshman and Reddon, 1983; Zwick and Velicer, 1986).

PA (Horn, 1965) has repeatedly been found to be the most consistently and frequently accurate method of factor retention (Hayton et al., 2004; Humphreys and Montanelli, 1975; Zwick and Velicer, 1986). For example, Zwick and Velicer (1986) generated several population correlation matrices with systematically varied sample sizes, variable numbers and component numbers and tested the ability of PA, the scree test and the KI rule to estimate the number of components present, given only the generated sample matrices. They found PA to be the most frequently accurate method that they examined. The KI rule consistently overestimated component number. Although the scree test more accurately estimated component number than the KI rule, it had a slight tendency towards overestimation.

Some researchers suggest that multiple criteria should be used for factor retention (Fabrigar et al., 1999; Zwick and Velicer, 1986). Many agree that PA should be performed (Hayton et al., 2004; Humphreys and Montanelli, 1975; Zwick and Velicer, 1986) and some suggest that this should be used in conjunction with the scree test (Fabrigar et al., 1999; Zwick and Velicer, 1986), with the solutions from each being examined for their theoretical plausibility and interpretability.

4.2.1.2 Factor Rotation

Although generally items load highly onto one component, they often load onto the other components with smaller loadings. This can make interpretation difficult (Field, 2005). Component rotation is a useful technique to use to maximise an items loading onto one component and minimise its loadings on the others (Tabachnick and Fidell, 2007a). There are two types of rotation available; orthogonal and oblique. Orthogonal rotation rotates components whilst keeping them independent of one another. This form of rotation should be used when components are conceptually considered to be uncorrelated, and when there is a negligible correlation between the extracted components. Oblique rotation allows components to correlate. This form of rotation should be used when the component structure is found to be correlated (Field, 2005), although this technique does not require components to be correlated (Tabachnick and Fidell, 2007a). Therefore, if an oblique
rotation finds that the solution with the best simple structure provides orthogonal components, the solution produced would be identical to that produced by an orthogonal rotation (Fabrigar et al., 1999). However, in some circumstances there are compelling reasons to use orthogonal rotation even though components are correlated: for instance, the requirement for orthogonal factors in subsequent analyses, or a theoretical need for orthogonal rotation (Tabachnick and Fidell, 2007a).

4.3 Freeman’s Factors

Freeman et al. (1996) asked 170 women who were seeking medical treatment for PMS to complete the DSR for two menstrual cycles. They conducted a PCA using orthogonal (Varimax) rotation on the premenstrual scores (sums of ratings for each symptom on cycle days 23-28) of women meeting PMS criteria. During this analysis, they used the KI rule (Kaiser, 1960) for factor retention and found five factors to have eigenvalues greater than one. However, as their fifth factor only had one item loading (food cravings), they examined and retained a forced four factor solution. Freeman’s factors, along with their item loadings and Cronbach’s alpha coefficients are displayed in Table 4.1 below.

Table 4.1 Factors derived from DSR items (Freeman et al., 1996)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Item</th>
<th>Factor loading</th>
<th>Cronbach’s α co-efficient (n=170)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mood</td>
<td>Anxiety</td>
<td>0.77</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>Irritability</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Depression</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nervous tension</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mood swing</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Feeling out of control</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>2. Behavioural</td>
<td>Poor coordination</td>
<td>0.76</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>Insomnia</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Confusion</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Headache</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Crying</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fatigue</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>3. Pain</td>
<td>Aches</td>
<td>0.76</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>Cramps</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Breast tenderness</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>4. Physical</td>
<td>Food craving</td>
<td>0.83</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>Swelling</td>
<td>0.63</td>
<td></td>
</tr>
</tbody>
</table>
4.4 Reasons for conducting PCAs

Some issues can be raised with Freeman et al.’s (1996) analysis. Firstly, it has been argued that factors should not be made up from less than three items (Zwick and Velicer, 1986). As can be seen in Table 4.1, Freeman et al.’s (1996) ‘physical’ factor comprised only the items ‘food craving’ and ‘swelling.’ Secondly, a good factor structure “makes sense” (Tabachnick and Fidell, 2007a, p608). In Freeman et al.’s (1996) model, the symptom ‘headache’ could load more meaningfully onto the ‘pain’ rather than the ‘behavioural’ factor, since the other items on the behavioural subscale do not relate to pain, whilst headaches can be painful. Most importantly, they used the K1 rule (Kaiser, 1960) to retain factors and opted for a forced four factor solution despite two factors having three or less items. As discussed in Section 4.1.2, this method has been heavily criticised as a method of factor retention (Fabrigar et al., 1999; Gorsuch, 1997; Zwick and Velicer, 1986).

PCA produces factors that are only applicable to the sample that has been studied (Field, 2005). Freeman et al. (1996) performed their analysis on women seeking medical treatment for PMS and who met a provisional diagnosis of PMS through the use of the Structured Clinical Interview for DSM-III-R (SCID, Spitzer et al., 1988). Although they confirmed PMS diagnosis with two cycles of daily symptom reporting, they did not state what criteria were used. The women participating in this research presented in this thesis also self-diagnosed themselves to suffer from PMS and women with co-morbidities were also excluded. However, PMS was only confirmed in women who demonstrated a 30% follicular to luteal symptom increase in two out of three menstrual cycles (see section 3.5.2.1). Therefore, as these criteria may have differed to those employed by Freeman et al. (1996), it was necessary to evaluate the factor structure produced from the women who participated in this research.

Freeman et al. only assessed women meeting PMS criteria when they were premenstrual. However, researchers often use PMS measures to assess symptoms during other menstrual cycle phases. Moreover, follicular to luteal change scores are often used in PMS diagnosis (NIMH, 1983) and in the assessment of treatment efficacy (e.g. Collins et al., 1993; De Souza et al., 2000; Doll et al., 1989; Freeman et al., 1990). Therefore, it was of interest and importance to examine whether the factor structure held for other menstrual cycle phases. Another aim of the thesis was to explore symptom reporting differences between women
who did and did not meet PMS criteria. Therefore, it was necessary to examine whether symptoms clustered together in a similar way for the women who were not categorised into the PMS diagnostic group.

Data from a previous study in the HARU indicated that the order of DSR items might influence reporting behaviour. Therefore, three alternative orders of the DSR were created by randomising the items, so that the DSR was presented to each participant in four orders (see section 3.1.1.6). The items ‘anger,’ ‘aggression,’ and ‘impulsiveness’ were added to the DSR to examine whether the addition of these under researched items could aid PMS diagnosis and assessment of treatment efficacy (see section 3.1.1.1). This meant that items were presented to participants in a different manner to that in Freeman et al.’s study, adding to the need to evaluate the factor structure produced.

A series of PCAs were performed to assess how items on the DSR grouped together in a PMS sample similar to that used by Freeman et al. (1996), and to assess the stability of these factors during different cycle phases, in different diagnostic groups and when items were presented to participants in different orders. It was also necessary to assess the effect of the addition of the items ‘anger,’ ‘aggression,’ and ‘impulsiveness.’ As there are a variety of statistical techniques available to extract factors (Cattell, 1966; Horn, 1965; Kaiser, 1960), which often produce different results (Hayton et al., 2004; Zwick and Velicer, 1986), the factor structures suggested by PA, the K1 rule and the scree plot were explored and are described below.

4.5 Method

4.5.1 Participants

The sample comprised of the 153 women who participated in the screening cycles for long enough for a diagnosis to be made i.e. at least two cycles (see section 3.5). This resulted in 68 PMS sufferers, 33 PMS non-reporters, 17 controls and 30 co-morbid drivers providing daily ratings for approximately 85-86 days (see section 4.5.3). 18 women belonging to the co-morbid group also met the 30% follicular to luteal symptom increase criteria (see section 3.5.1), while 12 did not. 142 of the 153 women for whom a diagnosis could be made provided daily ratings for the full three screening cycles, while 11 women withdrew from the study during the third screening cycle.
4.5.2 Measures
Participants completed the DSR with the items ‘anger,’ ‘aggression’ and ‘impulsiveness’ added to the original 17 items (Freeman et al., 1996) daily for up to three menstrual cycles. These 20 items were presented to participants in four different orders (see section 3.1.1.6). The seven DSR checklists in each daily diary booklet (see section 3.1.1) consisted of the same item order. However, diaries were administered to women in a semi random order (see section 3.1.1.6), such that each woman received approximately the same number of each diary order.

4.5.3 Procedure
Women were asked to complete the DSR each evening for three menstrual cycles. The full method employed for the screening cycles is detailed in Section 3.4. Each woman completed the DSR for approximately 85-86 days. The menstrual cycle was divided into four phases for the purposes of this research; bleeding (cycle days 1-4), follicular (days 5-10), luteal (days -1 to -6) and rest (remaining days between the follicular and luteal phases) (see section 3.3). Therefore, women who provided a full set of daily ratings throughout the screening cycles provided 12 DSR completions during the bleeding phase, 18 during the follicular phase and 18 during the luteal phase. As cycle length varies both between and within women (Chiazze et al., 1968; McIntosh et al., 1980; Sherwood, 2004), the number of diary completions during the rest phase varied.

4.6 Process of Analysis
A series of PCAs with orthogonal rotation were conducted using SPSS (version 14). PCA was favoured over FA to summarise the DSR, in a manner comparable to Freeman et al.’s analysis. Although it is possible that the components produced from the PCAs may correlate, orthogonal rotations were used because the components produced from the analyses were to be used to compare different groups of women using multivariate statistical tests (Tabachnick and Fidell, 2007a). The impact of using orthogonal, as opposed to oblique, rotations is likely to be small, as robustness across factor rotation methods has been reported (Zwick and Velicer, 1986). The alternative, to use oblique rotations followed by univariate statistical tests, would have resulted in a loss of degrees of freedom, and would not have allowed participants to have been compared along combinations of scores
on the components (Tabachnick and Fidell, 2007a). Varimax rotation (Kaiser, 1958) was chosen, based on Fabrigar et al.’s (1999) recommendations.

Factor retention was based on Parallel Analysis (Horn, 1965). The number of factors suggested by both the average and the 95th percentile eigenvalues from the random correlation matrices were calculated. If these differed, the 95th percentile values were favoured (e.g. Glorfeld, 1995; Harshman and Reddon, 1983; Zwick and Velicer, 1986). The factor structures suggested by the Scree test (Cattell, 1966) and the K1 rule (Kaiser, 1960) were also examined. Although the use of the K1 rule is problematic (Fabrigar et al., 1999; McWilliams et al., 2001; Zwick and Velicer, 1986), these solutions are included in the results presented in this chapter so that they can be compared to Freeman et al.’s (1996) factor structure.

PCAs were run on:

i). **Premenstrual model**: Firstly, the original 17-item scale was assessed, where those participants who most closely matched Freeman et al.’s (1996) sample were included in an initial PCA analysis i.e. all women who self-diagnosed with PMS, had PMS prospectively confirmed, and did not meet criteria for anxiety and depression and who were premenstrual (see section 3.5).

ii). **Women during other menstrual cycle phases**: The factor structure of the DSR was assessed during the other menstrual cycle phases of women meeting PMS criteria, by comparing the premenstrual model to PCAs produced from data collected during the ‘bleeding,’ ‘follicular’ and ‘rest’ phases (see section 3.3).

iii). **The four diary orders**: Order effects on the DSR were assessed by organising the data from women meeting PMS criteria according to each of the four diary orders and running four separate PCAs, which were then compared to the premenstrual model.

iv). **Premenstrual women in other diagnostic groups**: In order to determine whether the factor structure held for women not meeting PMS criteria, PCAs were performed on data collected during the premenstrual phase from women who belonged to other diagnostic groups (see sections 3.5.1 and 3.5.2.1). PCAs could only be conducted on those groups
with a large enough sample size. Although numerous suggestions have been put forward as to the adequate sample size on which to conduct a PCA (Comrey, 1978; Comrey and Lee, 1992; Kass and Tinsley, 1979; Tabachnick and Fidell, 2007a), only those samples having at least ten diary completions for each item were considered (Nunnally, 1978). Therefore, the following groups of women were analysed:

1. Controls: Women who did not self-diagnose themselves to suffer from PMS and who did not meet PMS criteria

2. PMS non-reporters: Women who did not self-diagnose themselves to suffer from PMS but who did meet PMS criteria

3. Co-morbidity with symptom increase: Women who met anxiety and/or depression criteria and who also met the 30% follicular to luteal increase criteria (see section 3.5.1)

4. Co-morbidity without symptom increase: Women who met anxiety and/or depression criteria but who did not meet the 30% follicular to luteal increase criteria (see section 3.5.1)

v). The 20 item modification of the DSR: In order to assess the factor structure of the 20-item scale, this process (i. to iv.) was repeated on the original 17, and additional three items.

4.7 Results

All PCAs were found to be satisfactory according to Field’s (2005) recommendations. All models were shown to have good sampling adequacy (Kaiser, 1974), as all Kaiser-Meyer-Olkin statistics were greater than 0.8 (min: 0.83, max: 0.94). There were sufficient correlations between variables in each PCA, indicated by the highly significant Bartlett’s sphericity statistics (p<0.01). The problem of multicollinearity was not evident, indicated by the determinants of the correlation matrices being greater than 1.00E-05. Moreover, all models were shown to fit the data well, indicated by less than 50% of residuals being greater than 0.05 (min: 32%, max: 49%).
The results from the PCAs that were performed on the 17- and 20-item scales are displayed in Tables 4.2-4.8. The factor solutions suggested by the scree plots are displayed in Table 4.2. Tables 4.3-4.8 display the factor solutions for the premenstrual model, and for each diary order, menstrual cycle phase and diagnostic group, for both the 17- and 20-item scales. Each table presents the factor solutions, including factor loadings and the percentage of variance explained by each factor. Cross-loading items are marked with an asterisk. Bold type indicates the primary factor on which an item loads; italic type the secondary. For presentation purposes, some items have been abbreviated. These items are indicated in the key.

Table 4.2. Summary table of the number of factors produced by the scree plot for each of the 17- and 20-item models

<table>
<thead>
<tr>
<th>Model</th>
<th>17-item scale</th>
<th>20-item scale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Premenstrual model</strong></td>
<td>Unclear (2 or 3)</td>
<td>2 factors</td>
</tr>
<tr>
<td><strong>Cycle phase:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Menstrual</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>- Follicular</td>
<td>Unclear (2 or 4)</td>
<td>Unclear (2 or 5)</td>
</tr>
<tr>
<td>- Rest</td>
<td>2</td>
<td>Unclear (2 or 4)</td>
</tr>
<tr>
<td><strong>Diary order:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- One</td>
<td>Unclear (2 or 3)</td>
<td>2</td>
</tr>
<tr>
<td>- Two</td>
<td>2</td>
<td>Unclear (2 or 4)</td>
</tr>
<tr>
<td>- Three</td>
<td>Unclear (2 or 4)</td>
<td>Unclear (2 or 4)</td>
</tr>
<tr>
<td>- Four</td>
<td>Unclear (2 or 4)</td>
<td>4</td>
</tr>
<tr>
<td><strong>Diagnostic groups</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- PMS non-reporters</td>
<td>3 factors</td>
<td>3</td>
</tr>
<tr>
<td>- Co-morbidity with symptom increase</td>
<td>Unclear (2 or 3)</td>
<td>3</td>
</tr>
<tr>
<td>- Co-morbidity without symptom increase</td>
<td>Too unclear to specify</td>
<td>Too unclear to specify</td>
</tr>
</tbody>
</table>
Table 4.3 Summary of factor loadings produced by PA for the 17-item models

<table>
<thead>
<tr>
<th>CYCLE PHASE</th>
<th>DIARY ORDER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ONE (n=310 days)</td>
</tr>
<tr>
<td>PREMENSTRUAL (n=1210 days)</td>
<td>2 fac. (50.17%)</td>
</tr>
<tr>
<td>MENSTRUAL (n=795 days)</td>
<td>2 fac. (57.40%)</td>
</tr>
<tr>
<td>FOLLICULAR (n=1206 days)</td>
<td>2 fac. (50.17%)</td>
</tr>
<tr>
<td>REST (n=2565 days)</td>
<td>2 fac. (43.11%)</td>
</tr>
</tbody>
</table>

Key for tables 4.3 - 4.8: *: cross-loading item (cross loading item in italics indicates lowest loading). Fac: factor. %: percentage of variance explained after rotation.

m.s: mood swings  n.t: nervous tension  b.t: breast tenderness  p.c: poor coordination  f.c: food cravings  control: feeling out of control

N.B. Numbers in brackets after each item denote factor loadings (only loadings greater than 0.4 displayed).
<table>
<thead>
<tr>
<th></th>
<th>CYCLE PHASE</th>
<th>DIARY ORDER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PREMENSTRUAL (n=1210 days)</td>
<td>MENSTRUAL (n=795 days)</td>
</tr>
<tr>
<td></td>
<td>FOLLICULAR (n=1206 days)</td>
<td>REST (n=2565 days)</td>
</tr>
<tr>
<td>3 fac. (56.90%)</td>
<td>2 fac. (57.40%)</td>
<td>4 fac. (56.19%)</td>
</tr>
<tr>
<td>4 fac. (58.13%)</td>
<td></td>
<td>4 fac. (58.13%)</td>
</tr>
<tr>
<td>1. (25.64%)</td>
<td>depression (.78)</td>
<td>control (.76)</td>
</tr>
<tr>
<td>depression (.83)</td>
<td>m.s. (.74)</td>
<td>anxiety (.71)</td>
</tr>
<tr>
<td>anxiety (.71)</td>
<td>crying (.69)</td>
<td>irritation (.68)</td>
</tr>
<tr>
<td>irritation (.68)</td>
<td>n.t. (.65)</td>
<td>confusion* (.48)</td>
</tr>
<tr>
<td>confusion* (.48)</td>
<td>p.c* (.51)</td>
<td>fatigue* (.45)</td>
</tr>
<tr>
<td></td>
<td>f.c* (.44)</td>
<td></td>
</tr>
<tr>
<td>2. (17.29%)</td>
<td>swelling (.75)</td>
<td>b.t (.75)</td>
</tr>
<tr>
<td>swelling (.75)</td>
<td>cramps (.77)</td>
<td>aches (.77)</td>
</tr>
<tr>
<td>aches (.57)</td>
<td>b.t (.61)</td>
<td>n.t* (.43)</td>
</tr>
<tr>
<td>fatigue (.52)</td>
<td>insomnia (.58)</td>
<td></td>
</tr>
<tr>
<td>confusion* (.51)</td>
<td>fatigue* (.57)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>f.c* (.53)</td>
<td>headache (.53)</td>
</tr>
<tr>
<td></td>
<td>p.c* (.51)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>confusion* (.48)</td>
<td></td>
</tr>
<tr>
<td>2. (12.10%)</td>
<td>swelling (.73)</td>
<td>b.t (.66)</td>
</tr>
<tr>
<td>swelling (.81)</td>
<td>fatigue (.64)</td>
<td>aches* (.55)</td>
</tr>
<tr>
<td>aches (.55)</td>
<td>fatigue* (.44)</td>
<td>cramps (.62)</td>
</tr>
<tr>
<td></td>
<td>n.t* (.43)</td>
<td></td>
</tr>
<tr>
<td>2. (13.27%)</td>
<td>swimming (.71)</td>
<td></td>
</tr>
<tr>
<td>2. (12.10%)</td>
<td>headache (.73)</td>
<td></td>
</tr>
<tr>
<td>2. (12.10%)</td>
<td>swelling (.71)</td>
<td></td>
</tr>
<tr>
<td>2. (15.62%)</td>
<td>aches (.76)</td>
<td>b.t (.80)</td>
</tr>
<tr>
<td>2. (15.62%)</td>
<td>cramps (.75)</td>
<td>m.s (.84)</td>
</tr>
<tr>
<td>2. (15.62%)</td>
<td>aches (.76)</td>
<td>b.t (.70)</td>
</tr>
<tr>
<td>2. (15.62%)</td>
<td>m.s (.84)</td>
<td>m.s* (.60)</td>
</tr>
<tr>
<td>2. (15.62%)</td>
<td>cramps (.55)</td>
<td>n.t (.56)</td>
</tr>
<tr>
<td>2. (15.62%)</td>
<td>headache (.46)</td>
<td></td>
</tr>
<tr>
<td>2. (15.62%)</td>
<td>headache*. (.45)</td>
<td></td>
</tr>
<tr>
<td>2. (15.62%)</td>
<td>swelling*. (.41)</td>
<td></td>
</tr>
<tr>
<td>2. (15.62%)</td>
<td>swelling*. (.41)</td>
<td></td>
</tr>
<tr>
<td>1. (31.27%)</td>
<td>anxiety (.83)</td>
<td>control*. (.64)</td>
</tr>
<tr>
<td>anxiety (.83)</td>
<td>depression*. (.66)</td>
<td>control*. (.70)</td>
</tr>
<tr>
<td>depression (.82)</td>
<td>m.s (.74)</td>
<td>m.s* (.60)</td>
</tr>
<tr>
<td>m.s (.77)</td>
<td>crying (.70)</td>
<td>n.t (.56)</td>
</tr>
<tr>
<td>irritation (.74)</td>
<td>irritation (.55)</td>
<td></td>
</tr>
<tr>
<td>irritation (.68)</td>
<td>n.t (.74)</td>
<td></td>
</tr>
<tr>
<td>n.t (.65)</td>
<td>crying (.69)</td>
<td></td>
</tr>
<tr>
<td>n.t*. (.41)</td>
<td>confusion (.62)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>fatigue*. (.49)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>f.c* (.49)</td>
<td></td>
</tr>
<tr>
<td>1. (15.87%)</td>
<td>swelling (.78)</td>
<td>m.s (.77)</td>
</tr>
<tr>
<td>swelling (.78)</td>
<td>control*. (.65)</td>
<td>control*. (.56)</td>
</tr>
<tr>
<td>b.t (.66)</td>
<td>headache*. (.44)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>insomnia*. (.40)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>fatigue*. (.41)</td>
<td></td>
</tr>
<tr>
<td>2. (16.56%)</td>
<td>swelling (.76)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>m.s (.84)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>control*. (.80)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>control*. (.55)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n.t (.56)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>anxiety*. (.56)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>irritation*. (.55)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>control*. (.43)</td>
<td></td>
</tr>
<tr>
<td>CYCLE PHASE</td>
<td>DIARY ORDER</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------</td>
<td></td>
</tr>
<tr>
<td>PREMENSTRUAL</td>
<td>MENSTRUAL</td>
<td></td>
</tr>
<tr>
<td>(n=1210 days)</td>
<td>(n=795 days)</td>
<td></td>
</tr>
<tr>
<td>3 fac. (56.90%)</td>
<td>2 fac. (57.40%)</td>
<td></td>
</tr>
<tr>
<td>4 fac. (56.19%)</td>
<td>4 fac. (58.13%)</td>
<td></td>
</tr>
<tr>
<td>4 fac. (56.92%)</td>
<td>5 fac. (63.50%)</td>
<td></td>
</tr>
<tr>
<td>4 fac. (66.05%)</td>
<td>4 fac. (63.81%)</td>
<td></td>
</tr>
<tr>
<td>3. (13.98%)</td>
<td>3. (10.74%)</td>
<td></td>
</tr>
<tr>
<td>aches (.75)</td>
<td>swelling (.80)</td>
<td></td>
</tr>
<tr>
<td>swelling (.80)</td>
<td>confusion (.71)</td>
<td></td>
</tr>
<tr>
<td>b.t (.68)</td>
<td>p.c (.69)</td>
<td></td>
</tr>
<tr>
<td>p.c (.70)</td>
<td>n.t* (.56)</td>
<td></td>
</tr>
<tr>
<td>headache (.63)</td>
<td>aches* (.49)</td>
<td></td>
</tr>
<tr>
<td>headache (.65)</td>
<td>anxiety* (.53)</td>
<td></td>
</tr>
<tr>
<td>insomnia (.54)</td>
<td>f.c* (.43)</td>
<td></td>
</tr>
<tr>
<td>insomnia (.64)</td>
<td>f.c* (.43)</td>
<td></td>
</tr>
<tr>
<td>cramps (.81)</td>
<td>f.c (.78)</td>
<td></td>
</tr>
<tr>
<td>p.c (.66)</td>
<td>cramps (.80)</td>
<td></td>
</tr>
<tr>
<td>cramps (.72)</td>
<td>aches (.72)</td>
<td></td>
</tr>
<tr>
<td>p.c* (.65)</td>
<td>cramps (.72)</td>
<td></td>
</tr>
<tr>
<td>confusion* (.50)</td>
<td>headache* (.49)</td>
<td></td>
</tr>
<tr>
<td>attention* (.40)</td>
<td>headache* (.62)</td>
<td></td>
</tr>
<tr>
<td>depression* (.49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. (9.16%)</td>
<td>4. (12.14%)</td>
<td></td>
</tr>
<tr>
<td>4. (12.17%)</td>
<td>4. (11.62%)</td>
<td></td>
</tr>
<tr>
<td>4. (11.77%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. (8.17%)</td>
<td>insomnia (.84)</td>
<td></td>
</tr>
<tr>
<td>PMS non-reporters (n=565 days)</td>
<td>Co-morbid with a symptom increase (n=316 days)</td>
<td>Co-morbid, no symptom increase (n=171 days)</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------------------------------------------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td><strong>PA</strong></td>
<td><strong>K1</strong></td>
<td><strong>PA</strong></td>
</tr>
<tr>
<td>2 fac (47.67%)</td>
<td>3 fac. (54.50%)</td>
<td>3 fac. (58.21%)</td>
</tr>
<tr>
<td>1. (29.36%)</td>
<td>1. (29.13%)</td>
<td>1. (23.61%)</td>
</tr>
<tr>
<td>n.t. (.80)</td>
<td>anxiety (.78)</td>
<td>control (.78)</td>
</tr>
<tr>
<td>anxiety (.80)</td>
<td>control (.74)</td>
<td>m.s. (.74)</td>
</tr>
<tr>
<td>control (.75)</td>
<td>anxiety (.74)</td>
<td>depression (.76)</td>
</tr>
<tr>
<td>crying (.64)</td>
<td>irritation (.74)</td>
<td>crying (.74)</td>
</tr>
<tr>
<td>m.s. (.63)</td>
<td>p.c. (.62)</td>
<td>confusion (.70)</td>
</tr>
<tr>
<td>insomnia (.61)</td>
<td>p.c. (.62)</td>
<td>m.s. (.67)</td>
</tr>
<tr>
<td>confusion (.61)</td>
<td>p.c. (.61)</td>
<td>depression (.66)</td>
</tr>
<tr>
<td>irritability (.61)</td>
<td>fatigue* (.46)</td>
<td>fatigue (.57)</td>
</tr>
<tr>
<td>depression (.57)</td>
<td>control* (.44)</td>
<td>f.c. (.55)</td>
</tr>
<tr>
<td>p.c. (.55)</td>
<td>headache* (.47)</td>
<td>f.c. (.45)</td>
</tr>
<tr>
<td>fatigue* (.50)</td>
<td>headache* (.40)</td>
<td></td>
</tr>
<tr>
<td>2. (18.31%)</td>
<td>2. (15.93%)</td>
<td>2. (19.19%)</td>
</tr>
<tr>
<td>swelling (.72)</td>
<td>irritability (.81)</td>
<td>swelling (.82)</td>
</tr>
<tr>
<td>b.t. (.71)</td>
<td>m.s. (.80)</td>
<td>aches (.78)</td>
</tr>
<tr>
<td>cramps (.67)</td>
<td>depression (.56)</td>
<td>b.t. (.75)</td>
</tr>
<tr>
<td>aches (.64)</td>
<td>f.c. (.52)</td>
<td>cramps (.68)</td>
</tr>
<tr>
<td>p.c. (.51)</td>
<td>headache* (.46)</td>
<td>insomnia (.46)</td>
</tr>
<tr>
<td>headache* (.48)</td>
<td>f.c*. (.47)</td>
<td>p.c*. (.47)</td>
</tr>
<tr>
<td>fatigue* (.43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. (15.80%)</td>
<td>3. (15.41%)</td>
<td>3. (12.58%)</td>
</tr>
<tr>
<td>aches (.72)</td>
<td>irritability (.83)</td>
<td>crying (.79)</td>
</tr>
<tr>
<td>swelling (.67)</td>
<td>m.s. (.75)</td>
<td>aches (.88)</td>
</tr>
<tr>
<td>cramps (.66)</td>
<td>f.c*. (.55)</td>
<td>depression (.63)</td>
</tr>
<tr>
<td>b.t. (.66)</td>
<td>control* (.52)</td>
<td>fatigue* (.62)</td>
</tr>
<tr>
<td>headache* (.52)</td>
<td>n.t*. (.43)</td>
<td>insomnia (.57)</td>
</tr>
<tr>
<td>fatigue* (.41)</td>
<td></td>
<td>f.c. (.43)</td>
</tr>
</tbody>
</table>
### Table 4.5 continued

<table>
<thead>
<tr>
<th>PMS non-reporters (n=565 days)</th>
<th>Co-morbids with a symptom increase (n=316 days)</th>
<th>Co-morbids, no symptom increase (n=171 days)</th>
<th>Controls (n=280 days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA 2 fac (47.67%)</td>
<td>PA 2 fac. (51.36%)</td>
<td>PA 3 fac. (56.59%)</td>
<td>PA 3 fac. (50.00%)</td>
</tr>
<tr>
<td>K1 3 fac. (54.50%)</td>
<td>K1 3 fac. (58.21%)</td>
<td>K1 5 fac. (69.90%)</td>
<td>K1 5 fac. (63.04%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. crying (.72)</td>
<td>4. crying* (.70)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>depression* (.69)</td>
<td>depression* (.63)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>irritability* (.55)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>m.s* (.49)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. (8.26%)</td>
<td>5. (9.46%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>insomnia (.75)</td>
<td>cramps (.62)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>headache (.59)</td>
<td>headache (.60)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>fatigue* (.56)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>aches* (.46)</td>
</tr>
</tbody>
</table>
Table 4.6. Summary of factor loadings produced by PA and the K1 rule for the different diagnostic groups for the 20-item models

<table>
<thead>
<tr>
<th>PMS non-reporters (n=565 days)</th>
<th>Co-morbid with a symptom increase (n=316 days)</th>
<th>Co-morbid, no symptom increase (n=171 days)</th>
<th>Controls (n=280 days)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PA</strong></td>
<td><strong>K1</strong></td>
<td><strong>K1</strong></td>
<td><strong>K1</strong></td>
</tr>
<tr>
<td>3 fac. (54.25%)</td>
<td>2 fac. (52.87%)</td>
<td>4 fac. (59.14%)</td>
<td>5 fac. (61.09%)</td>
</tr>
<tr>
<td>1. (21.33%)</td>
<td>1. (27.05%)</td>
<td>1. (27.65%)</td>
<td>1. (23.65%)</td>
</tr>
<tr>
<td>anxiety (.74)</td>
<td>anger (.82)</td>
<td>anger (.84)</td>
<td>anger (.87)</td>
</tr>
<tr>
<td>n.t. (.73)</td>
<td>aggression (.82)</td>
<td>irritability (.82)</td>
<td>m.s. (.76)</td>
</tr>
<tr>
<td>insomnia (.72)</td>
<td>irritability (.79)</td>
<td>m.s. (.76)</td>
<td>m.s. (.75)</td>
</tr>
<tr>
<td>control (.67)</td>
<td>m.s. (.78)</td>
<td>control (.73)</td>
<td>control (.72)</td>
</tr>
<tr>
<td>confusion (.61)</td>
<td>depression (.53)</td>
<td>p.c. (.61)</td>
<td>n.t. (.63)</td>
</tr>
<tr>
<td>p.c. (.55)</td>
<td>impulsiveness (.47)</td>
<td>anxiety* (.62)</td>
<td>anxiety* (.60)</td>
</tr>
<tr>
<td>impulsiveness* (.47)</td>
<td>pause* (.47)</td>
<td>confusion* (.50)</td>
<td>pause* (.48)</td>
</tr>
<tr>
<td>headache* (.46)</td>
<td>depression* (.49)</td>
<td>control* (.53)</td>
<td>depression* (.45)</td>
</tr>
<tr>
<td></td>
<td>impulsiveness* (.42)</td>
<td>aches* (.47)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>headache* (.46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. (18.89%)</td>
<td>2. (17.01%)</td>
<td>2. (13.05%)</td>
<td>2. (15.75%)</td>
</tr>
<tr>
<td>aggression (.83)</td>
<td>anxiety (.77)</td>
<td>anger (.86)</td>
<td>fatigue (.78)</td>
</tr>
<tr>
<td>anger (.82)</td>
<td>n.t. (.74)</td>
<td>aggression (.86)</td>
<td>aches (.72)</td>
</tr>
<tr>
<td>irritability (.80)</td>
<td>insomnia (.68)</td>
<td>irritability (.83)</td>
<td>p.c* (.68)</td>
</tr>
<tr>
<td>m.s. (.79)</td>
<td>control (.64)</td>
<td>m.s. (.74)</td>
<td>fatigue (.62)</td>
</tr>
<tr>
<td>depression (.50)</td>
<td>crying (.50)</td>
<td>control* (.60)</td>
<td>f.c* (.61)</td>
</tr>
<tr>
<td>f.c* (.41)</td>
<td></td>
<td>headache (.54)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. (17.38%)</td>
<td>3. (17.02%)</td>
<td>2. (13.04%)</td>
<td>2. (12.89%)</td>
</tr>
<tr>
<td>anxiety (.77)</td>
<td>anger (.86)</td>
<td>swelling* (.65)</td>
<td>f.c*. (.65)</td>
</tr>
<tr>
<td>n.t. (.74)</td>
<td>swelling (.78)</td>
<td>m.s. (.68)</td>
<td>p.c* (.68)</td>
</tr>
<tr>
<td>depression (.53)</td>
<td>cramps (.69)</td>
<td>control (.53)</td>
<td>control* (.50)</td>
</tr>
<tr>
<td>p.c. (.61)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>depression* (.44)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: * indicates loadings above .40.
Table 4.6 continued

<table>
<thead>
<tr>
<th>PMS non-reporters (n=565 days)</th>
<th>Co-morbid with a symptom increase (n=316 days)</th>
<th>Co-morbid, no symptom increase (n=171 days)</th>
<th>Controls (n=280 days)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PA</strong></td>
<td><strong>K1</strong></td>
<td><strong>PA</strong></td>
<td><strong>K1</strong></td>
</tr>
<tr>
<td>3 fac. (54.25%)</td>
<td>4 fac. (59.52%)</td>
<td>4 fac. (61.75%)</td>
<td>5 fac. (67.07%)</td>
</tr>
<tr>
<td>3. (14.03%)</td>
<td>3. (13.05%)</td>
<td>3. (10.92%)</td>
<td>3. (10.69%)</td>
</tr>
<tr>
<td>b.t (.69)</td>
<td>cramps (.73)</td>
<td>b.t (.77)</td>
<td>swelling (.79)</td>
</tr>
<tr>
<td>swelling (.68)</td>
<td>aches (.72)</td>
<td>swelling (.73)</td>
<td>f.c (.67)</td>
</tr>
<tr>
<td>aches (.68)</td>
<td>swelling (.62)</td>
<td>aches* (.69)</td>
<td>swelling (.75)</td>
</tr>
<tr>
<td>cramps (.66)</td>
<td>b.t (.60)</td>
<td>cramps (.65)</td>
<td>b.t (.53)</td>
</tr>
<tr>
<td>headache* (.51)</td>
<td>headache (.52)</td>
<td>f.c (.54)</td>
<td>b.t (.53)</td>
</tr>
<tr>
<td>f.c* (.43)</td>
<td>fatigue (.40)</td>
<td>insomnia (.48)</td>
<td></td>
</tr>
<tr>
<td><strong>fatigue</strong> (.41)</td>
<td></td>
<td>headache* (.48)</td>
<td></td>
</tr>
</tbody>
</table>

| 4. (11.04%)                   | 4. (10.13%)                                  | 4. (10.34%)                                | 4. (9.24%)           |
| impulsiveness (.68)           | crying (.75)                                 | b.t (.85)                                  | anxiety (.76)        |
| p.c (.61)                     | depression* (.65)                            | cramps (.76)                               | n.t (.74)            |
| confusion (.57)               | insomnia (.54)                               | swelling* (.41)                            | control* (.43)       |
| f.c (.43)                     |                                              | anxiety* (.40)                             |                       |

| 5. (9.13%)                    | 5. (9.00%)                                   | 5. (9.00%)                                 |                       |
| crying (.75)                  | cramps (.65)                                 | headache (.56)                             |                       |
| insomnia (.74)                | depression* (.46)                            | depression (.50)                           |                       |
|                               |                                              | fatigue* (.53)                             |                       |
|                               |                                              | aches* (.44)                               |                       |
Table 4.7 Summary of factor loadings produced by PA for the 20-item models

<table>
<thead>
<tr>
<th>CYCLE PHASE</th>
<th>DIARY ORDER</th>
</tr>
</thead>
<tbody>
<tr>
<td>PREMENSTRUAL (n=1210 days)</td>
<td>MENSTRUAL (n=795 days)</td>
</tr>
<tr>
<td>2 fac. (50.26%)</td>
<td>3 fac. (56.97%)</td>
</tr>
<tr>
<td>1. (28.37%)</td>
<td>1. (33.02%)</td>
</tr>
<tr>
<td>anger (.85)</td>
<td>aggression (.84)</td>
</tr>
<tr>
<td>m.s. (.79)</td>
<td>depression (.80)</td>
</tr>
<tr>
<td>irritability (.75)</td>
<td>irritability (.70)</td>
</tr>
<tr>
<td>depression (.74)</td>
<td>control (.74)</td>
</tr>
<tr>
<td>anxiety (.64)</td>
<td>anxiety (.70)</td>
</tr>
<tr>
<td>crying (.60)</td>
<td>crying (.60)</td>
</tr>
<tr>
<td>n.t* (.57)</td>
<td>crying (.68)</td>
</tr>
<tr>
<td>insensitivity (.44)</td>
<td>confusions* (.57)</td>
</tr>
<tr>
<td>confusion* (.45)</td>
<td>insensitivity*. (.52)</td>
</tr>
<tr>
<td>2. (21.89%)</td>
<td>2. (23.95%)</td>
</tr>
<tr>
<td>swelling (.74)</td>
<td>swelling (.80)</td>
</tr>
<tr>
<td>p.c. (.65)</td>
<td>aches (.76)</td>
</tr>
<tr>
<td>fatigue (.62)</td>
<td>cramps (.76)</td>
</tr>
<tr>
<td>confusion*. (.61)</td>
<td>b.t. (.61)</td>
</tr>
<tr>
<td>cramps (.60)</td>
<td>fatigue (.60)</td>
</tr>
<tr>
<td>headache (.60)</td>
<td>insomnia (.58)</td>
</tr>
<tr>
<td>aches (.59)</td>
<td>b.t. (.65)</td>
</tr>
<tr>
<td>b.t. (.57)</td>
<td>f.e*. (.55)</td>
</tr>
<tr>
<td>f.e*. (.48)</td>
<td>confusion*. (.53)</td>
</tr>
<tr>
<td>n.t* (.46)</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.7 continued

<table>
<thead>
<tr>
<th>CYCLE PHASE</th>
<th>DIARY ORDER</th>
</tr>
</thead>
<tbody>
<tr>
<td>PREMENSTRUAL</td>
<td>MENSTRUAL</td>
</tr>
<tr>
<td>(n=1210 days)</td>
<td>(n=795 days)</td>
</tr>
<tr>
<td>2 fac. (50.26%)</td>
<td>2 fac. (56.97%)</td>
</tr>
<tr>
<td>ONE</td>
<td>TWO</td>
</tr>
<tr>
<td>(n=310 days)</td>
<td>(n=274 days)</td>
</tr>
<tr>
<td>2 fac. (54.87%)</td>
<td>2 fac. (43.03%)</td>
</tr>
</tbody>
</table>

- 3. (10.88%)
- swelling (.72)
- b.t (.65)
- aches (.60)
- cramps (.43)
- insomnia* (.42)
- p.c (.73)
- confusion (.73)
- n.t* (.53)
- impulsiveness (.44)
- anxiety* (.51)
Table 4.8. Summary of factor loadings produced by the K1 rule for the 20-item models

<table>
<thead>
<tr>
<th>PREMENSTRUAL (n=1210 days)</th>
<th>MENSTRUAL (n=795 days)</th>
<th>FOLLICULAR (n=1206 days)</th>
<th>REST (n=2565 days)</th>
<th>CYCLE PHASE</th>
<th>DIARY ORDER</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 fac. (56.04%)</td>
<td>2 fac. (56.97%)</td>
<td>5 fac. (59.27%)</td>
<td>4 fac. (56.66%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. (27.20%)</td>
<td>1. (33.02%)</td>
<td>1. (17.00%)</td>
<td>1. (20.99%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>anger (.84)</td>
<td>aggression (.80)</td>
<td>depression (.80)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m.s. (.77)</td>
<td>n.t. (.66)</td>
<td>control (.64)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>depression (.75)</td>
<td>irritability (.77)</td>
<td>control*. (.64)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>irritability (.73)</td>
<td>control*. (.64)</td>
<td>depression*. (.62)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>control*. (.73)</td>
<td>impulsiveness (.53)</td>
<td>anxiety*. (.42)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anxiety*. (.64)</td>
<td>crying*. (.57)</td>
<td>impulsiveness (.52)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>crying*. (.61)</td>
<td>n.t* (.69)</td>
<td>crying (.68)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n.t* (.56)</td>
<td>confusion* (.57)</td>
<td>impulsiveness (.52)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>impulsiveness (.42)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. (15.09%)</td>
<td>2. (23.95%)</td>
<td>2. (16.05%)</td>
<td>2. (13.51%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>swelling (.76)</td>
<td>swelling (.80)</td>
<td>anger (.80)</td>
<td>swelling (.67)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.t. (.75)</td>
<td>cramps (.76)</td>
<td>irritability (.70)</td>
<td>cramps (.58)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p.c. (.63)</td>
<td>m.s. (.64)</td>
<td>impulsiveness (.48)</td>
<td>m.s*. (.52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f.c. (.56)</td>
<td>f.c* (.55)</td>
<td>T.c. (.46)</td>
<td>aches* (.44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fatigue (.50)</td>
<td>insomnia (.58)</td>
<td>headache (.55)</td>
<td>depression* (.68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>confusion* (.49)</td>
<td></td>
<td></td>
<td>p.c* (.40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. (19.56%)</td>
<td>2. (12.70%)</td>
<td>2. (19.56%)</td>
<td>2. (20.18%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>anger (.87)</td>
<td>aggression (.84)</td>
<td>swelling (.75)</td>
<td>aggression (.85)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.t. (.74)</td>
<td>cramps (.58)</td>
<td>agitation (.81)</td>
<td>m.s*. (.84)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m.s*. (.72)</td>
<td>aches* (.56)</td>
<td>control*. (.60)</td>
<td>control* (.50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>depression* (.58)</td>
<td>insomnia* (.48)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>control*. (.50)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. (18.92%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. (22.15%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>depression (.83)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anger (.81)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aggression* (.74)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>crying (.73)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m.s*. (.66)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>irritability (.63)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>insomnia* (.53)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>control*. (.52)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anxiety*. (.43)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 4.8 continued.

<table>
<thead>
<tr>
<th>PREMENSTRUAL</th>
<th>MENSTRUAL</th>
<th>FOLLICULAR</th>
<th>REST</th>
<th>ONE</th>
<th>TWO</th>
<th>THREE</th>
<th>FOUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n=1210 days)</td>
<td>(n=795 days)</td>
<td>(n=1206 days)</td>
<td>(n=2565 days)</td>
<td>(n=310 days)</td>
<td>(n=274 days)</td>
<td>(n=331 days)</td>
<td>(n=283 days)</td>
</tr>
<tr>
<td>3 fac. (56.04%)</td>
<td>2 fac. (56.97%)</td>
<td>5 fac. (59.27%)</td>
<td>4 fac. (56.66%)</td>
<td>4 fac. (65.22%)</td>
<td>5 fac. (60.92%)</td>
<td>4 fac. (65.02%)</td>
<td>4 fac. (63.50%)</td>
</tr>
<tr>
<td>3. (13.76%) aches (.71) cramps (.68) headache (.64) insomnia (.55) n.t* (.43)</td>
<td>3. (9.74%) headache (.73) fatigue (.64) aches* (.57)</td>
<td>3. (11.66%) confusion (.69) anxiety* (.63) n.t (.63) p.c* (.61) control* (.42)</td>
<td>3. (13.43%) cramps (.76) aches (.74) headache (.64) insomnia* (.64) fatigue* (.46) swelling* (.42)</td>
<td>3. (11.96%) n.t (.85) aches (.74) headache (.64) insomnia* (.64) fatigue* (.46) swelling* (.42)</td>
<td>3. (13.01%) crying (.81) anxiety (.74) confusion (.45) p.c* (.43) aches* (.42) n.t* (.43)</td>
<td>3. (11.84%) b.t (.84) swelling (.78) f.c (.47) fatigue*.46) p.c* (.47)</td>
<td>4. (8.55%) swelling (.78) b.t (.75) aches* (.46)</td>
</tr>
</tbody>
</table>
i). Premenstrual model
As illustrated in Table 4.2, the scree plot did not suggest a clear factor solution for the premenstrual model. The K1 rule produced 3 factors from the current data set (see Table 4.4). The first factor, which explained 25.54% of the variance, seemed to comprise the ‘psychological’ symptoms. Factors two (17.29%) and three (13.98%) explained less variance, and were made up from the ‘physical’ symptoms. When PA was used as the factor retention method, a two factor solution was produced. This solution explained 50.17% of the variance, with each factor explaining almost equal amounts (see Table 4.3). The seven ‘psychological’ items that grouped together to form factor one were identical to those constituting factor one in the K1 solution, while the ten ‘physical’ symptoms that formed factor two were identical to those items that were divided between factors two and three by the K1 rule.

ii). Women with PMS during other menstrual cycle phases
Although the scree plot for the follicular phase model did not provide a clear factor solution, the scree plots for the menstrual and rest cycle phases indicated a clear two factor solution. The K1 rule produced diverse solutions (2-4 factors) across cycle phases, although the item pairs ‘mood swings’ and ‘irritability,’ and ‘breast tenderness’ and ‘swelling,’ generally loaded together, as they did across diary orders. The other items grouped together in very different ways across cycle phases, and had variable factor loadings, especially the items ‘cramps’ and ‘confusion.’

In contrast to the diverse solutions produced by the K1 rule, PA seemed to produce two relatively consistent factors across cycle phases. The seven items that formed factor one in the premenstrual model, and which consistently loaded onto factor one across diary orders, also loaded onto factor one across cycle phases. Six of the ten items that comprised the ‘physical’ factor in the premenstrual model consistently loaded onto factor two across cycle phases, although ‘headache’ had a factor loading of <0.4 for the follicular model. The other four items showed slight variation in their groupings, although the items ‘fatigue,’ ‘food cravings,’ and ‘poor coordination’ did load primarily onto the second factor in two of the three cycle phase models. As for the premenstrual model, the item ‘depression’ had the highest factor loading on factor one, while ‘swelling’ was highest on factor two across cycle phases. ‘Confusion’ often demonstrated cross-loading between the factors. However,
unlike in the premenstrual model, this item primarily loaded onto factor one across cycle phases.

iii). The four diary orders

The scree plots did not suggest a clear factor solution for the majority of the models assessing the separate diary orders (see Table 4.2). Although the scree plots suggested that there was a possible two factor solution for all diary orders, a three factor solution also appeared possible for diary order one, whilst a four factor solution seemed feasible for diary orders three and four.

Table 4.4 illustrates the many differences apparent when comparing the solutions suggested by the KI rule for the separate diary orders. Although the item pairs ‘mood swings’ and ‘irritability,’ and ‘breast tenderness’ and ‘swelling,’ generally loaded on the same factor across models, solutions differed in the number of factors (3-5), groupings of items and item factor loadings. In contrast, PA (table 4.3) produced two relatively consistent factors across diary orders. The seven items that constituted the ‘psychological’ factor in the premenstrual model also loaded on to factor one across all diary orders, while seven of the ten items that loaded onto the ‘physical’ factor in the premenstrual model consistently loaded onto factor two. The other three items that loaded onto factor two in the premenstrual model showed slight variation in their affiliations. ‘Insomnia’ loaded onto factor two in three out of the four diary orders. Item loadings remained fairly consistent across diary orders, with the largest variation (0.22) occurring for the items ‘anxiety’ and ‘nervous tension’ between diary orders one and two.

iv). Premenstrual women in other diagnostic groups

In line with the general pattern observed thus far, the scree plots did not provide a clear factor solution for the co-morbid groups. However, the scree plots for the control and PMS non-reporter groups suggested a three factor solution. As indicated in Table 4.5, the K1 rule suggested quite a similar three factor solution for the PMS non-reporters and the co-morbid group who also showed a symptom increase. However, dissimilar five factor solutions were suggested for the control group and the co-morbid group who did not show a premenstrual symptom increase.
As Table 4.5 illustrates, PA produced more similar factor solutions across groups, although some inconsistencies remained. The two groups that demonstrated at least a 30% premenstrual increase in symptoms showed a similar structure to the premenstrual model, although transference of some items occurred from factor one to two in both of these groups. The model produced from the controls resulted in items being clustered in a very similar way to the premenstrual model, although their ‘physical’ symptoms were subdivided into two separate factors. The largest variation from the premenstrual model was found for the co-morbid group who did not demonstrate a 30% premenstrual increase in symptoms. However, it should be noted that this group also had the smallest number of diary days (n=171) and as such would be expected to be the least stable.

v). The 20 item modification of the DSR
When assessing the solutions produced by the K1 rule for the 20-item scale (see Table 4.8), many similarities were apparent between these models and those produced when only the 17-items were included. The premenstrual solution remained almost identical, with the additional three items (anger, aggression and impulsiveness) loading onto the first factor. However, in some models, the addition of these items did alter the existing structure by dividing factors into two (diary orders one and two) and by moving items between factors. Nevertheless, the additional symptoms ‘anger’ and ‘aggression’ loaded together across all models, and always clustered with the symptoms ‘irritability’ and ‘mood swings.’ The additional item ‘impulsiveness’ showed no such consistency. The models remained similar to those produced from the 17-items, and diverse structures were produced across diary orders and cycle phases.

As can be seen in Table 4.7, the use of PA produced a much more stable pattern across models. The premenstrual model, which explained 50.26% of the variance, was identical to that provided by the 17-item scale, with the additional items loading onto the ‘psychological’ factor. This, or a very similar pattern, was also found for the separate diary orders. The largest difference here was the movement of the items ‘nervous tension’ and ‘confusion’ from factor one to two in diary order two. A similar solution to the premenstrual model was found for the ‘menstrual’ phase. However, slight differences were found for the models based on the data collected during the ‘follicular’ and ‘rest’ phases, as three factors were produced, most of which consisted of different items (see Table 4.7).
The factor solutions produced by PA for the separate diagnostic groups appeared somewhat different to the premenstrual model, and to the structures produced when only the 17-items were considered. Many of the deviations from the premenstrual model were due to factors sub-dividing into a greater number of factors, although some items also moved between factors. Although there were these differences in factor structures, certain symptoms (e.g. anxiety, nervous tension, control; aggression, anger, irritability; breast tenderness, swelling) consistently clustered together.

As for the 17-item models, the solutions produced by the scree plot were very uncertain (see Table 4.2). In the four that were clear, three produced a two factor (table 4.7), and one a four factor solution (table 4.8).

4.8 Discussion and the chosen solution
A series of PCAs were performed to assess how items on the DSR grouped together in a sample similar to that used by Freeman et al. (1996), and to assess the stability of these factors during different cycle phases, in different diagnostic groups and when items were presented to participants in alternative orders. These analyses were also conducted to assess the effect of the addition of the items ‘anger,’ ‘aggression,’ and ‘impulsiveness’ on the factor structure. As there are a variety of statistical techniques available to extract factors (Cattell, 1966; Horn, 1965; Kaiser, 1960), and these can produce different results (Hayton et al., 2004; Zwick and Velicer, 1986), the factor structures suggested by PA, the K1 rule and the scree plot were explored.

The most direct comparison with Freeman et al.’s factor structure is the factor solution produced using the K1 rule for the premenstrual model, as this is the closest analysis to that performed by Freeman et al., with respect to the population and analysis employed. As Table 4.4 indicates, this solution produced three factors from the current data set, in comparison to Freeman et al.’s forced four (see Table 4.1). Factor one closely resembled Freeman’s ‘mood’ factor, with the addition of the item ‘crying.’ However, there were few other similarities. The items that made up Freeman et al.’s ‘physical,’ ‘pain,’ and ‘behavioural’ factors were divided between the second and third factors, with the items grouped together very differently. The purpose of PCA is to summarize a data set (Field, 2005; Zwick and Velicer, 1986). Therefore, these differences in factor structure may simply
arise from different women contributing to the analyses. In the current research, each woman was included several times, since she completed up to six premenstrual DSR scales in each of the three screening cycles (see section 3.3). The differences in factor structure could also arise from differences in the samples recruited. For example, Freeman et al. performed their analysis on an American sample in 1996, while here the sample was a UK sample recruited over a decade later. In addition, although Freeman et al. also excluded women who suffered from anxiety and depression, and only included women who self-diagnosed themselves to suffer from PMS, who had their PMS diagnosis prospectively confirmed, they did not specify how this was done. The methods employed may or may not have been as stringent as the criteria used in this research (see section 3.5.2.1). However, these differences could also have arisen from the addition of items to the scale, which may have caused women to respond to the original symptoms in a different way.

The K1 rule produced diverse factor solutions across diary orders and cycle phases. Although a few items tended to cluster together consistently, solutions differed in the number of factors, item factor loadings and groupings of the other items on factors. The K1 rule also yielded dissimilar factor solutions for the women who did not belong to the PMS diagnostic group. These models remained similar when the items ‘anger,’ ‘aggression’ and ‘impulsiveness’ were added to the scale. Therefore, diverse factor solutions were produced for the 20-item scale across diary orders, cycle phases and diagnostic groups.

The solutions produced by the K1 rule should be given little credence. This rule has been shown to be notoriously inaccurate (Fabrigar et al., 1999; Gorsuch, 1997; Hayton et al., 2004), and some recommend that it should not be used at all (Zwick and Velicer, 1986). Advocates of the technique suggest that it is only accurate when the number of variables is less than 30 and the resulting communalities (after extraction) are all greater than 0.7, or when the sample size is greater than 250 and the average communality greater than 0.6 (Field, 2005). Although the majority of samples were greater than 250, the average communality often fell just below 0.6. Therefore, even if this were a sound technique, the solutions provided would be questionable.

The scree plot did not suggest a clear factor solution for half of the models that were assessed. These problems of clarity arose due to there being a gradual decline in
eigenvalues, with no clear ‘elbow.’ In some cases, there was more than one break in the line, or more than one line that could be drawn through the lower eigenvalues. These are all familiar problems associated with the scree plot (Hayton et al., 2004; Zwick and Velicer, 1986). In the majority of models where a clear solution was apparent, the solutions generated were identical to those produced when PA was used.

PA has been shown to be the most consistently and frequently accurate method of factor retention (Hayton et al., 2004; Humphreys and Montanelli, 1975; Zwick and Velicer, 1986). The models produced using this method appeared more stable than those generated through the use of the K1 rule. A clear two-factor solution was produced for the premenstrual model for the 17- and 20- item scales, with the items being clearly divided into ‘psychological’ and ‘physical’ symptoms. Few differences were produced when items were presented in alternative orders. Slight differences became apparent for other cycle phases. Although a clear two factor model was produced for the 17-item scale in all phases, a couple of items did move between factors. Differences became slightly more apparent for the 20-item scale, where three factors were produced for the ‘follicular’ and ‘rest’ phases, with these comprising slightly different symptoms. These variations were expected, as PMS is a cyclical condition, where symptoms are experienced for approximately 7-10 days before the onset of menstruation, and only disappear a few days into menses (Altshuler et al., 2001; Bäckström et al., 1983; Endicott et al., 1986; Milewicz and Jedrzejuk, 2006; Reid, 1991; Yonkers et al., 2008). Therefore, the factor structure would be expected to be stable during the premenstrual and menstrual cycle phases, while stability of factor structure would not necessarily be expected during the more asymptomatic ‘follicular’ and ‘rest’ phases. However, researchers often use PMS measures to assess symptoms across the menstrual cycle. Moreover, follicular to luteal change scores are often used in PMS diagnosis (NIMH, 1983) and the assessment of treatment efficacy (e.g. Collins et al., 1993; De Souza et al., 2000; Doll et al., 1989; Freeman et al., 1990). Therefore, researchers should be cautious when they assess symptoms through use of component scores during these cycle phases.

The two factor structure became slightly unstable when women other than those who self-diagnosed themselves to suffer from PMS and who met PMS criteria were considered. Although many similarities to the premenstrual model were present, items showed the
tendency to move between factors, and in the case of the ‘controls,’ one factor sub-divided. These differences may be explained by the fact that the scale was developed to measure premenstrual symptoms in women meeting PMS criteria. Therefore, it is understandable that the factor structure may vary in women not suffering from PMS. Nonetheless, researchers use such measures to compare symptom reports from women meeting PMS criteria with other groups of women.

The results presented in this chapter clearly demonstrate that Freeman et al.’s (1996) factor structure was not replicated in the sample recruited for this research. In order to be able to draw comparisons from the analyses presented in this thesis with those from previous studies, the analyses in the following chapters were conducted using Freeman et al.’s factor structure. However, as this factor structure is not representative of the way in which the women recruited for this research experience premenstrual symptoms, the analyses were also conducted using the components derived in this chapter. As PCAs should be conducted on a similar population to the one on which the scale is to be used (Fabrigar et al., 1999; Gorsuch, 1997), the premenstrual factor solutions produced from PA for the 17- and 20-item scales were chosen for the further research presented in this thesis (see Tables 4.3 and 4.7). A central aim of this thesis was to explore premenstrual symptoms in women meeting PMS criteria, in comparison to other groups of women. The premenstrual model was produced from data obtained during the premenstrual phase in women meeting PMS criteria, across diary orders. Furthermore, given that there was very little effect of item order, the premenstrual model was chosen, rather than a model based on a specific order, as this model was based on a much larger sample size of daily diaries, making the solution more reliable. This 17-item structure differed in the number of factors to that produced by Freeman et al. However, the psychological factor was almost identical to Freeman et al.’s mood factor. Moreover, if Freeman et al.’s ‘behavioural,’ ‘pain,’ and ‘physical’ factors were merged into one, this would correspond very closely to the physical factor identified in the current analysis. This discrepancy could be explained by the fact that Freeman et al. used the K1 rule for factor extraction, which has consistently been found to overestimate the numbers of factors that should be retained (Fabrigar et al., 1999; McWilliams et al., 2001; Zwick and Velicer, 1986).
Therefore, as the data presented subsequently in this thesis will be examined using the factor solutions produced from the premenstrual 17- and 20-item models, the reliability of these scales was examined. Cronbach’s alpha is the most common measure of scale reliability (Field, 2005) and should be calculated separately for each subscale (Cronbach, 1951). Values should be in the region of 0.7-0.8 (Field, 2005; Kline, 1999). As the inclusion of repeated daily diary completions of the same women would probably inflate the reliability of the scale, each woman was included once only by taking the average of each of her diary completions in that cycle phase. As can be seen in Table 4.9, Cronbach’s alpha for each subscale was excellent for both scales, ranging from 0.90-0.94.

Table 4.9 Final models with factor loadings and Cronbach’s α coefficients (N=66)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Label</th>
<th>17-item scale items</th>
<th>Factor loadings</th>
<th>20-item scale items</th>
<th>Factor loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Psychological Depression</td>
<td></td>
<td>0.77</td>
<td>Anger</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>Feeling out of control</td>
<td></td>
<td>0.77</td>
<td>Aggression</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>Mood swing</td>
<td></td>
<td>0.76</td>
<td>Mood swing</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>Anxiety</td>
<td></td>
<td>0.73</td>
<td>Irritability</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Irritability</td>
<td></td>
<td>0.71</td>
<td>Depression</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Crying</td>
<td></td>
<td>0.68</td>
<td>Feeling out of control</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Nervous tension</td>
<td></td>
<td>0.67</td>
<td>Anxiety</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>(α = 0.92)</td>
<td></td>
<td></td>
<td>Crying</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Nervous tension</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Impulsiveness</td>
<td>0.44</td>
</tr>
<tr>
<td>B</td>
<td>Physical Swelling</td>
<td></td>
<td>0.76</td>
<td>Swelling</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Cramps</td>
<td></td>
<td>0.62</td>
<td>Poor coordination</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>Poor coordination</td>
<td></td>
<td>0.61</td>
<td>Fatigue</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>Breast tenderness</td>
<td></td>
<td>0.61</td>
<td>Confusion</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>Aches</td>
<td></td>
<td>0.60</td>
<td>Cramps</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>Headache</td>
<td></td>
<td>0.58</td>
<td>Headache</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>Fatigue</td>
<td></td>
<td>0.57</td>
<td>Aches</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>Confusion</td>
<td></td>
<td>0.54</td>
<td>Breast tenderness</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>Insomnia</td>
<td></td>
<td>0.52</td>
<td>Insomnia</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>Food cravings</td>
<td></td>
<td>0.44</td>
<td>Food cravings</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>(α = 0.90)</td>
<td></td>
<td></td>
<td>(α = 0.90)</td>
<td></td>
</tr>
</tbody>
</table>

For a reliable scale, the deletion of any one item should not increase the overall reliability, by increasing the overall alpha (Field, 2005). The overall alpha was not increased by the deletion of any item on the ‘psychological’ factor of the 17-item or 20-item scale, or on the
\textit{physical} factor, which is identical for each scale. This suggests that the inclusion of all items improves the reliability of the scale.

### 4.8.1 Factor Scores

Factor scores are estimates of the score participants would have received on each factor had the factor been measured directly (Tabachnick and Fidell, 2007a). Factor scores are extremely useful for many forms of analysis, including multiple regression and multivariate analysis of variance, as there are usually fewer factors than observed variables (Tabachnick and Fidell, 2007a) and because factor scores are almost completely uncorrelated if factors are orthogonal (Field, 2005; Tabachnick and Fidell, 2007a).

Factor scores can be produced through a variety of methods (Field, 2005; Grice, 2001; Tabachnick and Fidell, 2007a). Simple methods, such as averaging or summing the scores on the items that load highly on each factor can be used (Grice, 2001). However, this approach does not take into account the relevant importance of the items for the factor (Bogner and Wiseman, 2002). More sophisticated methods, which take this into account are known as ‘exact methods’ (Field, 2005; Tabachnick and Fidell, 2007a). The ‘regression method’ calculates factor scores by multiplying standardised scores on items with their factor score coefficients (Tabachnick and Fidell, 2007a). By doing so, this method produces factor scores that represent the relationship between the items and factors, taking into account the initial relationship between items (Field, 2005; Grice, 2001). If the factor scores are used in subsequent analyses, the regression method carries more information originally contained in the factors to those new analyses than a ‘simple method’ would (Grice and Harris, 1998). This method is commonly used by statisticians to demonstrate how factor scores should be produced (Field, 2005; Tabachnick and Fidell, 2007a) and has been widely used in research settings (e.g. Bogner and Wiseman, 2002; Hart et al., 1988). This is also the default setting on many statistical packages, including SPSS, which calculates factor scores by firstly multiplying the inverse correlation matrix with the factor loadings to produce the factor score coefficients, and then by multiplying the factor score coefficients with the standardized variables (Z scores).

Therefore, the regression method was used to produce factor scores for the DSR-17 and DSR-20. The factor score coefficients for the premenstrual models of the DSR-17 and
DSR-20 were multiplied with participants’ scores on the variables. Z scores were not needed, as the factors were all measured on the same scale (Grice and Harris, 1998; Tabachnick and Fidell, 2007a).

4.9 Summary
A series of PCAs were conducted to investigate the factor structure of the DSR based on the data collected for this thesis. Various techniques to assess how many factors should be retained were compared. The majority of solutions suggested by the scree plot lacked clarity, whilst those suggested by the K1 rule were diverse. PA produced a stable two factor solution in the premenstrual phase, which divided items into ‘psychological’ and ‘physical’ categories when only the original 17 items were considered. This structure remained when the additional items were included. These factor solutions remained stable when items were presented to participants in alternative orders, and only varied slightly during other menstrual cycle phases. The reliability of these scales was demonstrated through the examination of the Cronbach’s alpha coefficients. Based on these analyses, subsequent data analysis presented in this thesis will examine the two factor 17- and 20- item DSR scales, through the use of factor scores produced using the regression method. These scales will be referred to as the DSR-17 and DSR-20 respectively. The data will also be assessed according to Freeman et al.’s (1996) model using the factor totals for her four factor solution, so that comparisons can be made with the existing literature. This original scale will be referred to as the DSR.
Chapter 5: Characterising PMS

5.1 Introduction
This chapter aimed to improve the characterisation of the Premenstrual Syndrome by exploring the psychological (impulsivity and aggression) and physiological (serotonin, hormone and proinflammatory cytokine) profiles of women who self-diagnosed with PMS and who met PMS criteria, with the psychological and physiological profiles of other groups of women.

5.1.1 Psychological profiles
Women suffering from PMS often report that their most distressing symptoms include aggression (Endicott et al., 1999; Hylan et al., 1999; Landen and Eriksson, 2003; Lurie and Borenstein, 1990; Warner and Bancroft, 1990) and impulsivity (Elliott, 2002; Halbreich, 2003; Halbreich et al., 2007; Hallman et al., 1987; Hsu et al., 2007; Matsumato et al., 2007). Moreover, women maintain that these symptoms premenstrually are the symptoms which motivate them to seek treatment (Bancroft et al., 1993; Hartlage and Arduino, 2002). However, the variability of aggressive and impulsive symptoms across the menstrual cycle has been the subject of little research. The limited research that has been conducted was reviewed in Section 1.4.2.

The study presented in this chapter firstly aimed to determine whether women suffering from PMS demonstrate variable profiles of aggressive and impulsive symptoms across the menstrual cycle. This was addressed by assessing the aggression and impulsivity levels reported by PMS sufferers across four cycle phases (see section 3.3) of three menstrual cycles, through the administration of the BPAQ (Buss and Perry, 1992) and the BIS-11 (Patton et al., 1995). Women were defined as PMS sufferers in this research if they self-diagnosed with PMS, met the 30% increase criterion on the DSR (Freeman et al., 1996), did not exceed a follicular anxiety score of 51 (STAI; Spielberger et al., 1983b) and a follicular depression score of 12 (BDI; Beck et al., 1961) (see sections 3.5.1 and 3.5.2).

Furthermore, this study aimed to determine whether cyclical mood changes represent an abnormal symptom profile, or are variations of a normal cyclical pattern. This was addressed by comparing the aggressive and impulsive symptom profiles reported by PMS
sufferers with those reported by normally cycling women, who did not consider themselves to experience problematic PMS symptoms (see section 3.5.3). The 30% increase criterion was used as the diagnostic method to confirm PMS in this research (see section 3.5.2.4). The application of this criterion appeared to divide women experiencing ‘normal’ cyclicity into two groups, driven by their follicular symptom severity scores (controls and PMS non-reporters) (see section 3.5.2.3). In line with previous research (e.g. Gallant et al., 1992a; Morse et al., 1988), approximately one third of women who did not self-diagnose with PMS did not meet a PMS diagnosis when this criterion was used (controls). However, approximately two thirds of women who did not self-diagnose with PMS did meet the 30% increase criterion (PMS non-reporters) (see section 3.5.2.1). As discussed in Section 3.5.2.3, women belonging to the PMS non-reporter group were considered to be ‘false positives,’ in the sense that the use of the 30% increase criterion resulted in a positive PMS diagnosis in this group of women who did not self-report problematic premenstrual symptoms, and whose actual premenstrual symptom ratings were almost identical to the control group. As such, women belonging to the PMS non-reporter group were considered to experience ‘normal cyclicity’ (Gallant et al., 1992a; see also section 3.5.2.3). Therefore, the control and PMS non-reporter groups were combined into one group of normally cycling women, to provide a comparison of the impulsivity and aggression profiles reported by PMS sufferers with those reported by women experiencing ‘normal cyclicity’.

Some researchers compare the symptom profiles reported by PMS sufferers with those reported by a control group comprising both controls and PMS non-reporters, in order to determine whether cyclical mood changes represent an abnormal symptom profile, or are variations of a ‘normal’ cyclical pattern (e.g. Hallman et al., 1987). However, other researchers address this question by comparing the symptom profiles reported by PMS sufferers, with those reported by control women, recruited on the basis that they do not consider themselves to have PMS, and then by selecting from this group those women who do not meet PMS criteria (e.g. Bond et al., 2003; De Ronchi et al., 2005; Howard et al., 1994). This second strategy could result in the exclusion of a large proportion of women who do not consider themselves to experience problematic PMS symptoms (PMS non-reporters) from the control group in studies that have used liberal PMS diagnostic criteria (Gallant et al., 1992a; Morse et al., 1988; see also section 3.5.2.1). Therefore, the aggression and impulsivity profiles reported by PMS sufferers were compared with those
reported by the *PMS non-reporters* and *controls* as separate groups, to assess the impact of excluding PMS non-reporters from control groups. Furthermore, some researchers disguise the purpose of their research and allocate women to groups after they have completed symptom reports to reduce the likelihood of demand characteristics influencing reporting behaviour (e.g. Christensen and Oei, 1988; Dougherty et al., 1998; Keenan et al., 1992). Since many women who do not report distressing PMS symptoms may meet liberal PMS diagnostic criteria (PMS non-reporters) (e.g. Gallant et al., 1992a; Morse et al., 1988; see also section 3.5.2.1), retrospective designation of women as PMS sufferers could result in the inclusion of women who do not endorse the view that they have PMS in PMS groups in research settings. This issue is discussed in relation to the aggression and impulsivity profiles that these groups reported in the research presented in this chapter.

In order to confirm a PMS diagnosis, PMS symptoms should be limited to the luteal phase (Freeman, 2003; Frye and Silverman, 2000; Halbreich et al., 2007; Lampe, 2005; see also section 1.3.1.4). Therefore, researchers should screen for co-morbidities. Although some researchers identify and exclude women with co-morbidities from their PMS groups (e.g. Bond et al., 2003; De Ronchi et al., 2005; Dougherty et al., 1998; Haskett et al., 1984; Howard et al., 1994), others do not consider whether women are suffering from co-morbid conditions (e.g. Howard et al., 1988; Van der Ploeg, 1987; Watts et al., 1980). This can result in women with co-morbidities being considered PMS sufferers. In order to assess the impact of including women with co-morbidities in PMS groups in research settings, the aggression and impulsivity profiles reported by *PMS sufferers* were compared with those reported by women with co-morbidities. Women were defined as *co-morbid* in this research if they met criteria for anxiety and/or depression (see section 3.5.1), the most common co-morbid conditions, in the follicular phase, regardless of whether or not they met the 30% increase criterion on the DSR.

### 5.1.2 Physiological profiles

The aetiology of PMS is not clearly understood (see section 1.5; for a review see Halbreich, 2003). However, it has been suggested that women with PMS appear more sensitive to normal cyclical fluctuations in steroid hormones, which influence neurotransmitter function in the central nervous system (Eriksson et al., 2002; Rapkin, 1992; Reid and Yen, 1981; Steiner et al., 1997a). Increasing evidence suggests that serotonin plays an important role
(Eriksson et al., 2002; Halbreich, 1997; Rapkin, 1992; Steiner, 1997a; Yonkers et al., 2008). The serotonergic system interacts with other physiological mechanisms (Steiner et al., 2003b), including testosterone production (Eriksson et al., 1994; Virkkunen et al., 1994) and the cytokine network (Gemma et al., 1997; 2003; Hardin-Pouzet et al., 1996; Linthorst et al., 1994; Pousset et al., 1996; Silverman et al., 1989), leading to the proposal that PMS symptoms result from fluctuations in the activity of these systems (Konecna et al., 2000; Sunblad et al., 1994). However, no previous research has prospectively examined proinflammatory cytokine levels in PMS sufferers, while research assessing testosterone levels is limited. Moreover, the relationship between the serotonergic system and proinflammatory cytokine and testosterone levels has not been causally elucidated. Therefore, the study presented in this chapter aimed to determine whether serotonin, steroid hormone and proinflammatory cytokine levels are associated with PMS symptoms. This was addressed by comparing the physiological profiles (serotonin, steroid hormone and proinflammatory cytokine) of PMS sufferers with the physiological profiles of normally cycling women (controls and PMS non-reporters). The physiological profiles of PMS non-reporters were then compared with those of the PMS sufferer and control groups to investigate whether these comparisons could provide objective evidence that the PMS non-reporters should be considered to experience ‘normal’ rather than ‘abnormal’ symptom profiles (see section 3.5.2.3).

5.2 Method

5.2.1 Participants
The sample comprised a total of 142 women who completed the three screening cycles (see sections 3.4.3). Of these 66 were PMS sufferers, 48 normally cycling women (32 PMS non-reporters; 16 controls) and 28 co-morbid. Follicular and luteal blood samples, collected within the targeted days, were obtained from 21 PMS sufferers and 42 normally cycling women (29 PMS non-reporters and 13 controls).

5.2.2 Measures

5.2.2.1 Self-report measures
Participants were asked to complete the BPAQ (Buss and Perry, 1992) and the BIS-11 (Patton et al., 1995) at the end of each week throughout the three screening cycles, by reflecting on the previous week. For the purposes of this research, the menstrual cycle was
divided into four cycle phases; bleeding, follicular, rest and luteal (see section 3.3). Questionnaire completions from cycle day seven (representing cycle days 1-7) signified the bleeding phase, from cycle day fourteen (days 8-14) the follicular phase and from cycle day one of the next cycle the luteal phase (days -6 to -1). The rest phase varied in length between and within women (see section 3.3). Therefore, if two sets of questionnaires had been completed between the follicular and luteal phases, the completions from the week following the follicular phase were taken to represent the rest phase. If three sets of questionnaires had been completed, the middle set was used for this purpose (see section 3.3).

5.2.2.2 Biological measures
5.2.2.2.1 Gonadal steroids
Two blood samples (each 18mls) were taken during the course of the screening cycles to assess follicular and luteal hormone (FSH, LH, oestradiol, progesterone, prolactin and testosterone), serotonin (5-HT and 5-HIAA) and cytokine (IL-1β, IL-6, IL-8, IFN-γ and TNF-α) levels. Samples were collected as described in Section 3.4.4.3.

The presence and quantity of FSH, LH, oestradiol, progesterone, prolactin and testosterone were detected in vitro from each sample simultaneously using the automated biochip array analyser, Evidence (e.g. FitzGerald et al., 2005; Molloy et al., 2005). This method for the simultaneous analysis of these hormones was chosen for its sensitivity and cost effectiveness.

Each plasma sample was analysed using a Randox biochip, by pipetting the plasma onto specific test regions containing antibodies and antigen complexes specific to the different hormones under analysis. Levels of prolactin, FSH and LH were detected through a sandwich chemiluminescent immunoassay, by assessing the extent of binding of an antibody labelled with horse-radish peroxidase (HRP). When the analyte binds with the antibody labelled with HRP, the HRP fluoresces. The amount of chemiluminescence emitted was assessed through the use of digital imaging technology. The quantity of each hormone was identified by comparing the light signals generated from specific areas of the biochip with those from a calibration curve. As increased levels of prolactin, FSH and LH
lead to increased binding of the antibody labelled with HRP, the amount of chemiluminescence emitted reflected the concentration of the hormone.

Levels of testosterone, progesterone and oestradiol were detected through a competitive chemiluminescent immunoassay. Here, the analyte and a secondary antibody labelled with HRP compete to bind with the primary antibody. If there is more of the analyte present, then less of the secondary antibody labelled with HRP can bind with the primary antibody. Therefore, there is an inverse relationship between the amount of the analyte present and the chemiluminescence emitted. Lower emission of chemiluminescence reflects greater levels of these hormones.

5.2.2.2.2 Cytokine analysis
The presence and concentration of IL-1β, IL-6, IL-8, IFN-γ and TNF-α were detected simultaneously through the use of the LINCOplex multiplex assay kit (e.g. Hildesheim, 2002; Kaplan et al., 2007). This method is a sensitive and cost effective technique (Hildesheim, 2002).

Samples were analysed using plates containing 96 wells. One well was used for each sample under analysis. Each well contained antibodies highly specific to each cytokine, paired with a bead set internally stained with a mixture of fluorochromes, giving the bead set a unique colour code. The plasma samples, which were previously diluted with bead dilutant (1:4), were then mixed with these heterogeneous bead sets in the well, which allowed the analytes to be captured by the antibodies. After washing, a secondary antibody labelled with biotin (a reporter dye) was then added. The beads were then analysed in a Luminex analyser, where the beads passed singly through two lasers. One laser was used to excite the internal dyes of the bead to identify the analyte being measured, while another identified the reporter signal, which determined the strength of the interaction, thus quantifying the analyte.

5.2.2.2.3 Serotonin
The plasma samples were taken with the intention of measuring 5-HT and 5-HIAA levels. These samples were taken between December 2005 and June 2007 and were stored at -80° centigrade until the end of the study. Unfortunately the protocol supplied by the Perinatal
Research Group at St James’s University Hospital did not consider the need to carefully control food intake prior to sampling, or the need to analyse the samples within three weeks of them being taken. This rendered the samples unsuitable for the analysis of 5-HIAA and 5-HT.

5.2.3. Procedure
Women were asked to complete one set of questionnaires each week throughout the three screening cycles. This was done through the administration of the daily diary booklets (see section 3.1.1), which comprised one copy of the BPAQ (Buss and Perry, 1992) and the BIS-11 (Patton et al., 1995), which women were required to complete at the end of each phase by reflecting on the week during which they had completed daily ratings. PMS sufferers and normally cycling women (PMS non-reporters and controls) were also asked to provide blood samples during the follicular (cycle days 5-10) and luteal (days -6 to -1) phases of the third screening cycle. See Section 3.4 for the detailed method of the screening cycles.

5.3 Data Processing and Analysis
5.3.1 Self-report measures
Comparisons of the aggression and impulsivity profiles reported by the PMS sufferer, normally cycling (controls, PMS non-reporters), and co-morbid groups were undertaken using profile analysis. Each symptom was assessed separately, as there was no theoretical reason to combine them (Tabachnick and Fidell, 2007a). Repeated measures univariate analysis of variance (RM ANOVA) was used for the analysis of total scale scores, as this was the only DV being considered. Doubly multivariate analysis of variance was used for the simultaneous assessment of subscales of each measure in order to avoid inflation of familywise Type I error rates, which occurs when separate analyses are performed on each subscale (Bray, 1995; Field, 2005; Tabachnick and Fidell, 2007a). Pillai’s Trace was chosen as the multivariate test statistic, as this has most power in studies with more than one DV (Field, 2005), and is widely advocated for general use (Haase and Ellis, 1987; Olson, 1976; Tabachnick and Fidell, 2007a).

Simple effects and interaction contrast analyses were conducted to explore the results from the RM ANOVAs and doubly MANOVAs, in accordance with Tabachnick and Fidell’s (2007a) recommendation. Bonferroni corrections were applied to the α levels to control for
inflated familywise Type I error (Tabachnick and Fidell, 2007b). This was done for each analysis by dividing the $\alpha$ level by the number of tests that could have been conducted i.e. the most conservative approach. As PMS sufferers demonstrate their lowest symptom levels during the follicular phase, and their greatest levels during the luteal phase, interaction contrast analyses were conducted to compare the follicular to luteal change scores between groups in the presence of a statistically significant phase by group interaction.

5.3.2 Biological measures

Comparisons of the steroid hormone and cytokine concentrations of the PMS sufferers and normally cycling women (controls and PMS non-reporters) during the follicular and luteal phases were made using RM ANOVAs. Each steroid hormone was assessed separately as it was not expected that all hormones would change in the same way. These hormones vary quite differently across the cycle (see section 1.1). Therefore, combining the steroid hormone data in a multivariate analysis could result in the obscuring of specific hormone effects. The cytokine data was also analysed univariately, given the exploratory nature of the study (see section 1.5.3.2.3).

5.3.3 Data Screening

5.3.3.1 Self-report measures

Prior to the analysis described above, data were screened for missing data, outliers and normality.

5.3.3.1.1 Missing data

Only data from women who provided a complete set of weekly measures (BPAQ and BIS-11) throughout the screening cycles were included in the analyses presented in this chapter. Missing data was fairly consistent across the two measures, as if data was missing, it was usually missing for both measures (see Table 5.1).

<table>
<thead>
<tr>
<th>Table 5.1 Numbers of women on which the analyses were conducted</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMS</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>BPAQ</td>
</tr>
<tr>
<td>BIS-11</td>
</tr>
</tbody>
</table>
5.3.3.1.2 Sample sizes

Although the diagnostic groups consisted of different numbers of women (see Table 5.1) this is not of concern, as each hypothesis in profile analysis is tested as in a one-way design, and only creates a difficulty when there is more than one between subjects DV (Tabachnick and Fidell, 2007a).

5.3.3.1.3 Outliers

Outlying scores were not excluded from the analysis since the main concern was to examine how participant’s scores changed across time and in order to maintain the number of valid observations. Therefore, participants’ scores were retained, even if they were outliers (greater or less than 2sds from the mean) relative to the rest of the sample.

5.3.3.1.4 Normality

To check whether the data were normally distributed, both Z scores for skewness and the Kolmorgorov-Smirnov (K-S) test were consulted for each group, for each DV (factor), at each time-point. Where Z-scores for skewness were found to be greater than 2.58 (Field, 2005) and the K-S was significant, the data were considered to significantly deviate from a normal distribution. In repeated measures designs, transformation of one factor makes it necessary to transform all factors making up that measure. Transformations were not considered necessary in the study presented in this chapter, since the majority of BPAQ and BIS-11 DVs did not deviate from a normal distribution. Although the physical subscale of the BPAQ was found to be positively skewed for the PMS group at certain time points, the BPAQ DVs were not transformed, as the physical subscale of the BPAQ for the other groups, and the majority of the other BPAQ factors for all groups, did not deviate significantly from a normal distribution.

5.3.3.1.5 Sphericity

Sphericity was not a concern for the current study, since profile analysis was conducted, and this circumvents the need for the sphericity (Tabachnick and Fidell, 2007a). However, when simple effects and interaction contrast analyses were conducted, the Greenhouse Geisser’s correction was applied when Mauchly’s (1940) test indicated that sphericity was not met.
5.3.3.2 Biological measures

5.3.3.2.1 Missing data and protocol violations

Follicular data collection depended on prompt notification from the participants that they had begun their period. If participants did not contact the PI around the time that their period was expected, every effort was made by the PI to contact them. Occasionally this was not possible which resulted in some data being collected outside of targeted days. Missing data mostly arose in luteal phase data collection from women participating in the clinical trial (PMS sufferers), due to earlier or later than expected onset of menses. The visits were arranged based on the prediction of participants’ cycle length. Although only PMS sufferers who reported regular menstrual cycles were entered onto the study, there were many occasions when women did not begin their period when expected. If participants began their period early, this resulted in data being collected in the bleeding phase of their following cycle. Conversely, if participants began their period later than expected, data was collected before participants entered their luteal phase. The collection of data during targeted days was also compromised by the clinical trial spanning Christmas and summer holidays. Data collection began in November 2005 and ended in June 2007. Although great lengths were taken to avoid missing data, four of the PMS sufferers who completed the study each missed one visit. Very little missing data arose in luteal phase data collection from the group of normally cycling women (controls and PMS non-reporters), as these women were asked to continually return for additional blood sampling if their period did not begin within the required time-frame (see section 3.4.5).

Missing biological data also resulted from unsuccessful blood sampling e.g. vein collapsed during sampling. After a sample had been missed, many women returned later in the day for a second attempt. However, some samples were still unsuccessful.

5.4 Results

5.4.1 Psychological symptom profiles for the three diagnostic groups

The psychological (aggression and impulsivity) symptom profiles that were reported by the PMS sufferers were firstly compared with those that were reported by the normally cycling (control and PMS non-reporter) and co-morbid groups.
5.4.1.1 Aggression
Participants rated their aggression from “extremely characteristic” to “extremely uncharacteristic” on a weekly basis (see section 3.1.1). The scale was scored in such a way that a low score reflected low aggression.

5.4.1.1.1 BPAQ total scale scores
A 3 x 4 x 3 RM ANOVA was conducted on the BPAQ-total scores. The three diagnostic groups formed the between subjects factor; PMS sufferer, normally cycling (controls and PMS non-reporters) and co-morbid groups, formed on the basis of their symptom report profiles described in Section 3.5.2.1. The four cycle phases and the three screening cycles formed the within subjects IVs. There was a significant main effect of cycle (F(2, 216=3.50, p<0.05, partial \(\eta^2=0.031\)) but no significant cycle by group interaction (F(4, 216) = 1.93, p>0.05). There was a significant main effect of phase (F(3, 324) = 36.45, p<0.001, partial \(\eta^2 =0.25\)) and a significant main effect of group (F(2, 108) = 6.61, p<0.01, partial \(\eta^2 =0.11\)). However, the significant phase by group interaction (F(6, 324) = 2.97, p<0.05, partial \(\eta^2 =0.052\)) qualifies these main effects (Tabachnick and Fidell, 2007a), and suggests that the pattern of aggression reported over the cycle differed between groups. Figure 5.1 displays the BPAQ-total profiles of the three diagnostic groups over the averaged screening cycles (i.e. mean of screening cycles 1 to 3).
A simple effects analysis was performed to compare the groups in the follicular and luteal phases. A significant main effect of group was found during each phase ($F(2, 109) = 11.49$, $p<0.0125$, partial $\eta^2 = 0.17$ and $F(2, 109) = 6.07$, $p<0.0125$, partial $\eta^2 = 0.10$ respectively). Bonferroni pairwise comparisons revealed that the co-morbids reported significantly greater BPAQ-total scores than the PMS sufferers and normally cycling women in the follicular phase, while the PMS and co-morbid groups reported significantly greater levels than the normally cycling women in the luteal phase ($p<0.05$).

A simple effects analysis was performed to examine the mean differences in BPAQ-total reporting over the four cycle phases for each group. A significant main effect of cycle phase was found for the PMS sufferers and normally cycling women ($F(3, 150) = 40.29$, $p<0.0125$, partial $\eta^2 = 0.45$ and $F(3, 114) = 9.45$, $p<0.0125$, partial $\eta^2 = 0.20$ respectively), indicating that these groups reported variable patterns of aggression across the cycle. Whilst the co-morbids displayed this trend, differences between cycle phases just failed to reach statistical significance when familywise Type I error rate was controlled for ($F(3, 63) = 4.87$, $p=0.015$, partial $\eta^2 = 0.19$).
The BPAQ-total profiles displayed in Figure 5.1 suggest that the PMS group show a greater follicular to luteal increase in aggression than the normally cycling and co-morbid groups. This was confirmed through an interaction contrasts analysis, which compared the follicular to luteal change scores between groups. This indicated that the groups showed significantly different increases in aggression between these phases \( F(2, 109) = 5.52, p<0.0125, \text{partial } \eta^2=0.092 \). Bonferroni pairwise comparisons revealed that the PMS sufferers showed a significantly greater increase in BPAQ reporting between these phases than the normally cycling women \( (p=0.009) \), and showed a trend towards reporting a greater increase than the co-morbids \( (p=0.063) \).

5.4.1.1.2 The BPAQ subscales

A doubly multivariate analysis of variance was performed on the verbal, physical, hostility and anger BPAQ subscales over the three screening cycles and four cycle phases, with the three diagnostic groups as the between subjects factor. There was no significant main effect of cycle (multivariate \( F(8, 101) = 1.29, p>0.05 \)) and no significant cycle by group interaction (multivariate \( F(16, 204) = 0.76, p>0.05 \)), indicating that all groups were consistent in their reporting of aggression over the three screening cycles for the linear combination of the BPAQ subscales. There was a significant main effect of phase (multivariate \( F(12, 97) = 9.19, p<0.001, \text{partial } \eta^2= 0.53 \)) and significant main effect of group (multivariate \( F(8, 212) = 6.03, p<0.001, \text{partial } \eta^2= 0.19 \)). Moreover, the significant phase by group interaction (multivariate \( F(24, 196) = 2.01, p<0.01, \text{partial } \eta^2= 0.20 \)) indicated that the pattern of aggression reported over the cycle differed between groups.

Given that there were some significant multivariate effects, the univariate tests of the main effects of phase and group and the phase by group interaction were examined (Field, 2005; Tabachnick and Fidel, 2007a). These effects are summarized in Table 5.2.
Table 5.2 Multivariate and univariate statistics for the main effects of phase and group, and for the phase by group interaction

<table>
<thead>
<tr>
<th>Phase</th>
<th>Statistic</th>
<th>Partial $\eta^2$</th>
<th>Group</th>
<th>Statistic</th>
<th>Partial $\eta^2$</th>
<th>Phase by Group</th>
<th>Statistic</th>
<th>Partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multivariate</td>
<td>$F(12, 97)=9.19^*$</td>
<td>0.53</td>
<td>$F(8, 212)=6.03^*$</td>
<td>0.19</td>
<td>$F(24, 196)=2.01^*$</td>
<td>0.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Univariate:</td>
<td>F(3, 324)=45.53*</td>
<td>0.30</td>
<td>F(2, 108)=7.72*</td>
<td>0.13</td>
<td>F(6, 324)=4.20*</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Anger</td>
<td>F(3, 324)=31.24*</td>
<td>0.22</td>
<td>F(2, 108)=0.80</td>
<td>0.02</td>
<td>F(6, 324)=3.63*</td>
<td>0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Verbal</td>
<td>F(3, 324)=24.35*</td>
<td>0.18</td>
<td>F(2, 108)=13.75*</td>
<td>0.20</td>
<td>F(6, 324)=2.57†</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Hostility</td>
<td>F(3, 324)=18.74*</td>
<td>0.15</td>
<td>F(2, 108)=1.31</td>
<td>0.02</td>
<td>F(6, 324)=1.06</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NB. * Bonferroni corrected $p<0.0125$  $p<0.05$

Table 5.2 shows that there were significant univariate main effects of phase for all BPAQ subscales. There were significant main effects of group for anger and hostility. The phase by group interactions were significant for the anger and verbal subscales and marginally significant for hostility ($p=0.027$, Bonferroni corrected $\alpha = 0.0125$). Figures 5.2 to 5.5 display the aggression profiles of the three groups over the averaged screening cycles for each BPAQ subscale.
A simple effects analysis of the luteal phase showed that there was a significant main effect of group for the anger subscale (F(2, 109) = 8.28, p<0.0125, partial η²=0.13). Bonferroni pairwise comparisons revealed that the co-morbids and PMS sufferers rated anger...
significantly higher than the normally cycling women. All groups reported a variation of anger symptom reporting across the cycle (smallest $F(3, 63)=6.15$, $p<0.0125$, partial $\eta^2=0.23$). Figure 5.2 suggests that the PMS sufferers displayed the greatest variability in anger reporting. This was confirmed through an interaction contrasts analysis, which showed that the groups displayed significantly different increases in anger between the follicular and luteal phases ($F(2, 109) = 9.41$, $p<0.0125$, partial $\eta^2=0.15$). Bonferroni pairwise comparisons revealed that the PMS sufferers reported a significantly greater increase between these phases than the normally cycling and co-morbid groups ($p<0.05$).

5.4.1.1.2.2 Verbal aggression

A simple effects analysis of the luteal phase showed that there was a trend towards an effect of group for verbal aggression ($F(2, 109) = 2.46$, $p=0.09$, partial $\eta^2=0.043$). As for the anger subscale, all groups displayed variable patterns of verbal aggression across the cycle (smallest $F(3, 63) = 6.17$, $p<0.0125$, partial $\eta^2=0.23$). Figure 5.3 suggests that verbal aggression is most variable for the PMS group. This was confirmed through an interaction contrasts analysis, which showed that the groups reported significantly different increases in verbal aggression between the follicular and luteal phases ($F(2, 109) = 6.69$, $p<0.01$, partial $\eta^2=0.11$). The PMS sufferers showed a significantly greater increase between these phases than the normally cycling women ($p=0.002$), and showed a trend towards reporting a greater increase than the co-morbits ($p=0.055$).

5.4.1.1.2.3 Hostility

Hostility profiles (see Figure 5.4) for the three groups appear similar to those for anger (see Figure 5.2). A simple effects analysis of the luteal phase showed that there was a significant effect of group ($F(2, 109) = 10.25$, $p<0.0125$, partial $\eta^2=0.16$). As for the anger subscale, the PMS and co-morbid groups reported significantly greater hostility than the normally cycling women. The PMS sufferers and normally cycling women reported variable levels of hostility across the cycle (smallest $F(3, 114) = 6.65$, $p<0.0125$, partial $\eta^2=0.15$). Although the co-morbits showed this trend (see Figure 5.4), their mean differences between phases did not reach statistical significance when multiple testing was controlled for ($F(3, 63) = 3.51$, $p=0.037$, partial $\eta^2=0.14$). Similar to anger and verbal aggression, the PMS group displayed the greatest variability in hostility. An interaction contrasts analysis revealed that the groups showed significantly different increases in hostility between the
follicular and luteal phases (F(2, 109) = 4.81, p<0.05, partial η²=0.08). Bonferroni pairwise comparisons revealed that the PMS sufferers showed a significantly greater increase between these phases than the normally cycling women (p<0.05).

5.4.1.1.2.4 Physical aggression
As shown in Figure 5.5, the three groups showed similarity in their physical aggression profiles, in terms of the pattern and severity of symptoms that they reported. This was confirmed by the univariate analyses displayed in Table 5.2. Although a significant main effect of phase was evident (F(3, 321) = 14.08, p<0.0125, partial η²=0.12), there was no significant phase by group interaction, suggesting that the groups reported similar patterns of low physical aggression levels across the cycle, each with a slight luteal increase.

5.4.1.2 Impulsivity (BIS-11)
5.4.1.2.1 BIS 11 total scale scores
A 3 (cycle) x 4 (phase) x 3 (group) RM ANOVA was conducted on the BIS-11 total scores. There was no significant main effect of cycle (F(2, 224) = 0.49, p>0.05) and no significant cycle by group interaction (F(4, 224) = 2.15, p>0.05), indicating that all groups were consistent in their reporting of impulsivity over the three screening cycles. There was a significant main effect of phase (F(3, 336) = 31.9.3, p<0.001, partial η²= 0.22) and significant main effect of group (F(2, 112) = 7.35, p<0.01, partial η²=0.12). Moreover, the significant phase by group interaction (F(6, 336) = 3.74, p<0.01, partial η²= 0.06) indicated that the pattern of impulsivity reported over the cycle differed between groups. Figure 5.6 displays the BIS-11 profiles of the three groups over the averaged screening cycles.
Figure 5.6 The BIS-11 profiles of the three diagnostic groups averaged over the three screening cycles

Figure 5.6 suggests that the greatest group separation occurs in the follicular and luteal phases. A simple effects analysis showed that there was a significant main effect of group during each of these phases (F(2, 112) = 13.30, p<0.0125, partial η²= 0.19 and F(2, 112)=4.99, p<0.0125, partial η²=0.08 respectively). Bonferroni pairwise comparisons revealed that co-morbids reported significantly greater impulsivity levels than the PMS sufferers and normally cycling women during the follicular phase, and significantly greater levels than the normally cycling women during the luteal phase (p<0.05). The PMS sufferers showed a trend towards reporting greater levels than the normally cycling women during the luteal phase (p=0.078).

The BIS-11 total profiles displayed in Figure 5.6 suggest that the PMS sufferers and normally cycling women display a curvilinear pattern of impulsivity across the cycle, in comparison to the co-morbids who report more stable levels. This was confirmed through a simple effects analysis of phase differences for each group, which revealed that the PMS sufferers and normally cycling women (F(3, 159) = 40.82, p<0.0125, partial η²=0.44 and F(3, 111) = 9.22, p<0.0125, partial η²=0.20 respectively) displayed variable levels of impulsivity across the cycle, while the co-morbids did not (F(3, 66) = 3.29, p>0.0125, partial η²=0.13 respectively).
The BIS-11 profiles displayed in Figure 5.6 suggest that the PMS sufferers show a greater follicular to luteal increase in impulsivity than the normally cycling and co-morbid groups. This was confirmed through an interaction contrasts analysis, which compared the follicular to luteal change scores between groups. This indicated that the groups demonstrated significantly different increases in impulsivity between these phases (F(2, 112) = 7.08, p<0.0125, partial η²=0.11). Bonferroni pairwise comparisons revealed that the PMS sufferers reported a significantly greater increase in impulsivity between these phases than the normally cycling and co-morbid groups (p<0.05).

5.4.1.2.2 BIS-11 subscales

A 3 (cycle) x 4 (phase) x 3 (group) doubly multivariate analysis of variance was performed on the attentional, motor and non-planning BIS-11 subscales. There was no significant main effect of cycle (multivariate F(6, 108 = 1.15, p>0.05) and no significant cycle by group interaction (multivariate F(12, 218) = 1.59, p>0.05). There was a significant main effect of phase (multivariate F(9, 105) = 10.13 p<0.001, partial η²= 0.47) and significant main effect of group (multivariate F(6, 224) = 9.27, p<0.001, partial η²=0.20). However, these main effects were qualified by the significant phase by group interaction (multivariate F(18, 212) = 1.93, p<0.05, partial η²= 0.14).

Given that there were some significant multivariate effects, the univariate tests of the main effects of phase and group and the phase by group interaction were examined (Field, 2005; Tabachnick and Fidell, 2007a). These effects are summarized in Table 5.3

<table>
<thead>
<tr>
<th>Phase/Group by</th>
<th>Statistic</th>
<th>Partial η²</th>
<th>Statistic</th>
<th>Partial η²</th>
<th>Statistic</th>
<th>Partial η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multivariate</td>
<td>F(9, 105)=10.13*</td>
<td>0.47</td>
<td>F(6, 224)=9.27*</td>
<td>0.20</td>
<td>F(18, 212)=1.93*</td>
<td>0.14</td>
</tr>
<tr>
<td>Univariate:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attentional:</td>
<td>F(3, 339)=30.99*</td>
<td>0.22</td>
<td>F(2, 113)=23.85*</td>
<td>0.30</td>
<td>F(6, 339)=3.99*</td>
<td>0.07</td>
</tr>
<tr>
<td>Motor:</td>
<td>F(3, 339)=3.07*</td>
<td>0.03</td>
<td>F(2, 113)=1.19*</td>
<td>0.02</td>
<td>F(6, 339)=0.88</td>
<td>0.02</td>
</tr>
<tr>
<td>Non-planning:</td>
<td>F(3, 339)=34.08*</td>
<td>0.23</td>
<td>F(2, 113)=5.40*</td>
<td>0.09</td>
<td>F(6, 339)=4.23*</td>
<td>0.07</td>
</tr>
</tbody>
</table>

NB. *Bonferroni corrected p<0.017  † p<0.05  ^p<0.07
Table 5.3 shows that there were significant univariate main effects of phase and group for the attentional and non-planning BIS-11 subscales. The phase by group interactions were also significant for these subscales. Figures 5.7 to 5.9 display the impulsivity profiles of the three groups over the averaged screening cycles for each BIS-11 subscale.

Figure 5.7 The attentional impulsivity profiles of the three groups across the averaged screening cycles

Figure 5.8 The motor impulsivity profiles of the three groups across the averaged screening cycles

Figure 5.9 The non-planning impulsivity profiles of the three groups across the averaged screening cycles
5.4.1.2.2.1 Attentional impulsivity

For the attentional subscale, a simple effects analysis showed that there was a significant main effect of group in the follicular and luteal phases (F(2, 112) = 37.61, p<0.0125, partial \( \eta^2 = 0.40 \) and F(3, 112) = 11.66, p<0.0125, partial \( \eta^2 = 0.17 \) respectively). The co-morbid group reported attentional impulsivity significantly greater than both other groups in the follicular phase. The co-morbid and PMS groups reported significantly greater levels than the normally cycling women during the luteal phase (p<0.05). The PMS sufferers and the normally cycling women displayed variable profiles of attentional impulsivity across the cycle (F(3, 159)=39.32, p<0.0125, partial \( \eta^2 = 0.43 \) and F(3, 111) = 10.85, p<0.0125, partial \( \eta^2 = 0.23 \) respectively), while the co-morbid did not (F(3, 66) = 2.21, p>0.0125, partial \( \eta^2 = 0.09 \)). The attentional impulsivity profiles displayed in Figure 5.7 suggest the PMS sufferers displayed the greatest variability in attentional impulsivity reporting. This was confirmed through an interaction contrasts analysis, which showed the groups to differ significantly in their increase in attentional impulsivity between the follicular and luteal phases (F(2, 112) = 7.70, p<0.001, partial \( \eta^2 = 0.12 \)). Bonferroni pairwise comparisons revealed that the PMS sufferers reported a significantly greater increase on the attentional impulsivity subscale between these phases than the normally cycling and co-morbid groups (p<0.05).

5.4.1.2.2 Non-planning impulsivity

Non-planning impulsivity profiles (Figure 5.9) for the three groups appear similar to those for attentional impulsivity (Figure 5.7). As was found for the attentional subscale, a simple effects analysis showed that there was a significant main effect of group in the follicular and luteal phases (F(2, 112) = 9.20, p<0.0125, partial \( \eta^2 = 0.14 \)) and F(2, 112) = 4.57, p<0.0125, partial \( \eta^2 = 0.08 \) respectively). Again, the co-morbid group reported non-planning impulsivity significantly greater than both other groups during the follicular phase, whilst the co-morbid and PMS groups reported levels significantly greater than the normally cycling women premenstrually. Similarly to the attentional subscale, the PMS sufferers and normally cycling women displayed a variable pattern of non-planning impulsivity across the cycle (F(3, 159) = 43.77, p<0.0125, partial \( \eta^2 = 0.45 \) and F(3, 111) = 9.36, p<0.0125, partial \( \eta^2 = 0.20 \) respectively), while co-morbid did not (F(3, 66) = 3.06, p>0.0125. partial \( \eta^2 = 0.12 \)). The groups were found to show significantly different increases in non-planning impulsivity between the follicular and luteal phases (F(2, 112) = 9.70, p<0.001, partial
\(\eta^2=0.15\). in exactly the same way as they did for attentional impulsivity. The PMS sufferers reported a significantly greater increase between these phases than the normally cycling and co-morbid groups.

5.4.1.2.2.3 Motor impulsivity
As shown in Figure 5.8, the three groups showed great similarity in their motor impulsivity profiles, in terms of the pattern and severity of symptoms that they reported. This was confirmed by the univariate analyses displayed in Table 5.3, through the non-significant main effects of phase and group, and through the non-significant phase by group interactions. This suggests that all groups report similar stable patterns of motor impulsivity across the cycle.

5.4.2 Psychological symptom profiles for the four diagnostic groups
The psychological symptom profiles (aggression and impulsivity) that were reported by the PMS sufferers were then compared with those that were reported by the control, PMS non-reporter and co-morbid groups.

5.4.2.1 Aggression
5.4.2.1.1 BPAQ total scale scores
The BPAQ-total scores were examined through a 3 (cycle) x 4 (phase) x 4 (group) RM ANOVA. There was no significant main effect of cycle (F(2, 214)= 1.43, p>0.05) and no significant cycle by group interaction (F(6, 214) = 1.30, p>0.05). There was a significant main effect of phase (F(3, 321) = 24.53, p<0.001, partial \(\eta^2 =0.19\)), significant main effect of group (F(3, 107) = 4.744, p<0.001, partial \(\eta^2 =0.12\)) and significant phase by group interaction (F(9, 321) = 2.10, p<0.05, partial \(\eta^2 =0.06\)). Figure 5.10 displays the BPAQ profiles of the four groups over the averaged screening cycles.
Figure 5.10 The BPAQ-total profiles of the four groups averaged over the three screening cycles

Figure 5.10 suggests that the greatest group separation occurs in the follicular and luteal cycle phases. A simple effects analysis showed that there was a significant effect of group at each of these phases ($F(3, 108) = 8.98$, $p<0.0125$, partial $\eta^2 = 0.20$ and $F(3, 108) = 4.10$, $p<0.0125$, partial $\eta^2 = 0.10$ respectively). Bonferroni pairwise comparisons revealed that the co-morbid groups reported significantly greater BPAQ-total scores than the PMS and PMS non-reporter groups in the follicular phase, while the PMS and co-morbid groups reported significantly greater levels than the PMS non-reporters in the luteal phase ($p<0.05$).

A simple effects analysis of phase differences for each group revealed that the PMS and PMS non-reporter groups ($F(3, 150) = 40.29$, $p<0.0125$, partial $\eta^2 = 0.45$ and $F(3, 81) = 9.40$, $p<0.0125$, partial $\eta^2 = 0.26$ respectively) reported variable patterns of aggression across the cycle, while the controls ($F(3, 30) = 0.99$, $p>0.0125$, partial $\eta^2 = 0.09$) did not. Although the co-morbid groups displayed greater variability in aggression reporting than the controls, differences between cycle phases did not reach statistical significance when familywise Type I error rate was controlled for ($F(3, 63) = 4.87$, $p=0.015$, partial $\eta^2 = 0.19$).
An interaction contrasts analysis revealed that the four groups differed in their severity of symptom change between the follicular and luteal phases (\(F(3, 108) = 3.91, \ p<0.0125, \ \eta^2=0.098\)). However, Bonferroni pairwise comparisons revealed that there were no significant differences between the groups, although there was a trend for the PMS group to show a greater increase in BPAQ reporting between these phases than the controls (\(p=0.064\)).

5.4.2.1.2 The BPAQ subscales

The BPAQ subscale scores were examined through a 3 (cycle) x 4 (phase) x 4 (group) doubly multivariate analysis of variance. There was no significant main effect of cycle (multivariate \(F(8, 100) = 1.45, \ p>0.05\)) and no significant cycle by group interaction (multivariate \(F(24, 306) = 0.80, \ p>0.05\)), indicating that all groups were consistent in their reporting of aggression over the three screening cycles for the linear combination of the BPAQ subscales. There was a significant main effect of phase (multivariate \(F(12, 96) = 6.54, \ p<0.001, \ \eta^2= 0.45\)) and significant main effect of group (multivariate \(F(12, 318) = 4.23, \ p<0.001. \ \eta^2= 0.14\)). Moreover, the significant phase by group interaction (multivariate \(F(36, 294) = 1.79, \ p<0.01. \ \eta^2= 0.18\)) indicated that the pattern of aggression reported over the cycle differed between groups.

Given that there were some significant multivariate effects, the univariate tests of the main effects of phase and group and the phase by group interaction were examined (Field, 2005; Tabachnick and Fidell, 2007a). These effects are summarized in Table 5.4.

Table 5.4 Multivariate and univariate statistics for the main effects of phase and group, and for the phase by group interaction

<table>
<thead>
<tr>
<th>Phase Group</th>
<th>Phase</th>
<th>Group</th>
<th>Phase by Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase</td>
<td>Statistic</td>
<td>Partial (\eta^2)</td>
<td>Statistic</td>
</tr>
<tr>
<td>Multivariate</td>
<td>(F(12, 96)=6.54^*) 0.45</td>
<td>F(12, 318)=4.23* 0.14</td>
<td>F(36, 294)=1.79^* 0.18</td>
</tr>
<tr>
<td>Univariate:</td>
<td>-Anger (F(3, 321)=31.19^*) 0.23</td>
<td>F(3, 107)=5.41* 0.13</td>
<td>F(9, 321)=2.87* 0.07</td>
</tr>
<tr>
<td></td>
<td>-Verbal (F(3, 321)=21.15^*) 0.17</td>
<td>F(3, 107)=1.73 0.05</td>
<td>F(9, 321)=2.82* 0.07</td>
</tr>
<tr>
<td></td>
<td>-Hostility (F(3, 321)=14.75^*) 0.12</td>
<td>F(3, 107)=9.16* 0.20</td>
<td>F(9, 321)=2.05† 0.054</td>
</tr>
<tr>
<td></td>
<td>-Physical (F(3, 321)=14.08^*) 0.12</td>
<td>F(3, 107)=0.92 0.03</td>
<td>F(9, 321)=0.71 0.02</td>
</tr>
</tbody>
</table>

NB. *\(p<0.0125\) † \(p<0.05\)
Table 5.4 shows that there were significant univariate main effects of phase on all BPAQ subscales. There were significant main effects of group for anger and hostility. The phase by group interactions were significant for the anger and verbal subscales and marginally significant for hostility (p=0.034, Bonferroni corrected α = 0.0125). Figures 5.11 to 5.14 display the aggression profiles of the four groups over the averaged screening cycles for each BPAQ subscale.

Figure 5.11 The anger profiles of the four groups averaged over the three screening cycles

Figure 5.12 The verbal aggression profiles of the four groups averaged over the three screening cycles

Figure 5.13 The hostility profiles of the four groups averaged over the three screening cycles

Figure 5.14 The physical aggression profiles of the four groups averaged over the three screening cycles
5.4.2.1.2.1 Anger

For the anger subscale, a simple effects analysis of the luteal phase showed that there was a significant main effect of group \( (F(3,108) = 5.54, p<0.0125, \eta^2=0.13) \). The co-morbid and PMS groups rated anger significantly higher than the PMS non-reporters. All groups reported a variation of anger symptom reporting across the cycle (smallest \( F(3,63)=6.15, p<0.0125, \eta^2=0.23 \) except the controls \( F(3,30)=1.34, p>0.0125, \eta^2=0.12 \)). Figure 5.11 suggests that the PMS group displayed the greatest variability in anger reporting. This was confirmed through an interaction contrasts analysis, which showed that the groups displayed significantly different increases in anger between the follicular and luteal phases \( (F(3,108) = 6.39, p<0.0125, \eta^2=0.15) \). Bonferroni pairwise comparisons revealed that the PMS group reported a significantly greater increase between these phases than all other groups \( (p<0.05) \).

5.4.2.1.2.2 Verbal aggression

A simple effects analysis of the luteal phase showed that there was a trend towards an effect of group for verbal aggression \( (F(3,108) = 2.32, p=0.079, \eta^2=0.061) \), such that the PMS group showed a trend towards reporting greater verbal aggression than the PMS non-reporters \( (p=0.079) \). As for the anger subscale, all groups (smallest \( F(3,81) = 5.83, p<0.0125, \eta^2=0.18 \) except the controls \( F(3,30)=2.11, p>0.0125 \)) displayed variable patterns of verbal aggression across the cycle. Figure 5.12 suggests that verbal aggression is most variable for the PMS group. This was confirmed through an interaction contrasts analysis, which showed that the groups reported significantly different increases in verbal aggression between the follicular and luteal phases \( (F(3,108) = 4.65, p<0.01, \eta^2=0.11) \). The PMS group showed a significantly greater increase between these phases than the PMS non-reporters and controls. Unlike the anger subscale, there were no statistically significant differences between the co-morbid and PMS groups.

5.4.2.1.2.3 Hostility

Hostility profiles (see Figure 5.13) for the four groups appear similar to those for anger (see Figure 5.11). A simple effects analysis of the luteal phase showed that there was a significant main effect of group \( (F(3,108) = 6.79, p<0.0125, \eta^2=0.16) \). As for the anger subscale, the PMS and co-morbid groups reported significantly greater hostility than the PMS non-reporters. In addition, the co-morbid also reported significantly greater levels
than the controls. The PMS and PMS non-reporter groups reported variable levels of hostility across the cycle (smallest $F(3, 81) = 8.29, p<0.0125, \eta^2=0.23$). while the controls did not ($F(3, 30)=0.32, p>0.0125$). Although the co-morbids showed this trend (see Figure 5.13), their mean differences between phases did not reach statistical significance when multiple testing was controlled for ($F(3, 63) = 3.51, p=0.037, \eta^2=0.14$).

Similar to anger and verbal aggression, the PMS group displayed the greatest variability in hostility. An interaction contrasts analysis revealed that the groups showed significantly different increases in hostility between the follicular and luteal phases ($F(3, 108) = 3.96, p<0.01, \eta^2=0.10$). The PMS group showed a greater increase between these phases than the controls ($p<0.05$).

### 5.4.2.1.2.4 Physical aggression

As shown in Figure 5.14, the four groups showed similarity in physical aggression profiles, in terms of the pattern and severity of symptoms that they reported. This was confirmed by the univariate analyses displayed in Table 5.4. Although a significant main effect of phase was evident ($F(3, 321) = 14.08, p<0.0125, \eta^2=0.12$), there was no significant phase by group interaction, suggesting that the groups reported similar patterns of low physical aggression levels across the cycle, each with a slight luteal increase.

### 5.4.2.2 Impulsivity (BIS-11)

#### 5.4.2.2.1 BIS 11 total scale scores

A 3 (cycle) x 4 (phase) x 4 (group) RM ANOVA was conducted on the BIS-11 total scores. There was no significant main effect of cycle ($F(2, 222) = 0.41, p>0.05$) and no significant cycle by group interaction ($F(6, 22) = 1.48, p>0.05$). There was a significant main effect of phase ($F(3, 333) = 19.82, p<0.001, \eta^2= 0.15$) and significant main effect of group ($F(3, 111) = 5.13, p<0.01, \eta^2=0.12$). However, these effects were qualified by the significant phase by group interaction ($F(9, 333) = 3.19, p<0.01, \eta^2= 0.08$). Figure 5.15 displays the BIS-11 profiles of the four groups over the averaged screening cycles.
Figure 5.15 The BIS-11 profiles of the four diagnostic groups averaged over the three screening cycles

Figure 5.15 suggests that the greatest group separation occurs in the follicular phase. A simple effects analysis was conducted to explore group differences in impulsivity in the follicular and luteal phases. This showed that there was a significant main effect of group at each of these phases (F(3, 111) = 10.49, p<0.0125, partial $\eta^2= 0.22$ and F(3, 111)=3.30, p<0.0125, partial $\eta^2=0.08$ respectively). Bonferroni pairwise comparisons revealed that co-morbids reported significantly greater impulsivity levels than the PMS non-reporters at each phase, and greater levels than the PMS group during the follicular phase (p<0.05).

The BIS-11 total profiles displayed in Figure 5.15 suggest the PMS and PMS non-reporter groups display a curvilinear pattern of impulsivity across the cycle, in comparison to the co-morbids and controls who report more stable levels. This was confirmed through a simple effects analysis of phase differences for each group, which revealed that the PMS and PMS non-reporter groups (F(3, 159) = 40.82, p<0.0125, partial $\eta^2=0.44$ and F(3, 78) = 10.88, p<0.0125, partial $\eta^2=0.30$ respectively) displayed variable levels of impulsivity across the cycle, while the controls and co-morbids did not (F(3, 30) = 0.27, p>0.0125, partial $\eta^2=0.027$ and F(3, 66) = 3.29, p>0.0125, partial $\eta^2=0.13$ respectively).
An interaction contrasts analysis revealed that the four groups differed in their severity of symptom change between the follicular and luteal phases (F(3, 111) = 6.25, p<0.0125, partial η²=0.14). Bonferroni pairwise comparisons showed that the PMS group reported a significantly greater increase in impulsivity between these phases than the controls and co-morbid conditions (p<0.05). However, the PMS and PMS non-reporter groups did not differ statistically (p>0.05).

5.4.2.2 BIS-11 subscales

A 3 (cycle) x 4 (phase) x 4 (group) doubly multivariate analysis of variance was performed on the attentional, motor and non-planning BIS-11 subscales. There was no significant main effect of cycle (multivariate F(6, 107) = 1.62, p>0.05) and no significant cycle by group interaction (multivariate F(18, 327) = 1.16, p>0.05). There was a significant main effect of phase (multivariate F(9, 104) = 7.27, p<0.001, partial η²= 0.39) and significant main effect of group (multivariate F(9, 336) = 5.87, p<0.001, partial η²=0.14). However, these main effects were qualified by the significant phase by group interaction (multivariate F(27, 318) = 1.54, p<0.05, partial η²= 0.12).

Given that there were some significant multivariate effects, the univariate tests of the main effects of phase and group and the phase by group interaction were examined (Field, 2005; Tabachnick and Fidell, 2007a). These effects are summarized in Table 5.5.

<table>
<thead>
<tr>
<th>Phase Group Phase by Group</th>
<th>Statistic</th>
<th>Partial η²</th>
<th>Statistic</th>
<th>Partial η²</th>
<th>Statistic</th>
<th>Partial η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multivariate</td>
<td>F(9, 104) = 7.27*</td>
<td>0.39</td>
<td>F(9, 336) = 5.87*</td>
<td>0.14</td>
<td>F(27, 318) = 1.54*</td>
<td>0.12</td>
</tr>
<tr>
<td>Univariate:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Attentional</td>
<td>F(3, 336) = 19.37*</td>
<td>0.15</td>
<td>F(3, 112) = 16.02*</td>
<td>0.30</td>
<td>F(9, 336) = 3.41*</td>
<td>0.08</td>
</tr>
<tr>
<td>- Motor</td>
<td>F(3, 336) = 1.92</td>
<td>0.02</td>
<td>F(3, 112) = 1.14</td>
<td>0.03</td>
<td>F(9, 336) = 0.68</td>
<td>0.02</td>
</tr>
<tr>
<td>- Non-planning</td>
<td>F(3, 336) = 21.86*</td>
<td>0.16</td>
<td>F(3, 112) = 3.65*</td>
<td>0.09</td>
<td>F(9, 336) = 3.51*</td>
<td>0.09</td>
</tr>
</tbody>
</table>

NB. *Bonferroni corrected p<0.017
Table 5.5 shows that there were significant univariate main effects of phase and group for the attentional and non-planning BIS-11 subscales. The phase by group interactions were also significant for these subscales. Figures 5.16 to 5.18 display the impulsivity profiles of the four groups over the averaged screening cycles for each BIS-11 subscale.

**Figure 5.16** The attentional impulsivity profiles of the four groups across the averaged screening cycles

**Figure 5.17** The motor impulsivity profiles of the four groups across the averaged screening cycles

**Figure 5.18** The non-planning impulsivity profiles of the four groups across the averaged screening cycles
5.4.2.2.1 Attentional impulsivity
For the attentional subscale, a simple effects analysis showed that there was a significant main effect of group in the follicular and luteal phases (F(3, 111) = 27.08, p<0.0125, partial $\eta^2$=0.42 and F(3, 111) = 7.84, p<0.0125, partial $\eta^2$=0.17 respectively). The co-morbid group reported attentional impulsivity significantly greater than all groups in the follicular phase and greater than the controls and PMS non-reporters premenstrually. There was also a trend for the PMS group to report higher levels than the PMS non-reporters premenstrually (p=0.085). The PMS and PMS non-reporter groups displayed variable profiles of attentional impulsivity across the cycle (F(3, 159)=39.32, p<0.0125, partial $\eta^2$=0.43 and F(3, 78) = 12.95, p<0.0125, partial $\eta^2$=0.33 respectively), while the controls and co-morbids did not (F(3, 30) = 0.83, p>0.0125, partial $\eta^2$=0.08 and F(3, 66) = 2.21, p>0.0125, partial $\eta^2$= 0.09 respectively). The attentional impulsivity profiles displayed in Figure 5.16 suggest the PMS group displayed the greatest variability in attentional impulsivity reporting. This was confirmed by an interaction contrasts analysis, which showed the groups to differ significantly in their increase in attentional impulsivity between the follicular and luteal phases (F(3, 111) = 7.06, p<0.001, partial $\eta^2$=0.16). Bonferroni pairwise comparisons revealed that the PMS group reported a significantly greater increase on the attentional impulsivity subscale between these phases than the control and co-morbid groups (p<0.05).

5.4.2.2.2 Non-planning impulsivity
Non-planning impulsivity profiles (Figure 5.18) for the four groups appear similar to those for attentional impulsivity (Figure 5.16). For the non-planning subscale, a simple effects analysis showed that although there was no significant main effect of group in the luteal phase (F(3, 111) = 3.05, p>0.0125, partial $\eta^2$=0.08), the main effect of group was significant in the follicular phase (F(3, 111) = 6.96, p<0.0125, partial $\eta^2$=0.16). The co-morbid group reported non-planning impulsivity to be significantly higher than the PMS and PMS non-reporter groups in the follicular phase. Similarly to the attentional subscale, the PMS and PMS non-reporter groups displayed a variable pattern of non-planning impulsivity across the cycle (F(3, 159) = 43.77, p<0.0125, partial $\eta^2$=0.45 and F(3, 78) = 9.82, p<0.0125, partial $\eta^2$=0.27 respectively), while the controls and co-morbids did not (F(3, 30) = 0.51, p>0.0125, partial $\eta^2$=0.05 and F(3, 66) = 3.06, p>0.0125, partial $\eta^2$= 0.12 respectively). The groups were found to show significantly different increases in non-
planning impulsivity between the follicular and luteal phases ($F(3, 111) = 7.94, p<0.001$, partial $\eta^2=0.18$), in exactly the same way as they did for attentional impulsivity. The PMS group reported a significantly greater increase between these phases than the control and co-morbid groups.

### 5.4.2.2.3 Motor impulsivity

As shown in Figure 5.17, the four groups showed great similarity in their motor impulsivity profiles, in terms of the pattern and severity of symptoms that they reported. This was confirmed by the univariate analyses displayed in Table 5.5, through the non-significant main effects of phase and group, and through the non-significant phase by group interactions. This suggests that all groups report similar stable patterns of motor impulsivity across the cycle.

### 5.4.3 Summary of self-report measures

Table 5.6 summarises the statistical differences that were apparent between the PMS sufferers and the other groups of women on the BPAQ and BIS-11, when the participants were categorised into both three (PMS sufferers, normally cycling women and co-morbids) and four (PMS sufferers, controls, PMS non-reporters and co-morbids) diagnostic groups. Table 5.6 displays the statistical differences that were apparent between the PMS sufferers and the other diagnostic groups, in terms of their luteal scores and their follicular to luteal changes scores. It also indicates which groups displayed a variation of symptom reporting across the cycle on each of these measures. The physical aggression BPAQ subscale and motor impulsivity BIS-11 subscale are omitted from the table, as the groups reported comparable profiles on these measures, when women were categorised into both three and four diagnostic groups (see Tables 5.2, 5.3, 5.4 and 5.5).

For example, Table 5.6 shows that when the women were categorised into three diagnostic groups (PMS, normally cycling, co-morbids), the normally cycling women had significantly lower BPAQ-total scores than the PMS sufferers during the luteal phase. Thenormally cycling women had a significantly smaller follicular to luteal change score than the PMS sufferers, while the co-morbids showed a trend towards having a smaller follicular to luteal change score than the PMS sufferers. The PMS, normally cycling and co-morbid groups all showed a significant main effect of phase for aggression (BPAQ-total score).
Moreover, Table 5.6 shows that when the women were categorised into four diagnostic groups (PMS, controls, PMS non-reporters co-morbids), the PMS non-reporters had significantly lower BPAQ-total scores than the PMS sufferers premenstrually. The PMS and PMS non-reporter groups both showed a significant main effect of phase for aggression, whilst the co-morbids showed this trend (BPAQ-total score).
Table 5.6 Phase effects and luteal and follicular to luteal statistical differences between the PMS sufferers and the other groups of women, on the BPAQ and BIS-11 (and their subscales), when women were categorised into both three and four diagnostic groups

<table>
<thead>
<tr>
<th>Measure</th>
<th>Group</th>
<th>Luteal Score</th>
<th>Follicular to luteal change score</th>
<th>Groups showing significant effect of cycle phase</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Three groups</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPAQ-total</td>
<td>Normally cycling</td>
<td>-</td>
<td>-</td>
<td>PMS</td>
</tr>
<tr>
<td></td>
<td>co-morbids</td>
<td>~ (-)</td>
<td></td>
<td>Normally cycling</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Co-morbids (trend)</td>
</tr>
<tr>
<td>-BPAQ anger</td>
<td>Normally cycling</td>
<td>-</td>
<td>-</td>
<td>PMS</td>
</tr>
<tr>
<td></td>
<td>co-morbids</td>
<td>-</td>
<td></td>
<td>Normally cycling</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Co-morbids</td>
</tr>
<tr>
<td>-BPAQ verbal</td>
<td>Normally cycling</td>
<td>-</td>
<td>-</td>
<td>PMS</td>
</tr>
<tr>
<td></td>
<td>co-morbids</td>
<td>~ (-)</td>
<td></td>
<td>Normally cycling</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Co-morbids (trend)</td>
</tr>
<tr>
<td>-BPAQ hostility</td>
<td></td>
<td>-</td>
<td>-</td>
<td>PMS</td>
</tr>
<tr>
<td></td>
<td>co-morbids</td>
<td>~ (-)</td>
<td></td>
<td>Normally cycling</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Co-morbids (trend)</td>
</tr>
<tr>
<td>BIS-11 total</td>
<td>Normally cycling</td>
<td>~ (-)</td>
<td>-</td>
<td>PMS</td>
</tr>
<tr>
<td></td>
<td>co-morbids</td>
<td>-</td>
<td></td>
<td>Normally cycling</td>
</tr>
<tr>
<td>-BIS attentional</td>
<td></td>
<td>-</td>
<td>-</td>
<td>PMS</td>
</tr>
<tr>
<td></td>
<td>co-morbids</td>
<td>-</td>
<td></td>
<td>Normally cycling</td>
</tr>
<tr>
<td>- BIS non-planning</td>
<td></td>
<td>-</td>
<td>-</td>
<td>PMS</td>
</tr>
<tr>
<td></td>
<td>co-morbids</td>
<td>-</td>
<td></td>
<td>Normally cycling</td>
</tr>
<tr>
<td><strong>Four groups</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPAQ-total</td>
<td>Non-reporters</td>
<td>-</td>
<td></td>
<td>PMS</td>
</tr>
<tr>
<td></td>
<td>Controls</td>
<td></td>
<td></td>
<td>Non-reporters</td>
</tr>
<tr>
<td></td>
<td>co-morbids</td>
<td></td>
<td></td>
<td>Co-morbids (trend)</td>
</tr>
<tr>
<td>-BPAQ anger</td>
<td>Non-reporters</td>
<td>-</td>
<td>-</td>
<td>PMS</td>
</tr>
<tr>
<td></td>
<td>Controls</td>
<td></td>
<td></td>
<td>Non-reporters</td>
</tr>
<tr>
<td></td>
<td>co-morbids</td>
<td>-</td>
<td></td>
<td>Co-morbids</td>
</tr>
<tr>
<td>-BPAQ verbal</td>
<td>Non-reporters</td>
<td>~ (-)</td>
<td>-</td>
<td>PMS</td>
</tr>
<tr>
<td></td>
<td>Controls</td>
<td></td>
<td></td>
<td>Non-reporters</td>
</tr>
<tr>
<td></td>
<td>co-morbids</td>
<td>-</td>
<td></td>
<td>Co-morbids</td>
</tr>
<tr>
<td>60</td>
<td>-BPAQ hostility</td>
<td>-</td>
<td></td>
<td>PMS</td>
</tr>
<tr>
<td></td>
<td>Non-reporters</td>
<td>-</td>
<td></td>
<td>Non-reporters</td>
</tr>
<tr>
<td></td>
<td>Controls</td>
<td></td>
<td></td>
<td>Co-morbids (trend)</td>
</tr>
<tr>
<td>BIS-11 total</td>
<td>Non-reporters</td>
<td>-</td>
<td></td>
<td>PMS</td>
</tr>
<tr>
<td></td>
<td>Controls</td>
<td></td>
<td></td>
<td>Non-reporters</td>
</tr>
<tr>
<td></td>
<td>co-morbids</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-BIS attentional</td>
<td></td>
<td>~ (-)</td>
<td>-</td>
<td>PMS</td>
</tr>
<tr>
<td></td>
<td>co-morbids</td>
<td>-</td>
<td></td>
<td>Non-reporters</td>
</tr>
<tr>
<td>-BIS non-planning</td>
<td></td>
<td>-</td>
<td>-</td>
<td>PMS</td>
</tr>
<tr>
<td></td>
<td>co-morbids</td>
<td>-</td>
<td></td>
<td>Non-reporters</td>
</tr>
</tbody>
</table>

**KEY:** Less than PMS group (-), Greater than PMS group (+), Trend towards less than PMS group (~ (-))
5.4.4 Biological measures

5.4.4.1 PMS sufferers vs. normally cycling women

5.4.4.1.1 Steroid hormones

Figures 5.19 to 5.24 display the mean steroid hormone concentrations for the PMS sufferers and normally cycling women (PMS non-reporters and controls) during the follicular and luteal phases of the third screening cycle.
Six RM ANOVAs were conducted, one for each steroid hormone measured during the study. The two diagnostic groups formed the between subjects factor; PMS sufferers and normally cycling women (PMS non-reporter and control groups). The two cycle phases formed the within subjects factor. The results of the univariate tests of the main effects of cycle phase, group and of the cycle phase by group interactions can be found in Table 5.7.

Table 5.7 Univariate statistics for the effects of cycle phase, group, and cycle phase by group interactions for the steroid hormone analyses

<table>
<thead>
<tr>
<th></th>
<th>Cycle phase</th>
<th>Group</th>
<th>Cycle phase by Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
<td>Partial η²</td>
<td>Statistic</td>
</tr>
<tr>
<td>Oestradiol</td>
<td>F(1, 61)=1.65</td>
<td>0.03</td>
<td>F(1, 61)=3.28</td>
</tr>
<tr>
<td>Progesterone</td>
<td>F(1, 61)=35.89**</td>
<td>0.37</td>
<td>F(1, 61)=0.53</td>
</tr>
<tr>
<td>Testosterone</td>
<td>F(1, 61)=2.28</td>
<td>0.04</td>
<td>F(1, 61)=7.68**</td>
</tr>
<tr>
<td>LH</td>
<td>F(1, 61)=25.00**</td>
<td>0.29</td>
<td>F(1, 61)=4.02</td>
</tr>
<tr>
<td>FSH</td>
<td>F(1, 61)=40.35**</td>
<td>0.40</td>
<td>F(1, 61)=2.38</td>
</tr>
<tr>
<td>PRL</td>
<td>F(1, 61)=4.66*</td>
<td>0.07</td>
<td>F(1, 61)=1.86</td>
</tr>
</tbody>
</table>

** p<0.01  * p<0.05

It is evident from Table 5.7 that there were significant main effects of cycle phase for progesterone, LH, FSH and prolactin, such that women exhibited significantly greater levels of LH and FSH during the follicular phase (Figures 5.22 and 5.23) and significantly greater levels of progesterone and prolactin during the luteal phase (Figures 5.20 and 5.24). There was a significant main effect of group for testosterone, such that the PMS sufferers exhibited significantly greater testosterone concentrations than the normally cycling women (Figure 5.21). The cycle phase by group interaction was non-significant for testosterone, indicating that the PMS sufferers had higher testosterone concentrations than the normally cycling women during both the follicular and luteal phases.
5.4.4.1.2 Cytokines

Figures 5.25 to 5.29 display the mean cytokine concentrations for the PMS sufferers and normally cycling women (PMS non-reporters and controls) during the follicular and luteal phases of the third screening cycle.

Figure 5.25 Mean IL-1β concentration

Figure 5.26 Mean IL-6 concentration

Figure 5.27 Mean IL-8 concentration

Figure 5.28 Mean TNF-α concentration

Figure 5.29 Mean IFN-γ concentration
Five 2 (cycle phase) x 2 (group) RM ANOVAs were conducted on the cytokine data. The univariate tests of the main effects of cycle phase, group and of the cycle phase by group interactions of these analyses can be found in Table 5.8.

Table 5.8 Univariate statistics for the effects of cycle phase, group, and the cycle phase by group interactions for the cytokine analyses

<table>
<thead>
<tr>
<th>Cycle phase</th>
<th>Group</th>
<th>Cycle phase by Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistic</td>
<td>Partial η²</td>
<td>Statistic</td>
</tr>
<tr>
<td>- IL-1β</td>
<td>F(1, 61)=0.22</td>
<td>0.004</td>
</tr>
<tr>
<td>- IL-6</td>
<td>F(1, 61)=0.12</td>
<td>0.002</td>
</tr>
<tr>
<td>- IL-8</td>
<td>F(1, 61)=0.01</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>- TNF-α</td>
<td>F(1, 61)=2.25</td>
<td>0.04</td>
</tr>
<tr>
<td>- IFN-γ</td>
<td>F(1, 61)=0.09</td>
<td>0.002</td>
</tr>
</tbody>
</table>

* p<0.05

Table 5.8 shows that there was no significant main effect of cycle phase for any of the cytokines measured. Figures 5.25 to 5.29 indicate that the PMS sufferers had greater concentrations of all cytokines during the follicular and luteal phases than the normally cycling women. These main effects of group were significant for IL-6, IL-8 and TNF-α. The non-significant cycle phase by group interactions revealed that these effects were not limited to a specific cycle phase, but that PMS sufferers had greater levels of IL-6, IL-8 and TNF-α overall.
5.4.4.2 PMS sufferers, PMS non-reporters and controls

5.4.4.2.1 Steroid hormones

Figures 5.30 to 5.35 display the mean steroid hormone concentrations for the PMS sufferers, PMS non-reporters and controls during the follicular and luteal phases of the third screening cycle.

Figure 5.30 Mean oestradiol concentration

Figure 5.31 Mean progesterone concentration

Figure 5.32 Mean testosterone concentration

Figure 5.33 Mean LH concentration

Figure 5.34 Mean FSH concentration

Figure 5.35 Mean PRL concentration
Six 2 (cycle phase) x 3 (group) RM ANOVAs were conducted, one for each steroid hormone measured during the study. The results of the univariate tests of the main effects of cycle phase, group and of the cycle phase by group interactions can be found in Table 5.9.

### Table 5.9 Univariate statistics for the effects of cycle phase, group, and cycle phase by group interactions for the steroid hormone analyses

<table>
<thead>
<tr>
<th>Cycle phase</th>
<th>Group</th>
<th>Cycle phase by Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
<td>Partial η^2</td>
</tr>
<tr>
<td>Oestradiol</td>
<td>F(1, 60)=2.65</td>
<td>0.042</td>
</tr>
<tr>
<td>Progesterone</td>
<td>F(1, 60)=37.97**</td>
<td>0.39</td>
</tr>
<tr>
<td>Testosterone</td>
<td>F(1, 60)=0.99</td>
<td>0.016</td>
</tr>
<tr>
<td>LH</td>
<td>F(1, 60)=21.82**</td>
<td>0.27</td>
</tr>
<tr>
<td>FSH</td>
<td>F(1, 60)=42.57**</td>
<td>0.42</td>
</tr>
<tr>
<td>PRL</td>
<td>F(1, 60)=6.11*</td>
<td>0.092</td>
</tr>
</tbody>
</table>

** p<0.001 * p<0.01

It is evident from Table 5.9 that there were significant main effects of cycle phase for progesterone, LH, FSH and prolactin, such that women exhibited significantly greater levels of LH and FSH during the follicular phase (Figures 5.33 and 5.34) and significantly greater levels of progesterone and prolactin during the luteal phase (Figures 5.31 and 5.35). There was a significant main effect of group for testosterone. Figure 5.32 indicates that PMS sufferers had significantly greater testosterone concentrations than the PMS non-reporter and control groups. The cycle phase by group interaction was non-significant, indicating that the PMS sufferers had greater testosterone concentrations than the PMS non-reporter and control groups during both the follicular and luteal phases.

#### 5.4.4.2.2 Cytokines

Figures 5.36 to 5.40 display the mean cytokine concentrations for the PMS sufferers, PMS non-reporters and controls during the follicular and luteal phases of the third screening cycle.
Figure 5.36 Mean IL-1β concentration

Figure 5.37 Mean IL-6 concentration

Figure 5.38 Mean IL-8 concentration

Figure 5.39 Mean TNF-α concentration

Figure 5.40 Mean IFN-γ concentration
Five 2 (cycle phase) x 3 (group) RM ANOVAs were conducted on the cytokine data. The univariate tests of the main effects of cycle phase, group and of the cycle phase by group interactions of these analyses can be found in Table 5.10

Table 5.10 Univariate statistics for the effects of cycle phase, group, and the cycle phase by group interactions for the cytokine analyses

<table>
<thead>
<tr>
<th>Cycle phase</th>
<th>Group</th>
<th>Cycle phase by Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
<td>Partial η²</td>
</tr>
<tr>
<td>- IL-1β</td>
<td>F(1, 60)=0.006</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>- IL-6</td>
<td>F(1, 60)=0.089</td>
<td>0.001</td>
</tr>
<tr>
<td>- IL-8</td>
<td>F(1, 60)=0.004</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>- TNF-α</td>
<td>F(1, 60)=0.60</td>
<td>0.01</td>
</tr>
<tr>
<td>- IFN-γ</td>
<td>F(1, 60)=0.28</td>
<td>0.005</td>
</tr>
</tbody>
</table>

*p<0.05

Table 5.10 shows that there was no significant main effect of cycle phase for any of the cytokines measured. Figures 5.36 to 5.40 indicate that PMS sufferers had greater concentrations of all cytokines, except IFN-γ, during the follicular and luteal phases than the PMS non-reporter and control groups. These main effects of group were significant for IL-8 and TNF-α. The non-significant cycle phase by group interactions revealed that these effects were not limited to specific cycle phases, but that PMS sufferers had greater levels of IL-8 and TNF-α overall.
5.5 Discussion

5.5.1 Self-report measures

The study presented in this chapter firstly compared the aggression and impulsivity profiles reported by PMS sufferers who self-diagnosed themselves as suffering from the condition, and who met a PMS diagnosis through the use of the 30% increase criterion, with those reported by normally cycling women, who did not consider themselves to experience problematic PMS symptoms (controls and PMS non-reporters). Furthermore, the aggression and impulsivity profiles reported by PMS sufferers were compared with those reported by women who met criteria for anxiety and/or depression (co-morbids), most of whom also self-reported having PMS. Aggression and impulsivity levels were assessed using validated and widely used measures (see sections 3.1.1.2 to 3.1.1.5).

PMS sufferers were found to report cyclicity of all forms of aggression and impulsivity, with the exception of motor impulsivity. They demonstrated a curvilinear pattern of aggression (BPAQ-total score, physical, verbal, anger and hostility subscales) and impulsivity (BIS-11 total score, attentional and non-planning subscales) across the cycle, and reported the highest level of these symptoms during the luteal phase. Normally cycling women also reported variable levels of aggression (BPAQ-total score, physical, verbal, anger and hostility) and impulsivity (BIS-11 total score, attentional and non-planning) across the cycle. Their symptom profiles on these measures followed a similar curvilinear pattern to those of the PMS sufferers, albeit with smaller follicular to luteal increments.

Although the co-morbids reported stable profiles of aggression (BPAQ total score) and impulsivity (BIS-11 total score, attentional, non-planning), they reported variable levels of anger, hostility and verbal aggression across the cycle. Interestingly, the co-morbids reported greater levels of impulsivity (BIS-11 total score, attentional and non-planning impulsivity) and aggression (total BPAQ score, physical aggression, anger and hostility) across the cycle than the PMS sufferers and normally cycling women. Although it is possible that the co-morbid group comprised over-reporters, or women who tended to use the extremes of the scales, this seems unlikely, as the variances of their scores were comparable to those of the other groups (Figures 5.1 to 5.9). If these women reported higher symptom levels across the cycle due to a tendency to use the extreme ends of the scales, then it would be expected that they would demonstrate greater variances. The PMS
sufferers reported statistically comparable premenstrual levels to the co-morbids for all forms of aggression and impulsivity that were assessed (total scale scores and subscales).

5.5.2 ‘Normal’ and ‘abnormal’ symptoms
The aggression and impulsivity profiles reported by normally cycling women (PMS non-reporters and controls) were compared with those reported by PMS sufferers, to establish what separates ‘normal’ from ‘abnormal’ symptom profiles. Both the normally cycling women and the PMS sufferers demonstrated a curvilinear pattern of aggression and impulsivity across the menstrual cycle (total scale scores and subscales). However, the PMS sufferers reported a significantly greater increase in aggression (BPAQ total score, verbal, anger and hostility) and impulsivity (BIS-11 total scores, attentional and non-planning) between the follicular and luteal phases than the normally cycling women. Moreover, the PMS sufferers reported significantly greater levels of aggression (BPAQ total score, anger and hostility) and impulsivity (attentional and non-planning) than the normally cycling women during the luteal phase. As both women experiencing ‘normal’ and ‘abnormal’ cycles seemed to display variable profiles of aggression and impulsivity across the cycle, cyclicity of aggression and impulsivity may fall along a continuum. ‘Abnormality,’ or ‘caseness’ with respect to PMS diagnosis, could be said to arise when the follicular to luteal symptom increase, and the premenstrual symptom severity, exceed a certain level. As such, ‘normal’ and ‘abnormal’ cycles appear to differ quantitatively rather than qualitatively. Further work should determine the appropriate level for “caseness.”

5.5.3 Heterogeneous PMS and control groups
5.5.3.1 Exclusion of PMS non-reporters from controls groups
Some researchers investigate whether cyclical mood changes represent an abnormal symptom profile, or are variations of a ‘normal’ cyclical pattern, by comparing the symptom profiles of PMS sufferers with those reported by ‘control’ women who do not self-diagnose with PMS and who do not meet PMS criteria (e.g. Bond et al., 2003; De Ronchi et al., 2005; Hallman et al., 1987). If this approach was adopted here, through the comparison of the symptom profiles reported by PMS sufferers with those reported by control women, then the findings would suggest that ‘normal’ women report stable aggression (BPAQ total, verbal aggression, anger and hostility) and impulsivity (BIS-11 total, attentional and non-planning) levels across cycle phases, while women experiencing
'abnormal' cycles report variable levels. As such it would seem that 'normal' and 'abnormal' cycles differ qualitatively rather than quantitatively. Moreover, the findings would suggest that women experiencing 'normal' menstrual cycles do not report significantly different aggression (anger and hostility) and impulsivity (attentional and non-planning) levels to PMS sufferers during the luteal phase.

However, this comparison is overly simplistic when liberal PMS criteria have been used to form diagnostic groups. Previous research has demonstrated that the use of liberal PMS criteria can result in a large proportion of women who do not consider themselves to experience problematic PMS symptoms meeting a PMS diagnosis (Gallant et al., 1992a; Morse et al., 1988). The use of the 30% increase criterion in this research resulted in the confirmation of a PMS diagnosis in approximately two thirds of women who did not self-diagnose with PMS (PMS non-reporters) (see section 3.5.2.1). As discussed in Section 3.5.2.3, these women reported premenstrual symptoms at an almost identical severity to control women, who did not self-diagnose with PMS and who did not meet the 30% increase criterion, during the luteal phase. In fact, the main difference between the PMS non-reporter and control groups was that the PMS non-reporters reported symptoms at a slightly lower severity than the controls during the follicular phase. The 30% increase criterion bases its diagnosis on the change in symptoms that are reported between the follicular and luteal phases. Hence its use resulted in the division of normally cycling women into two groups on the basis of women's follicular phase symptom severity (see section 3.5.2.3). As such, the use of the 30% increase criterion resulted in the confirmation of a PMS diagnosis in a group of women who should not have been considered to be PMS sufferers, but who should instead have been considered to experience 'normal cyclicity,' and included in the control group (false positives). If researchers only select women who do not meet a PMS diagnosis following the use of liberal PMS criteria, then it would appear that PMS sufferers are only being compared with an unusual minority of normally cycling women, who report slightly raised symptom levels during the follicular phase. The differences in results that were produced in this research when the PMS non-reporters and controls comprised the group of normally cycling women, compared to when only control women were included, demonstrates that the omission of PMS non-reporters from control groups can dramatically alter research findings. When only control women were included, normally cycling women were shown to report stable levels of aggression and impulsivity.
across the cycle. However, when a more representative ‘control’ sample was selected (PMS non-reporters and controls), the ‘normal symptom profile’ was shown to be cyclical, with the follicular to luteal change in symptoms and the premenstrual symptom severity separating ‘normal’ from ‘abnormal’ symptom profiles. These results demonstrate the importance of taking into account the diagnostic criteria that have been used to form diagnostic groups when selecting a ‘control’ sample in PMS studies.

5.5.3.2 Impact of retrospective group allocation
Disguising the purpose of the research and allocating women to groups after they have completed symptom reports is sometimes employed to reduce the likelihood of demand characteristics influencing reporting behaviour (e.g. Christensen and Oei, 1988; Dougherty et al., 1998; Keenan et al., 1992). Retrospective designation of women as PMS sufferers following the use of liberal PMS criteria could result in PMS non-reporters being considered PMS sufferers rather than being included in ‘control’ groups. Although the aggression and impulsivity symptom profiles of the PMS non-reporters followed a similar curvilinear pattern to those of the PMS sufferers in this research, the PMS non-reporters demonstrated smaller follicular to luteal increments in symptoms. In fact, the PMS non-reporters demonstrated significantly smaller follicular to luteal change scores on the anger and verbal subscales of the BPAQ than the PMS sufferers. Furthermore, the PMS non-reporters experienced lower aggression and impulsivity levels than the PMS sufferers during the luteal phase (total scale scores and subscales). These differences reached statistical significance for the BPAQ-total score, anger and hostility, whilst this trend was shown for verbal aggression and attentional impulsivity. Therefore, the inclusion of PMS non-reporters in PMS groups following retrospective group allocation could result in an underestimation of the premenstrual symptom severity and the cyclicity of symptoms that PMS sufferers experience.

5.5.3.3 Inclusion of women with co-morbidities in PMS groups
Although PMS symptoms should be limited to the luteal phase (Freeman, 2003; Frye and Silverman, 2000; Halbreich et al., 2007; Lampe, 2005; see also section 1.3.1.4), many researchers do not consider whether women are suffering from co-morbid conditions (e.g. Howard et al., 1988; Van der Ploeg, 1987; Watts et al., 1980). This could result in women with co-morbidities being considered PMS sufferers. The main difference between the
aggression and impulsivity profiles that were reported by the PMS sufferer and co-morbid groups was that the co-morbid groups reported more stable but elevated symptom levels across the cycle, and demonstrated smaller follicular to luteal increments in these symptoms (see Table 5.6). The co-morbid groups reported a significantly smaller symptom increase between these phases than the PMS sufferers for anger, attentional and non-planning impulsivity. Moreover, whilst the PMS sufferers demonstrated impulsivity levels that varied by cycle phase, the co-morbid groups reported higher but stable levels across the cycle (BIS-11 total score, attentional and non-planning). Therefore, failure to identify co-morbidities in PMS studies could obscure the cyclicity of symptom reporting that PMS sufferers experience.

5.5.4 Effect of recruitment strategy

It must be acknowledged that the recruitment methods that were employed in this study may have influenced participants' reporting behaviour. Two forms of recruitment method were employed. Firstly, women considering themselves to experience normal menstrual cycles were recruited to take part in a study advertised as assessing changes in mood and behaviour across the menstrual cycle. These women formed the group of normally cycling women (PMS non-reporters and controls) (see section 3.5.3). Secondly, women considering themselves to suffer from PMS were recruited to participate in the clinical trial detailed in Chapter 8. All PMS sufferers and the majority of co-morbid groups were recruited via this advertising (see section 3.7). Although some of these women (see Figure 3.1) were immediately informed that they were not eligible to take part in the clinical trial and completed symptom reports in order to gain more insight into their menstrual cycle, more than half of these women completed symptom reports during screening cycles that were used to assess their eligibility to enter the treatment phase of the clinical trial i.e. to confirm prospectively that they met criteria for PMS (see section 3.6). The desire to try a therapeutic intervention to alleviate their symptoms may have motivated these women to report higher symptom levels during these screening cycles in order to meet the entry criteria for the clinical trial (Gallant et al., 1992b), which may have resulted in an inflated estimation of symptom severity in these groups.

All groups of women were informed that the study focus involved the menstrual cycle. Aubuchon and Calhoun (1985) found that women reported significantly more negative psychological and somatic symptoms during the premenstrual and menstrual cycle phases if
they were informed that the study was interested in menstrual cycle symptomatology. However, these researchers did not consider whether women self-diagnosed with PMS or the impact of this on symptom profiles. Gallant et al. (1992b) found that knowledge of the study focus did not influence reporting behaviour in women who self-diagnosed with severe premenstrual symptoms and who met a provisional diagnosis of PMDD, and suggested that the topic of the menstrual cycle may have already been so salient for these women that the knowledge of the study focus did not draw any additional attention to it. However, Gallant et al. (1992b) found that awareness of the study focus resulted in slightly increased cyclicity of prospective symptom reporting in women who did not report problematic PMS symptoms. Therefore, knowledge of the study focus in this study may have resulted in demand characteristics and expectations increasing cyclicity of symptom reporting in the group of normally cycling women (PMS non-reporters and controls), as these women did not self-diagnose with PMS. However, any such exaggeration effects were likely to be minimised since these women were made explicitly aware that the study required women who experience ‘normal’ menstrual cycles. Therefore, the normally cycling women may have actually reported less cyclicity of symptoms in response to these demand characteristics.

5.5.5 Biological data
Unfortunately, the serotonin data (5-HT, 5-HIAA) could not be analysed (see section 5.3.2.2.3). Therefore the potential relationship between the serotonergic system, symptom profiles and diagnostic groups could not be explored. However, levels of the steroid hormones FSH, LH, oestradiol, progesterone, prolactin and testosterone, and the cytokines IL-1β, IL-6, IL-8, IFN-γ, TNF-α, were successfully assessed during the follicular and luteal phases of the third screening cycle for the PMS sufferers and normally cycling women (PMS non-reporters and controls). In line with previous research (Connolly, 2001; Dennerstein et al., 1993; Rubinow, 1992), the steroid hormone analysis revealed that the PMS sufferers and normally cycling women demonstrated similar profiles of FSH, LH, progesterone and oestradiol across the cycle phases.

The groups did however, exhibit significantly different concentrations of testosterone, such that PMS sufferers had higher levels than the normally cycling women during both the follicular and luteal phases. Increased androgen levels have been associated with PMS
symptoms, including anger, aggression and impulsive behaviour (Dabbs and Hargrove, 1997; Harris et al., 1996; Steiner et al., 2003b; Von der Pahlen et al., 2002). Moreover, the administration of an androgen antagonist, spironolactone, has been shown to benefit PMS symptoms in women with PMS who had high premenstrual testosterone levels (Burnet et al., 1991). This has led researchers to propose that testosterone may be involved in the production of PMS symptoms (Rapkin, 2003), especially irritability and impulsivity (Dougherty et al., 1997a; Sunblad et al., 1994), possibly through interaction with the serotonergic system (Steiner et al., 2003b). The findings from this research support this theory, and previous research which has found PMS sufferers to have higher testosterone levels than women without PMS during the follicular and luteal phases (Eriksson et al., 1992; 1994). However, more research is needed to clarify these findings, as other researchers have found testosterone levels not to differ between women with and without PMS (Bäckström and Aakvaag, 1981; Dougherty et al., 1997a; Watts et al., 1985). Moreover, research assessing testosterone levels in these groups of women in relation to serotonin activity is warranted.

PMS sufferers were also found to exhibit higher levels of all proinflammatory cytokines that were measured than the normally cycling women during both the follicular and luteal phases, with these differences reaching statistical significance for IL-6, IL-8 and TNF-α. Increased proinflammatory cytokine production has been associated with depression (Maes et al., 1995; Schlatter et al., 2001), anxiety (Connor et al., 1998), fatigue (Kronfol and Remick, 2000), headaches (Martelletti et al., 1993) and sleep disturbances (Kapas and Krueger, 1992; Vgontzas et al., 2005), all of which are symptoms characteristic of PMS. Moreover, increasing proinflammatory cytokine production through endotoxin administration induces anxiety, depression and decreased memory performance (Reichenberg et al., 2001), leading to the proposal that the up-regulation of proinflammatory cytokines may be involved in the production of PMS symptoms (Konecna et al., 2000). Cytokines have been shown to interact with the serotonergic system (Gemma et al., 1997; Linthorst et al., 1994; Pousset et al., 1996; Silverman et al., 1989). SSRIs, an effective PMS treatment (Dimmock et al., 2000; Eriksson et al., 2002; Yonkers et al., 2008), can normalise cytokine levels (Maes, 2001; Sluzewska et al., 1995; Van Gool et al., 2003). Therefore, it is possible that cytokines interact with the serotonergic system to produce PMS symptoms, in a manner similar to that suggested for the aetiology of
depression (Bonaccorso et al., 2002; Menkes and MacDonald, 2000). Although the results from this study suggest that increased proinflammatory cytokine production may be involved in the pathophysiology of PMS, no other studies have examined cytokine levels in this population. There is a need for further research to prospectively examine cytokine production in PMS sufferers and normally cycling women. Moreover, research assessing proinflammatory cytokine production in relation to the serotonergic system is warranted.

When the PMS non-reporters and controls were analysed as separate groups, they were shown to exhibit similar hormone (FSH, LH, oestradiol, progesterone, prolactin and testosterone) and cytokine (IL-1β, IL-6, IL-8, IFN-γ, TNF-α) concentrations to each other during the follicular and luteal phases. This resulted in both the PMS non-reporters and controls exhibiting lower testosterone and pro-inflammatory cytokine (IL-8 and TNF-α) levels than the PMS sufferers. As the PMS non-reporters displayed similar physiological profiles to the controls, which were different to those of the PMS sufferers, this further suggests that the women belonging to the PMS non-reporter group should not be considered to be PMS sufferers, but rather be considered to experience ‘normal cyclicity’ and be included in ‘control’ groups in research settings (see sections 3.5.2.3 and 5.5.3.1).

5.6 Summary
Variable profiles of aggression and impulsivity across the menstrual cycle should be considered ‘normal.’ ‘Abnormality,’ or caseness with respect to PMS diagnosis could be said to arise when the follicular to luteal symptom increase, and the premenstrual symptom severity, exceed a certain level. This pattern of results can become disguised if researchers do not carefully consider the diagnostic methods they have used when forming their diagnostic groups. The use of liberal PMS criteria can divide normally cycling women into two groups, and confirm a PMS diagnosis in a large proportion of women who do not believe they experience problematic PMS symptoms (PMS non-reporters). These women should not be regarded as PMS sufferers, but rather be considered to experience ‘normal’ cyclicity and included in ‘control’ groups in research settings. Comparing women who do and do not meet PMS criteria can result in the exclusion of PMS non-reporters from ‘control’ groups. Moreover, retrospective group allocation can result in the inclusion of PMS non-reporters in PMS groups. Furthermore, failure to identify co-morbidities can result in women with co-morbidities also being considered PMS sufferers. This
heterogeneity in the composition of control and PMS samples in research settings makes it difficult to draw meaningful comparisons between PMS studies. Researchers need to give careful consideration to how they form their diagnostic groups, taking into account the diagnostic criteria they have used, and to describe that classification system clearly.
Chapter 6: Dietary supplements and herbal remedies for Premenstrual Syndrome (PMS): A Systematic Research Review of the Evidence for their Efficacy

6.1 Introduction

Approximately three quarters of regularly menstruating women experience some physical and psychological symptoms of PMS (Burt and Stein, 2002; Haywood et al., 2002; Hylan et al., 1999; Ramcharan et al., 1992; Steiner and Wilkins, 1996; see also section 1.3), and these symptoms result in significant impairment in one or more areas of daily life for around 3-8% of women (Burt and Stein, 2002; Endicott et al., 1999; Halbreich et al., 2007; Haywood et al., 2002; Johnson, 1987; Pearlstein et al., 2005; Steiner and Born, 2000). As PMS symptoms usually last until menopause (Dimmock et al., 2000; Johnson, 2004), treatment for this condition must be effective, cost efficient and safe to use over a long time period (Johnson, 2004). Numerous treatments have been recommended for PMS (Bryant and Dye, 2004; Chakmakjan, 1983; Ismail and O’Brien, 2001; O’Brien, 1993).

Many advocate lifestyle changes for women with mild symptoms, such as exercise, relaxation and nutritional changes, including reduced caffeine, alcohol and unrefined sugar intake (Bryant and Dye, 2004; Connolly, 2001; Johnson, 2004). However, the studies that have assessed the efficacy of these interventions have been poorly controlled (Connolly, 2001; Johnson, 2004; Kirkby and Lindner, 1998). Bryant and Dye (2004) performed a review of the nutritional interventions used by PMS sufferers. However, although this covered a range of treatments, it was not systematic.

A variety of drugs are prescribed for women with more severe symptoms (Bendich, 2000; Eriksson et al., 2002; Yonkers et al., 2008). Although the aetiology of PMS remains unclear (Ismail and O’Brien, 2001; Halbreich, 2003; O’Brien, 1993), women with PMS appear more sensitive to normal cyclical fluctuations in steroid hormones (see section 1.5), which influence neurotransmitter function in the central nervous system (Eriksson et al., 2002; Rapkin, 1992; Reid and Yen, 1981; Steiner et al., 1997a). Therefore, medical treatment for PMS has focused on removing the ovarian steroid trigger through ovulation suppression (Eriksson et al., 2002; O’Brien, 1993; Yonkers et al., 2008) and on correcting

---

1 The publication and presentations listed at the beginning of this thesis were based on the work presented in this chapter.
neurotransmitter function through the administration of selective serotonin reuptake inhibitors, SSRIs (Connolly, 2001; Johnson, 2004).

Ovulation can be suppressed through the administration of drugs such as Danazol and gonadotropin-releasing hormone agonists (Connolly, 2001). Although this often effectively reduces many PMS symptoms (Eriksson et al., 2002; Yonkers et al., 2008), menopause is induced (Connolly, 2001; Yonkers et al., 2008). The administration of progesterone and oestrogen can surmount this effect, but can also result in the re-emergence of PMS symptoms (Connolly, 2001). Danazol has been shown to result in significant side effects (Johnson, 2004), while GRH agonists are expensive (Johnson, 2004), and cannot be used over a long time period, as they increase risk of osteoporosis (Connolly, 2001) and atherosclerotic cardiovascular disease (O’Brien, 1993). Moreover, some women with severe symptoms do not respond well to these drugs (Rapkin, 2003). Ovariectomy also suppresses ovulation (O’Brien, 1993) and reduces PMS symptoms (Connolly, 2001; Yonkers et al., 2008). However, this is an extremely drastic measure to relieve symptoms of PMS, and is associated with other symptoms, such as early menopause (Yonkers et al., 2008).

Serotonin transmission is facilitated through the administration of SSRIs, which have been shown to effectively reduce both physical and psychological PMS symptoms (Johnson, 2004; Steiner et al., 2003b; Wyatt et al., 2002; 2004). Fluoxetine (Miner et al., 2002; Pearlstein et al., 1997; Rapkin, 2003; Steiner et al., 1997b), citalopram (Wikander et al., 1998), paroxetine (Eriksson et al., 1995), sertraline (Freeman et al., 1999; Yonkers et al., 1997) and clomipramine (Sundblad et al., 1993) have all been shown to be effective PMS treatments. However, SSRIs can cause side effects, including gastrointestinal disturbances, insomnia, fatigue, headache, dry mouth, dizziness, tremor and sweating (Johnson, 2004). Intermittent SSRI treatment can be used (Connolly, 2001; Johnson, 2004; Sundblad et al., 1993), and this reduces the likelihood of side effects occurring (Johnson, 2004). This form of treatment is not as effective as continuous treatment for somatic symptoms (Connolly, 2001; Yonkers et al., 2008), and can result in side effects re-emerging each cycle (Eriksson et al., 2002). Moreover, long term use of SSRIs can result in decreased libido and delayed orgasm (Eriksson et al., 2002). These adverse effects often persist until treatment is ended (Yonkers et al., 2008).
Therefore, unless symptoms are severe, many women are reluctant to take these forms of treatment, while women planning a pregnancy are often unable to take them (Bendich, 2000). In contrast, less invasive lifestyle changes often do not benefit moderate to severe symptoms sufficiently (Bendich, 2000). Therefore, many women with PMS turn to dietary supplements and herbal remedies to treat their symptoms (Bendich, 2000; Domoney et al., 2003; Eriksson et al., 2002; Fugh-Berman and Kronenberg, 2003; Girman et al., 2003). Many women with PMS report that they find their doctors unhelpful (Domoney et al., 2003; O’Brien, 1993) and would rather self-medicate with over the counter medication (Domoney et al., 2003). Numerous dietary supplements and herbal remedies have been proposed to relieve PMS symptoms. However, rigorous scientific studies to test their efficacy are lacking (Bendich, 2000; Domoney et al., 2003; Eriksson et al., 2002).

This chapter presents a systematic review of dietary supplements and herbal remedies commonly used for PMS, including calcium, magnesium, vitamin B6, evening primrose oil, *Vitex agnus castus*, ginkgo biloba and St John’s Wort.

6.2 Method

6.2.1 Selection Criteria

For review inclusion, studies were required to include a placebo/comparison group, as the placebo effect has been shown to be large in women with PMS (Dimmock et al., 2000; Eriksson et al., 1995; Freeman and Rickels, 1999; Magos et al., 1986; Steiner et al., 1995; Yonkers et al., 1997; see also section 7.1.6). They were required to test the efficacy of one treatment, taken for at least one cycle. Treatment could occur throughout the cycle or in the luteal phase only. Participants had to be randomised to treatments in the case of parallel designs, or have treatment orders counterbalanced in the case of crossover designs.

Studies required participants of reproductive age, with PMS or PMDD, diagnosed prospectively or retrospectively. Studies employing outcome measures which examined combined PMS symptoms, global scores or specific symptoms e.g. cyclical breast pain were included. Few trials employed prospective diagnosis or assessment of efficacy hence, in order to provide a comprehensive review, retrospective measures were included. Trials including women taking oral contraceptives were also included.
6.2.2 Search Strategy

Trials were identified by searching Embase (1980 to 2008), Medline (1966 to 2008), AMED (1985-2008), Cinahl (1982 to 2008, PsycINFO (1967-2008), and the Cochrane Controlled Trials Register database.

A general search on these databases revealed dietary supplements and herbal remedies used for PMS, including vitamin B6, magnesium, calcium, *Vitex agnus castus*, evening primrose oil, St John’s Wort and ginkgo biloba supplements. Databases were searched using all Latin and English names for these supplements. Hence, the following keywords were used:

- Pms, pmt, pmdd, llpd, llpdd, premenstrual syndrome, pre menstrual syndrome, pre-menstrual syndrome, premenstrual tension, pre menstrual tension, pre-menstrual tension, premenstrual dysphoria, pre-menstrual dysphoria, pre menstrual dysphoria, premenstrual dysphoric disorder, pre-menstrual dysphoric disorder, pre menstrual dysphoric disorder, late luteal
- *Vitex, agnus castus, vitex agnus castus, vitex agnus-castus, chasteberry, castufemin, cefanorm, femicur, gynocastus, hewekliman, kytta-femin, strotan, agnomens*
- Evening primrose oil, oenothera biennis, evening primrose, primrose oil, oenothera, biennis, fever plant, oep, sundrop, essential fatty acids, efamol
- Calcium, calcium supplements, calcium therapy
- Magnesium, magnesium therapy
- Vitamin B6, vitamin B6, vitamin b-6, vitamin B-6, vitamin therapy, vitamins, pyridoxine, B-vitamins, B vitamins, pyridoxine hydrochloride
- St john's wort, st johns wort, hypericum perforatum, hypericum, perforatum, hypericin, hypericins, kira
- Gingko, gingko biloba, ginko, ginko biloba, biloba, living fossil, Japanese Silver Apricot, Kew Tree, Maidenhair Tree, Yinsing
- Alternative therapy, alternative therapies, herbal therapy, nutritional supplements
There were 119 articles remaining when duplicates were removed, with 38 articles kept as trials relevant to the research area. Nine articles did not meet the selection criteria. Seven had no placebo or comparison group (Berger et al., 2000; Brush et al., 1988; Larsson et al., 1989; Loch et al., 2000; Pal et al., 2003; Prilepskaya et al., 2006; Stevinson and Ernst, 2000). Two articles were excluded as they tested the efficacy of more than one treatment. Krutan Berman et al. (1990) tested the efficacy of pyridoxine for PMS, in combination with a dietary intervention, while Callender et al. (1988) tested evening primrose oil using efamol tablets which also included efavit (containing vitamin C, pyridoxine, niacin, and zinc sulfate). Hence, 29 studies meeting the selection criteria were retained (See tables 1-7). One additional article (Ockerman et al., 1986) was identified from the reference lists of review articles in this area.

6.3 Data synthesis for the dietary supplements and herbal remedies for PMS

Two studies of calcium, four of magnesium, twelve of vitamin B6, four of evening primrose oil, four of Agnus castus, three of St. John’s Wort and one of gingko biloba met the inclusion criteria for the review. Nineteen trials suggested some benefit of the treatment under investigation, while 11 trials found no such benefit.

Table 6.1 describes and evaluates these studies. The trials are ordered alphabetically and sub-divided into studies finding positive and negative effects for each treatment. Aspects of the methodological quality (e.g. sample size, design, dose and duration of intervention, screening and assessment tools employed) are considered in order to provide a context for discussion of the reliability of the results. The table assumes that studies excluded women taking the oral contraceptive pill, unless otherwise stated.
<table>
<thead>
<tr>
<th>Calcium (Ca)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reference</strong></td>
</tr>
<tr>
<td>Thys-Jacobs et al. (1989)</td>
</tr>
<tr>
<td>Thys-Jacobs et al. (1998)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Magnesium (Mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reference</strong></td>
</tr>
<tr>
<td>Facchinetti et al. (1991)</td>
</tr>
<tr>
<td>Walker et al. (1998)</td>
</tr>
<tr>
<td>De Souza (2000)</td>
</tr>
<tr>
<td>Walker et al. (2002)</td>
</tr>
<tr>
<td>Reference</td>
</tr>
<tr>
<td>----------------------------</td>
</tr>
<tr>
<td>Abraham &amp; Hargrove (1980)</td>
</tr>
<tr>
<td>Barr (1984)</td>
</tr>
<tr>
<td>Doll et al. (1989)</td>
</tr>
<tr>
<td>Kashanian et al. (2006)</td>
</tr>
<tr>
<td>Kendall &amp; Schnurr (1987)</td>
</tr>
<tr>
<td>Sharma et al. (2007)</td>
</tr>
<tr>
<td>Williams et al. (1985)</td>
</tr>
<tr>
<td>Reference</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>Diegoli et al.</td>
</tr>
<tr>
<td>Hagen et al.</td>
</tr>
<tr>
<td>Malmgren et al.</td>
</tr>
<tr>
<td>Smallwood et al.</td>
</tr>
<tr>
<td>Stokes &amp; Mendels,</td>
</tr>
<tr>
<td>Reference</td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>Ockerman et al. (1986)</td>
</tr>
<tr>
<td>Puolakka et al. (1985)</td>
</tr>
<tr>
<td>Collins et al. (1993)</td>
</tr>
<tr>
<td>Khoo et al. (1990)</td>
</tr>
</tbody>
</table>

**Vitis agnus castus (AC)**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample Size</th>
<th>Design</th>
<th>Dose and Duration</th>
<th>Diagnosis</th>
<th>Assessment Measures</th>
<th>Outcome</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atmaca et al. (2003).</td>
<td>N=41; 38 completed; 7% dropout (fluoxetine, 21; AC, 20) PMDD</td>
<td>R, SB PGs PC</td>
<td>Fluoxetine or AC (20-40mg/d) for 2 cycles</td>
<td>DSM-IV; Penn daily symptom reports for 2 cycles</td>
<td>DSM-IV criteria for PMDD; premenstrual score of the DSR, HAM-D (Hamilton, 1960), CGI-I; agreement of improvement by the authors</td>
<td>✓ No significant difference between groups on DSR, CGI-SI scores or responder rates</td>
<td>AC compared to fluoxetine. Participants and raters blinded but prescribing physician not.</td>
</tr>
<tr>
<td>Halaska et al. (1998)</td>
<td>N=100, Completion: AC 48, placebo 49 3% dropout</td>
<td>DB, PGs PC</td>
<td>2 x 30 AC drops (1.8ml)/d for 3 cycles</td>
<td>Not specified</td>
<td>VAS scale</td>
<td>✓ breast pain</td>
<td>Only women suffering from mastalgia &gt; 5 days/cycle included.</td>
</tr>
</tbody>
</table>
### Vitex agnus castus (AC) continued.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample Size</th>
<th>Design</th>
<th>Dose and Duration</th>
<th>Diagnosis</th>
<th>Assessment Measures</th>
<th>Outcome</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schellenberg (2001)</td>
<td>N=170, active, 86; placebo 84</td>
<td>R, DB PGs, PC</td>
<td>1 x 20mg AC tablet/d for 3 cycles</td>
<td>DSM-III-R</td>
<td>6 symptoms (3 somatic, 3 mood) on a VAS compared to previous 3 cycles; responder rate</td>
<td>✓ combined and individual symptoms (not bloating); responder rates- 52% AC v 24% placebo</td>
<td>Assessment scale and timing of symptom rating was unclear. OC users included.</td>
</tr>
<tr>
<td>Turner &amp; Mills (1993)</td>
<td>N=600, Completion: AC, 105; placebo 112 64% dropout</td>
<td>R, DB PGs, PC</td>
<td>Treatment for 3 mths</td>
<td>MDQ</td>
<td>MDQ at end of treatment; a shortened version administered at the end of cycles 1 and 2</td>
<td>X except for the symptom ‘feel jittery or restless’ and responder rates = 25%AC v 16% placebo</td>
<td>Dose and frequency of administration was unclear. Only women high in negative affect included. Huge drop out rate, but evenly distributed across groups.</td>
</tr>
</tbody>
</table>

### St John’s Wort (SJW)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample Size</th>
<th>Design</th>
<th>Dose and Duration</th>
<th>Diagnosis</th>
<th>Assessment Measures</th>
<th>Outcome</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hicks et al. (2004)</td>
<td>N=169, 125 completed; 26% dropout</td>
<td>R, DB PGs, PC-lactose/cellulose</td>
<td>2 x 300mg tablets standardized to 900µg hypericin/d for 2 cycles</td>
<td>Retrospective assessment; 25 symptoms rated daily for 1 cycle</td>
<td>VAS assessing 25 symptoms grouped into 6 categories (anxiety, craving, depression, hydration, other and total)</td>
<td>X all symptom subgroups</td>
<td>Although diagnosis was prospectively confirmed, this process was unclear.</td>
</tr>
<tr>
<td>Pakgohar et al. (2005)</td>
<td>N=70 (35 in each group)</td>
<td>R, DB PGs, PC</td>
<td>2 x 30 drops for 7 days prior to menses</td>
<td>Daily ratings for two cycles</td>
<td>Not specified</td>
<td>✓ symptom reduction</td>
<td>Persian article. Method of diagnosis and assessment of efficacy not specified. Not clear which symptoms benefited.</td>
</tr>
</tbody>
</table>

### Ginko Biloba

<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample Size</th>
<th>Design</th>
<th>Dose and Duration</th>
<th>Diagnosis</th>
<th>Assessment Measures</th>
<th>Outcome</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tamborini &amp; Taurelle (1993)</td>
<td>N=165 Congestive PMS symptoms</td>
<td>R, DB PGs, PC-MC</td>
<td>EGB 761 from cycle day 16 for 2 cycles</td>
<td>Observation of 1 menstrual cycle</td>
<td>Daily scale measuring congestion, breast tenderness and mood; Practitioner observation premenstrually before and after treatment</td>
<td>✓ congestive symptoms, particularly breast symptoms</td>
<td>Method of diagnosis not specified.</td>
</tr>
</tbody>
</table>

**KEY:** AC = agnus castus; B = blind; B6 = vitamin B6; Ca = calcium; CaCO3 = calcium carbonate; CGI-I = Clinical global Impression scale-severity of illness; CO = crossover; DB = double-blind; DSR = Daily Symptom Report (Freeman et al., 1996); EPO = evening primrose oil; GI = Global Improvement; HADS = Hospital Anxiety and Depression Scale (Zigmond and Snaith, 1983); HAM-D = Hamilton depression rating scale (Hamilton, 1960); HCI = hydrochloride; MC = multi-centred; MDQ = Menstrual Distress Questionnaire (Moos, 1968); Mg = magnesium; MgO = magnesium oxide; MHQ = Mental Health Questionnaire (Warner and Bancroft, 1990); Microg = microgram; MSQ = Menstrual Symptomatology Questionnaire (Abraham, 1980); OC = oral contraceptive; PAF = Premenstrual Assessment Form (Halbreich et al., 1982); PC = placebo controlled; PGs = parallel groups; R = randomised; SAS-M = Social adjustment scale (Cooper et al., 1982); SB = single-blind; SJW = St John’s Wort; STAI = State-Trait Anxiety Inventory (Spielberger, 1970); VAS = Visual analogue scale
6.3.1 Calcium

Two well designed trials rigorously assessed the efficacy of calcium for PMS (Thys-Jacobs et al., 1989; Thys-Jacobs et al., 1998). Both of these demonstrated a benefit of calcium supplementation for PMS symptoms. Thys-Jacobs et al. (1989) found calcium to benefit the symptoms of negative affect, water retention and pain. Retrospective analysis showed that 73% of women on calcium reported fewer premenstrual symptoms than the placebo group. Thys-Jacobs et al. (1998) confirmed these findings with a larger sample, and also found calcium to benefit food cravings. By the end of the second treatment cycle, all individual symptoms, except fatigue and insomnia, showed a significant response to calcium. By the end of the third cycle, a 48% reduction in combined symptom scores was found in the calcium group, compared to 30% in the placebo group.

Though diagnosis in both studies involved prospective measurement, Thys-Jacobs et al. (1989) only used one screening cycle for this purpose and supplied all diaries for this cycle together, which may have allowed women to identify a pattern in their symptom reports. Nevertheless, these studies were carefully designed, and the similarities between their findings suggest that calcium supplementation of at least three cycles may be of benefit to women suffering from PMS.

6.3.2 Vitamin B6 (pyridoxine)

A systematic review to evaluate the efficacy of vitamin B6 for the treatment of PMS identified 25 published trials and included nine trials, representing 940 patients (Wyatt et al., 1999). The overall odds ratio relative to placebo for efficacy of B6 was 2.32 (95% confidence interval 1.95 to 2.54). Wyatt et al. suggested that doses of vitamin B6 up to 100mg/day may benefit premenstrual depression and other symptoms. However, conclusions from the systematic review were limited by the methodological weaknesses of some of the trials included. Findings from studies assessing vitamin B6 in the current review are also mixed. Positive effects of B6 were found in seven studies (Abraham and Hargrove, 1980; Barr, 1984; Doll et al., 1989; Kashanian et al., 2006; Kendall and Schnurr, 1987; Sharma et al., 2007; Williams et al., 1985), while five studies failed to demonstrate an effect (Diegoli et al., 1998; Hagen et al., 1985; Malmgren et al., 1987; Smallwood et al., 1986; Stokes and Mendels, 1972). Those studies which demonstrated benefits appeared to be methodologically stronger, although some limitations were still apparent.
Williams et al. (1985) performed a large study of 434 women. Although no benefit was found for individual symptoms, after three cycles women in the pyridoxine group showed an 82% improvement relative to symptoms at the start of the study, while the placebo group only experienced a 70% improvement. However, PMS was diagnosed if one or more symptom in the cycle was relieved by menstruation and ‘other medication’ could be taken throughout the study. Furthermore, many women on both pyridoxine and placebo changed their dose, usually by increasing it, which seemed to lead to greater improvement in each group.

Six small-scale studies showed benefits of vitamin B6. Treatment with B6 improved total symptom score (Abraham and Hargrove, 1980; Barr, 1984; Sharma et al., 2007); psychiatric symptoms (Kashanian et al., 2006); emotional symptoms, including depression, irritability and tiredness (Doll et al., 1989); autonomic reactions, including dizziness and vomiting, and behavioural changes including poor performance and decreased social activity (Kendall and Schnurr, 1987). However, it was unclear when women rated symptoms for diagnosis in the Abraham and Hargrove (1980) study, as women completed the MDQ for 6 cycles, and treatment (3 month vitamin B6 and 3 months placebo) lasted this long. Sharma et al. (2007) did not discuss blinding procedures. Barr (1984) failed to specify how women were selected for the trial, though women were examined for physical causes for their symptoms, and excluded if any were found. Barr (1984) also assessed efficacy with an 8 symptom diary card that she herself described as ‘notoriously inaccurate in assessing results’ (p.425).

Smallwood et al. (1986) and Hagen et al. (1985) found no benefit of vitamin B6. Although there was a trend towards a reduction in breast tenderness compared to placebo treatment in the study by Smallwood et al. (1986), this did not reach statistical significance. However, only a small proportion of PMS symptoms were assessed and only women having symptoms of severe pain in the premenstrual half of the cycle for a minimum of six consecutive cycles were included. Treatment order in the Hagen et al. (1985) study was not fully counter-balanced, such that placebo was the second treatment for the majority of women. They found that women seemed to prefer their 2nd treatment, irrespective of what this was – for the majority this was placebo. Women were diagnosed via interview, and were only required to report regularly occurring problems commonly associated with PMS. Moreover, treatment efficacy was only retrospectively assessed at the end of treatment. Other studies which used intermittent
Pyridoxine treatment also reported negative findings (Diegoli et al., 1998; Stokes and Mendels., 1972; Malmgren et al., 1987).

Therefore, some evidence suggests that continuous vitamin B6 treatment at doses of 50 to 150 mg/day may be beneficial for some PMS symptoms, since the intermittent treatment (at 50 to 300mg/day) did not prove effective. However, more trials of better quality, using stricter diagnostic criteria, are required to confirm its benefit.

6.3.3 Magnesium

Trials assessing magnesium supplementation produced mixed findings. Both Walker et al. (1998) and De Souza (2000) found no beneficial magnesium effect compared with placebo for any symptom group measured after one treatment cycle. However, Walker et al. (1998) found magnesium to benefit symptoms of weight gain, swelling, breast tenderness and abdominal bloating after two cycles. Facchinetti et al. (1991) also found intermittent magnesium treatment to benefit total symptom scores, and specifically symptoms of negative affect and arousal after two treatment cycles. However, Walker et al. (2002) unexpectedly found their placebo (sorbital) group to have significantly reduced anxiety-related and total premenstrual symptoms after two cycles compared with the magnesium groups.

Many methodological limitations were apparent in these studies. All studies used relatively small samples. Only Facchinetti et al. (1991) excluded women taking oral contraceptives. Walker et al. (1998) confirmed PMS retrospectively and did not take any baseline measurements, and neither De Souza (2000) nor Walker et al. (1998) had a washout cycle between treatments. Therefore, trials with longer treatment durations, tighter controls and larger samples are required to evaluate Mg supplementation in PMS.

6.3.4 Evening Primrose Oil

The most methodologically sound study (Collins et al., 1993) to assess evening primrose oil found no benefit for mood or physical symptoms. Interestingly, women who stayed in the study the longest showed greatest improvement, suggesting a strong placebo or expectancy effect. Khoo et al. (1990) found no difference between their treatment and placebo groups in total PMS scores, or in psychological, fluid retention, breast or menstrual symptoms. It was not clear when symptoms were recorded for
diagnosis or treatment evaluation during the cycle, and it is possible that this was done retrospectively.

Two studies showed benefits of treatment with evening primrose oil. Puolakka et al. (1985) found benefits to total PMS score, and specifically depression, compared with placebo treatment. However, they did not specify the nature of their placebo treatment or if the study was double-blind. Ockerman et al. (1986) found 50% of their efamol group experienced moderate to complete relief of symptoms, compared with only one case in the placebo group. However, this short report failed to specify the method of diagnosis, the symptoms measured, or the measures used to assess treatment efficacy. Methods of randomisation were not discussed, and only women “resistant” to other therapies were included in the trial. Though some studies suggest evening primrose oil may benefit PMS symptoms (Ockerman et al., 1986; Puolakka et al., 1985), currently the evidence is not convincing.

6.3.5 Vitex agnus castus (Chasteberry)

Most trials evaluating Agnus castus treatment reported positive effects. Halaska et al. (1998) found reduction of cyclical breast pain, as did Schellenberg (2001) who also found positive results for combined PMS symptoms and specific symptoms of irritability, mood alteration, anger and headaches. Atmaca et al. (2003) compared Agnus castus treatment with fluoxetine, an effective PMS medication (e.g. Miner et al., 2002; Pearlstein et al., 1997; Steiner et al., 1997b) and found both treatments reduced total symptoms and severity of illness scores, but the treatment groups did not differ statistically. Both Atmaca et al. (2003) and Schellenberg (2001) noted an improvement in responder rates with Agnus castus. Atmaca et al. (2003) found Agnus castus to improve physical symptoms, including swelling, food cravings and cramps whereas fluoxetine improved psychological symptoms including depression, insomnia, nervous tension and feeling out of control. However, no placebo group was employed, making it impossible to separate pharmacological and placebo effects.

Turner and Mills (1993) report the only trial with negative findings. Although the symptom ‘feel jittery or restless’ was reduced in the Agnus castus group, no other benefits were found. However, there was some suggestion that Agnus castus might benefit water retention, as the difference between the Agnus castus and placebo groups
approached statistical significance (p=0.09), and more women on treatment than on placebo commented that their PMS symptoms had generally improved (25% v 16%).

Methodological limitations were apparent in these studies in terms of retrospective diagnosis (Turner and Mills, 1993) and evaluation of treatment efficacy (Schellenberg, 2001; Turner and Mills, 1993), lack of representative samples (Halaska et al., 1998; Turner and Mills, 1993) and double-blind procedures (Atmaca et al., 2003). Moreover, some authors failed to specify the dose and frequency of treatment used (Turner and Mills, 1993). Though some studies suggest *Agnus castus* may benefit PMS symptoms, trials with stricter controls and representative samples are needed.

### 6.3.6 *Hypericum perforatum* (St John’s Wort)

The three trials evaluating St John’s Wort reported positive effects (Hicks et al., 2004; Pakgohar et al., 2004; 2005). However, both trials reported by Pakgohar et al. were published in Persian, with only the abstracts available in English. Many details were omitted from the abstracts e.g. sample size (Pakgohar et al., 2004), the methods used for diagnosis and assessment of treatment efficacy (Pakgohar et al., 2004; 2005), and the type of symptoms that improved (Pakgohar et al., 2005). Hicks et al. (2004) found St John’s Wort to benefit all symptom subgroups that they studied in comparison to placebo treatment. However, these differences did not reach statistical significance. The authors suggested that their study had insufficient statistical power to demonstrate the efficacy of the dose they used (600mg, standardised to 0.3% of hypericin)/day). Most studies have used a dose of 900mg/day standardized to 0.3% hypericin for depression (Montgomery et al., 2000; Shelton et al., 2001; Szegedi et al., 2005), and this is the dose recommended for depression by most manufacturers of St John’s Wort. There is currently no recommended dose for PMS. Therefore, although some studies suggest St John’s Wort may be beneficial for PMS sufferers, more rigorously controlled studies using larger doses are needed.

### 6.3.7 *Gingko biloba*

Only one study has been performed to assess gingko biloba for PMS. Tamborini and Taurelle (1993) found gingko biloba to benefit congestive PMS symptoms, particularly breast tenderness. However, the sample of PMS sufferers studied was atypical in that women were required to report congestive premenstrual symptoms for 7 days per cycle for the 3 cycles prior to recruitment. Therefore, more studies using representative
samples are needed to determine whether gingko biloba may have a role in PMS treatment.

6.4 Discussion

Many women with PMS turn to dietary supplements and herbal remedies to treat their symptoms (Bendich, 2000; Domoney et al., 2003; Eriksson et al., 2002; Fugh-Berman and Kronenberg, 2003; Girman et al., 2003), despite the lack of established efficacy (Bendich, 2000; Domoney et al., 2003; Eriksson et al., 2002). This review included 30 trials that assessed the efficacy of seven different dietary supplements and herbal remedies for PMS. The most substantial positive evidence was found for calcium and continuous vitamin B6 treatment. Trials assessing magnesium and evening primrose oil produced conflicting findings, whilst insufficient data was found to advocate the use of vitex agnus castus, gingko biloba or St John’s Wort.

The effectiveness of vitamin B6 and calcium may shed light on the physiological processes involved in the aetiology, development or maintenance of PMS. There is some evidence that calcium levels vary across the cycle, with levels reducing premenstrually (Okey et al., 1930; Pandya et al., 1995), although evidence is not consistent (e.g. Gray et al., 1982). Women with PMS have been shown to have lower calcium levels than symptom-free women (Shamberger, 2003; Thys-Jacobs et al., 2007). Many similarities exist between PMS symptoms and hypocalcaemia, including irritability, fatigue, anxiety, depression, impaired memory and muscle cramps (Thys-Jacobs, 2000; Thys-Jacobs et al., 2007). Moreover, women with PMS have been shown to have a lower bone mass than symptom-free women (Thys-Jacobs et al., 1995). The effective use of calcium supplementation to reduce PMS symptoms (Thys-Jacobs et al. 1989, 1998) has been used to support the theory that PMS may be related to calcium deficiency (Shamberger, 2003; Thys-Jacobs et al., 1989; Thys-Jacobs and Alvir, 1995).

Pyridoxal phosphate, the active form of vitamin B6, is an essential cofactor in the synthesis of serotonin and dopamine in the hypothalamus (Brush et al., 1988; Connolly, 2001). Increasing evidence suggests that serotonin is important in the pathogenesis of PMDD (Eriksson et al., 2002; Halbreich, 1997; Rapkin, 1992; Steiner, 1997a). Women with PMS have been shown to have lower serotonin levels premenstrually (Ashby et al., 1988; Rapkin et al., 1987; Rasgon et al., 2000; Taylor et al., 1984). Many symptoms characteristic of PMS have been linked to serotonergic dysfunction (Bond et al., 2003;
Cleare and Bond 1995; Ho et al., 2001; Kahn et al., 1888; Russo et al., 2005; Steinberg et al., 1999). Selective serotonin reuptake inhibitors (SSRIs), which facilitate serotonergic transmission, effectively reduce PMS symptoms (Dimmock et al., 2000; Steiner et al., 2003; Szegedi et al., 2005). Therefore, supplementing women with vitamin B6 may reduce PMS symptoms by increasing serotonin levels.

Alternatively, increasing the synthesis of serotonin and dopamine in the hypothalamus may reduce PMS symptoms via reduction of prolactin (PRL) levels (Delitala et al., 1976; Kleijnen et al., 1990). *Agnus castus* treatment has also been proposed to alleviate PMS symptoms by this mechanism (Merz et al., 1996; Sliutz et al., 1993). *Vitex agnus castus* extracts and lisuride, a synthetic dopamine agonist, have been shown to inhibit PRL secretion in rat pituitary cells, an effect prevented by administering a dopamine receptor blocker (Sliutz et al., 1993). Although this mechanism appears plausible, the evidence to date for *Agnus castus* as an effective PMS treatment is lacking.

The studies considered in this review differed greatly in the diagnostic methods they used. It is generally accepted that prospective daily self-report measures are needed to confirm PMS (Connolly, 2001; Freeman, 2003; Johnson, 2004; Steiner and Wilkins, 1996). Some studies diagnosed PMS by using the DSM criteria, and confirmed their diagnoses prospectively. However, others relied upon retrospective diagnosis, which has been criticized (Ainscough, 1990; Connolly, 2001; Richardson, 1990), since this often results in inflated estimates of symptom severity (De Souza et al., 2000; see also section 1.3.1.3).

The methods used for the assessment of treatment efficacy also differed. Many studies used the total symptom score of a rating scale as their primary outcome measure, and simultaneously considered symptom clusters. Some also considered individual symptoms. This increases the chances of finding symptom effects. Assessment measures were used prospectively with daily symptom ratings in some trials. Others assessed treatments retrospectively, at the end of each cycle or at the end of treatment, using a variety of methods, including questionnaires, interviews (Loch et al., 2000; Kendall and Schnurr, 1987) and general practitioner assessments (Williams et al., 1985; Smallwood et al., 1986). Some authors did not specify their treatment assessment methods (Ockerman et al., 1986; Malmgren et al., 1987).
Women taking the oral contraceptive, which has previously been used as a PMS treatment (Girman et al., 2003), were not excluded from some studies. Other studies focused on specific groups of women, including women with ‘premenstrual tension depression’ (Stokes and Mendels, 1972), severe cyclical mastalgia (Smallwood et al., 1986) and congestive PMS symptoms (Tamborini and Taurelle, 1993). It is difficult to compare such studies with those examining samples with a range of PMS symptoms.

6.5 Conclusion
Randomised controlled trials of magnesium and evening primrose oil produced conflicting results, in contrast to the positive evidence for the efficacy of calcium and continual vitamin B6 treatment. There are currently insufficient data to advocate the use of ginkgo biloba, Vitex agnus castus and St John’s Wort, although preliminary data seems supportive. More consensus about the diagnosis, measurement and assessment of PMS is required. Randomised, double-blind, placebo-controlled trials using larger, representative samples, strict, prospectively confirmed diagnostic criteria and assessment of treatment efficacy would help to clarify the role of the many dietary supplements and herbal remedies that are currently used by PMS sufferers. Although much of the clinical research is preliminary and/or inadequately controlled, this review will be relevant to the practicing clinician looking for greater understanding of the alternative therapies available to their patients with PMS.
7.1 Introduction

7.1.1 St John’s Wort (SJW)

St John’s Wort (Hypericum perforatum) is a flowering plant that has been used in Europe for centuries to treat conditions such as neuralgia, neuroses, insomnia, anxiety and depression (Gaster and Holroyd., 2000; Linde et al., 2005; Nangia et al., 2000; Rasmussen, 1998). SJW is a popular antidepressant in many European countries and is available on prescription in Germany for the treatment of depression, anxiety, seasonal affective and sleep disorders (Bennett et al., 1998; Linde et al., 2005; Nangia et al., 2000).

7.1.2 Antidepressant properties of SJW

The antidepressant properties of SJW have received most scientific attention (Rasmussen, 1998). Several randomised controlled trials have been conducted which have compared the effectiveness of SJW with placebo or antidepressant treatment for depressive disorders (Seelinger and Mannel, 2007). A number of systematic reviews (Clement et al., 2006; Ernst, 1995; Gaster and Holroyd, 2000; Linde et al., 2005; Pilkington et al., 2006; Volz and Laux, 2000) and meta-analyses (Kim et al., 1999; Linde et al., 1996; Wernke et al., 2004; Williams et al., 2000) have been published which have assessed these studies. The majority of these reviews have concluded that SJW is superior to placebo and as effective as pharmacological antidepressant treatment for mild to moderate depression (Clement et al., 2006; Ernst, 1995; Gaster and Holroyd, 2000; Kim et al., 1999; Linde et al., 1996; Williams et al., 2000). However, the varied methodologies have resulted in differing estimates of effectiveness (Pilkington et al., 2006). A recent systematic review (Linde et al., 2005), which included 37 trials (4925 patients), concluded that the findings are more heterogeneous than previously thought, and proposed that this may partly be due to the variation in patient populations that have been studied. The most methodologically sound trials assessing patients with MD only provided evidence of a minimal benefit of SJW over and above placebo treatment (Linde et al., 2005). However, Linde et al. concluded that the evidence still suggested that SJW provides benefit for patients with mild to moderate depressive symptoms, who do not necessarily meet the criteria for MD.
Little is known about the mechanism of action by which SJW alleviates depression (Bennett et al., 1998; Nangia et al., 2000). The monoamine theory of depression proposes that depression results from a decreased concentration of monoamine neurotransmitters in the brain (Carlson, 1988), particularly serotonin and noradrenalin (Hirschfeld, 2000). Although common antidepressants differ in their mechanisms of action, they are believed to reduce depression by increasing brain monoamine levels or availability (Zanoli, 2004). Although antidepressant medication, such as SSRI treatment, usually inhibits the serotonin transporter within hours (Hirschfeld, 2000), therapeutic effects only result after a few weeks (Teufel-Mayer and Gleitz, 1997). Because of this observation, the monoamine theory of depression was revised to include involvement of receptor adaptation (Hirschfeld, 2000). Researchers have proposed that SJW may also reduce depression through action on the serotonergic (Perovic and Muller, 1995), noradrenergic and dopaminergic systems (Pilkington et al., 2006). Extracts of SJW have been shown to lead to increased levels of serotonin, noradrenalin and dopamine in the brain (Calapai et al., 1999; Calapai et al., 2001), and an up-regulation of the 5-HT1A and 2A receptors (Teufel-Mayer and Gleitz, 1997) in the frontal cortex (Muller et al., 1997). This has been proposed to be a possible mechanism of action of SJW.

Increased levels of pro-inflammatory cytokines may also be involved in the pathophysiology of depression (Dantzer et al., 1999; Maes et al., 1995; Smith, 1991; see also section 1.5.3.2.2). SJW has been shown to have anti-inflammatory effects (Hu et al., 2006; Winkler et al., 2004), leading to the proposal that the antidepressant properties of SJW arise from inhibition of proinflammatory cytokines (Thiele et al., 1994), particularly interleukin-6 (IL-6) (Fiebich et al., 2001; Thiele et al., 1994), which may facilitate regulation of corticotrophin releasing hormone (CRH) concentration (Thiele et al., 1994). Thiele et al. (1994) found that in vitro administration of hypericum resulted in suppression of IL-6 concentrations in both depressed and healthy patients.

Several researchers have suggested that the antidepressant properties of SJW arise from a combination of mechanisms, each being too weak on their own to produce a therapeutic effect (Bennett et al., 1998; Nangia et al., 2000; Raffa, 1998). Calapai et al. (2001) administered the hypericum extract Ph-50 to male mice who had/ had not had their IL-6 gene inactivated (knockout, wildtype). They found that hypericum administration reduced immobility duration during a forced swimming task in the
wildtype, but not the knockout mice. Moreover, the wildtype mice showed significantly
greater increases in diencephalic 5-HT and 5-HIAA content. Therefore, IL-6 may
mediate the antidepressant effects of hypericum through activation of the serotonergic
system (Calapai et al., 2001).

7.1.3 Active ingredients of SJW

SJW contains several bioactive compounds (Gaster and Holroyd, 2000; Linde et al.,
2005; Mennini and Gobbi, 2003; Zanoli, 2004), including naphthodianthrons
(hypericins), flavonoids, phloroglucinol derivatives (e.g hyperforin) and
proanthocyanidins (Mennini and Gobbi, 2004). There is still a lack of consensus as to
which of these is/ are responsible for the antidepressant effect of SJW (Mennini and
Gobbi, 2004; Muller et al., 1997; Nangia et al., 2000; Teufel-Mayer and Gleitz, 1997:
Wheatley, 1998; Zanoli, 2004), with some researchers arguing that it is likely to be due
to a combination of these compounds (Linde et al., 2005; Nangia et al., 2000).
Hypericin was originally considered to be the central component (Bennett et al., 1998;
Gaster and Holroyd, 2000; Mennini and Gobbi, 2004), and many extracts of SJW are
still standardized by their hypericin content (Gaster and Holroyd, 2000). Recently, it has
been argued that this compound does not play a significant role (Cott, 1997; Mennini
and Gobbi, 2004; Wheatley, 1998), and attention has turned to the role of hyperforin
(Cervo et al., 2002; Chatterjee et al., 1998; Laakmann et al., 1998; Muller et al., 1998;
Zanoli, 2004), which has been shown to inhibit the reuptake of serotonin, noradrenalin
and dopamine (Chatterjee et al. 1998; Muller et al., 1997; Singer et al., 1999; Zanoli,
2004).

Although the SJW products that have been used in most previous research were
standardised by their percentage hypericin content (Gaster and Holroyd, 2000; Linde et
al., 2005), usually to the value of 0.3% (Mennini and Gobbi, 2004), products vary
greatly in terms of other chemical constituents (Bennett et al., 1998; Linde et al., 2005;
Mennini and Gobbi, 2004), including hyperforin (Mennini and Gobbi, 2004). Most
studies do not even specify the exact content of the extracts used (Nangia et al., 2000).
This makes comparison of the literature problematic, especially as it remains uncertain
as to which components of SJW are involved in its therapeutic effect (Mennini and
Gobbi, 2004; Muller et al., 1997; Nangia et al., 2000; Teufel-Mayer and Gleitz, 1997;
7.1.4 SJW for PMS

SJW has been shown to be an effective treatment for mild to moderate depression (Clement et al., 2006; Ernst, 1995; Gaster and Holroyd, 2000; Kim et al., 1999; Linde et al., 1996; Williams et al., 2000), which is a core symptom of PMS (Eriksson et al., 1995; Steiner et al., 1995). There are some similarities between the symptoms of depression and PMS (Hicks et al., 2004; Steiner et al., 1995). Biochemical similarities also exist between these conditions (Hicks et al., 2004). Abnormality of serotonergic function has been demonstrated in depressed patients (Carlson, 1988; Hirschfeld, 2000; Zanoli, 2004) and in women suffering from PMS (Ashby et al., 1988; Bancroft et al., 1991; Halbreich et al., 2002; Rapkin, 1992; Rapkin et al., 1987; Taylor et al., 1984). Moreover, elevated levels of proinflammatory cytokines have been demonstrated in depressed patients (Dantzer et al., 1999; Maes et al., 1995; Smith, 1991) and have been shown to occur premenstrually (Cannon and Dinarello, 1985; Konecna et al., 2000; Polan et al., 1990). As SJW has been shown to influence the serotonergic system (Calapai et al., 1999; Calapai et al., 2001; Muller et al., 1997; Teufel-Mayer and Gleitz, 1997) and to suppress pro-inflammatory cytokine levels (Thiele et al., 1994; Fiebich et al., 2001), this herbal product deserves attention as a potential PMS treatment (Hicks et al., 2004). As PMS usually lasts until menopause, treatment must be effective and safe for use over a long period (Dimmock et al., 2000). SSRIs are often used in PMS treatment (Eriksson et al., 2002; Yonkers et al., 2008), and some believe that they should be the first line of therapy for sufferers with severe symptoms (Dimmock et al., 2000). However, SSRIs have been shown to result in side effects in PMS sufferers (Eriksson et al., 2002; Johnson, 2004; Yonkers et al., 2008). As tolerability has been shown to be excellent with the use of SJW (Miller, 1998; Nangia et al., 2000; Pilkington et al., 2006; Schulz, 2002; Seelinger and Mannel, 2007; Wheatley, 1998), this intervention could prove extremely beneficial for PMS sufferers if it was found to be effective. Moreover, SJW is cheaper than SSRI treatment (Bennett et al., 1998; Gaster and Holroyd, 2000). As SJW is currently available over the counter, evidence of efficacy is also needed to protect the women currently self-diagnosing with PMS and buying this product in this way.

7.1.7 Previous trials of SJW for PMS

As discussed in Section 6.3.6, three trials have been performed to test the efficacy of SJW for PMS, and all produced positive results (Hicks et al., 2004; Pakgohar et al., 2004; 2005). However, Pakgohar et al. (2004; 2005) omitted many methodological
features from their abstracts, which were the only translated parts of these Iranian articles, hence the quality of these studies could not be assessed. Hicks et al. (2004) performed a randomised, double-blind, placebo controlled, parallel groups study, to assess the effectiveness of SJW for PMS. No significant differences between SJW and placebo treatment were found. Combining symptom scores from both treatment cycles resulted in SJW appearing more beneficial than placebo treatment for all of the symptom subgroups assessed, but these combined effects did not reach statistical significance. The authors concluded that their non-significant findings may have resulted from insufficient power, rather than a lack of efficacy of SJW. A dose of 600mg/day was used. Most previous research which has assessed the efficacy of SJW for depression (Bennett et al., 1998; Miller, 1998; Montgomery et al., 2000; Shelton et al., 2001) has employed a higher dose of 900mg/day. This is also the dose recommended by most manufacturers of SJW to improve mood. Therefore, further research assessing the effectiveness of SJW as a PMS treatment, using a higher dose of 900mg/day, is warranted.

7.1.6 The placebo effect

Women participating in PMS treatment studies show a strong response to placebo treatment (Dimmock et al., 2000; Eriksson et al., 1995; Freeman and Rickels, 1999; Magos et al., 1986; Steiner et al., 1995; Yonkers et al., 1997). Researchers have documented placebo response rates between 23% and 94% (Magos et al., 1986; Rapkin., 2003; Steiner et al., 1995). Freeman and Rickels (1999) combined the data from two PMS treatment studies, each of which subjected one group of women to four consecutive cycles of placebo treatment. They found that 20% of these women reported a 50% or greater reduction in symptoms from baseline values in two to four of the placebo cycles. They labelled these women as having ‘sustained improvement’ and reported that most of their improvement occurred during the first two cycles. They also found that 42% of women showed ‘partial improvement,’ demonstrating a reduction in symptoms during one cycle of placebo treatment.

Researchers who have documented placebo response rates have usually done so through reporting the percentage of women taking placebo treatment who meet the study’s improvement criteria. This usually involves demonstrating a particular percentage reduction in the symptoms reported on treatment, compared to those reported at baseline, usually on the total score of the scale employed by the study (Freeman et al.,
Therefore, although the extent to which the placebo effect operates is fairly well documented, little research has examined which PMS symptoms respond to placebo treatment.

7.1.7 Aims
The primary objective of this study was to determine whether SJW was more beneficial than placebo treatment in relieving symptoms of PMS, in a 10 cycle double-blind randomised controlled crossover trial. A subsidiary aim was to explore the nature of the placebo effect often observed in PMS research, in terms of the symptoms involved and the duration of these effects.

7.2 Method
7.2.1 Sample size
At the time this study was designed, only one small pilot study testing the efficacy of SJW for PMS had been published (Stevinson and Ernst, 2000). This study used a pre-post test design and did not employ a placebo group or crossover design. Therefore, for the purpose of sample size estimation in the current study, response to placebo had to be estimated. This was done by calculating the half way point between the total DSR score at baseline and the 1st treatment cycle. The effect size was estimated from the published data (Stevinson and Ernst, 2000) using the following formula (Rosenthal and Rosnow, 1999):

\[ d = \frac{\text{Treatment} - \text{Placebo}}{\sigma_{\text{within}}} \]

\[ d = \frac{(\text{Treatment} - (\text{half way between the baseline measure and 1st treatment cycle measure}))}{(\sigma_{\text{within}}(\text{df within}/N))} \]

\[ d = \frac{((70.11 + ((128.42 - 70.11)/2)) - 57.63)}{(68.15 \sqrt{(18/19))}} \]

\[ d = 0.63 \]

Sample size was then estimated using the tables provided by Machin and Campbell (1987) using the equation \( d = |\mu_0 - \mu_1| / \text{population standard deviation} \), and rounding upwards to the nearest whole number. Machin and Campbell's (1987) Table 7.1 was used, as this is suitable for paired data, and therefore appropriate for cross-over designs, where participants receive each of two treatments at different times. A crossover design, where each woman acted as her own control, was employed to allow a smaller sample size to be used and to remove between subject variability between treatments.
From Table 7.1, for a two-sided test with $\alpha = 0.05$, power $(1 - \beta) = 0.80$, and $d = 0.63$: $m = 23$ for $d = 0.60$, and $m = 20$ for $d = 0.65$. Thus, approximately 22 women would correspond to $d = 0.63$.

7.2.2 Participant enrolment

A similar study carried out at the University of Leeds, which examined the effect of soy isoflavones on premenstrual symptom severity (Bryant et al., 2005), initially screened 41 women, which, following withdrawals left 23 women completing the study. Therefore, on the basis of this experience, it was determined that at least 41 women should enter the screening phase to ensure the required sample size of 22.

7.2.3 Participants

Volunteers were recruited for the study through various strategies (see section 3.4.1.1), all of which made it clear that the study required women who thought that they may suffer from PMS. Most participants were recruited from the Yorkshire area due to the frequency of visits demanded by the study protocol. However, some participants from outside the Leeds area were accepted onto the study, as they guaranteed that they would be able to attend all of the study visits required.

7.2.4 Inclusion/Exclusion Criteria

Participant eligibility was assessed through the use of preset inclusion and exclusion criteria (See Table 7.1). Participants were required to be between the ages of 18-45 years, so that they were old enough to have regular menstrual cycles, but young enough for it to be unlikely that menopausal symptoms were affecting their menstrual cycle. Throughout the study, it was essential to arrange study visits during specific cycle phases, which were predicted by the PI from participants’ average cycle lengths. Therefore, participants were required to have regular menstrual cycles (25-35 days). The use of pharmaceutical or herbal preparations, which could alter the hormonal profiles of women and which could therefore alter the timing and nature of the cycle, served as an exclusion criterion, while women taking prescribed or over the counter (OTC) medication for PMS were excluded to allow the sole effect of SJW on premenstrual symptomatology to be assessed. Women taking prescribed drugs (Appendix 3.6) which could interact with SJW were excluded. Manufacturer’s information also suggested that SJW could increase photosensitivity. Therefore women who had been diagnosed with photosensitivity were excluded. Eligibility was initially
assessed through completion of the recruitment questionnaire, which was followed up by a meeting with the study physician (Dr Julie Ayres) and PI.

Table 7.1 Inclusion and exclusion criteria for participation

<table>
<thead>
<tr>
<th>INCLUSION CRITERIA</th>
<th>EXCLUSION CRITERIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>From the Leeds area, or able to visit regularly</td>
<td>Taking prescribed or OTC medication for PMS</td>
</tr>
<tr>
<td>Aged 18-45 years</td>
<td>Using hormonal preparations, including oral contraception or HRT</td>
</tr>
<tr>
<td>Of good physical and psychological health, assessed by the clinic doctor</td>
<td>Pregnant, planning pregnancy or lactating</td>
</tr>
<tr>
<td>Regular menstrual cycles (25-35 days)</td>
<td>Photosensitive</td>
</tr>
<tr>
<td></td>
<td>Taking any of the preparations listed in Appendix 3.6</td>
</tr>
</tbody>
</table>

7.2.5 Participant progress

371 women completed the online recruitment questionnaire. Of these, 270 were not suitable to begin screening or lived too far away. These women were offered the opportunity to take part in postal screening (see sections 3.4.3 and 3.4.5). 101 women attended the briefing meeting about the clinical trial, of which 91 began screening. At the end of screening, 38 women were suitable and willing to continue onto the intervention, 34 of whom completed the full clinical trial. For a record of the progress of all participants through the whole study, see Figure 7.1.
7.2.6 Study design

The study conformed to a randomised, placebo-controlled, double-blind, crossover design across 10 menstrual cycles. All participants underwent prospective screening for three menstrual cycles, took placebo tablets for two cycles (placebo run-in), before being randomly allocated to one of two treatment conditions (continued placebo or SJW, 900mg/day). Women then underwent a washout cycle before crossing over to the other treatment condition (SJW (900mg/day) or placebo treatment). Women were given placebo tablets throughout the washout cycle to prevent them speculating about the study design.

Randomisation was achieved by allocating each participant a participation number in sequential order (starting with 01, 02 etc). Participants were then randomised to
treatment conditions by the chief pharmacist (Caroline Bedford) at Leeds General Infirmary. The randomisation was done using random number tables in blocks of ten, so that for every ten participants randomised, five received SJW followed by placebo and five received placebo followed by SJW. Random numbers, rather than labels such as A and B, were used to identify the two experimental conditions to prevent the PI inferring treatment orders. The master file of the allocation codes was kept by the hospital pharmacy, who along with the lead academic supervisor of the PI (Dr Louise Dye), held a copy in a sealed envelope to enable the codes to be broken in case of an emergency, when the knowledge of the investigational product might be essential for the clinical management of the participant. The code was broken twice during the study. One participant became pregnant during study cycle 10, while another had to take emergency hormonal contraception.

A two cycle placebo run-in was employed because women with PMS show a strong placebo response (Dimmock et al., 2000; Eriksson et al., 1995; Freeman and Rickels, 1999; Magos et al., 1986; Steiner et al., 1995; Yonkers et al., 1997). As most symptom reduction involved in this effect occurs during the first two cycles during which placebo treatment is taken (Freeman and Rickels, 1999), the placebo run-in minimised the likelihood of a placebo response clouding the treatment effect by permitting the placebo response to level out before treatment was administered. The placebo run-in also allowed the placebo effect to be examined.

7.2.7 Dose and clinical supplies
The only pilot study published at the time this study was designed (Stevinson and Ernst, 2000) used 300mg of SJW, standardised to 0.3% hypericin, and reported a significant reduction of PMS symptoms, with over two thirds of their sample demonstrating at least a 50% decrease in symptom severity. However, this study did not include a placebo group or crossover treatment. Thus, some of these premenstrual symptom reductions are probably accounted for by the placebo response. Most clinical studies which test the efficacy of SJW for depression (Bennett et al., 1998; Miller, 1998; Montgomery et al., 2000; Shelton et al., 2001) have administered a dose of 900mg/day, standardised to 0.3% hypericin, which is also the dose recommended by most manufacturers of SJW to improve mood. On the basis of these findings, a dose of 900mg/day was employed in this study. The SJW tablets were requested to be standardised to 0.3% hypericin, in line previous studies (Linde et al., 2005; Gaster and Holroyd, 2000).
Lichtwer Pharma provided 450mg tablets of SJW and identical placebo tablets, both produced according to Good Manufacturing Practice (GMP) guidelines. The elimination half-life is 24 hours for hypericin (Lichtwer Pharma, Appendix 8.1), and 9 hours for hyperforin (Nangia et al., 2000). The tablets were packed in identical blister packs of 15 tablets, labelled with the participant treatment code by Lichtwer Pharma in Berlin and sent to the pharmacy at Leeds General Infirmary. The PI made regular visits to the pharmacy to pick up the tablets to distribute to participants. The study medication was stored securely, at room temperature, by both the pharmacy and in the HARU. Each participant was given seven boxes of tablets throughout the entire study, each containing 75 tablets and were instructed how to store these. Participants were given boxes of tablets at regular intervals to try to ensure that they were taking the tablets from the appropriate boxes. Moreover, each box was clearly labelled with the study cycle number (cycles 4-10). Participants were instructed to begin the appropriate box of tablets on the first day of each menstrual cycle, to prevent them from knowing how long they were taking each treatment for, and were asked to return unused medication at each study visit.

7.2.8 Ethical and Regulatory Approval

Ethical approval to carry out the study was granted by the Institute of Psychological Sciences (16/12/2004) and the Leeds (West) Research Ethics Committee (23/6/2006). The Medicines and Healthcare products Regulatory Agency (MHRA) also granted approval for the clinical trial (23/9/2005).

7.2.9 Measures

7.2.9.1 Recruitment questionnaire (Appendix 8.2): All participants completed this questionnaire at the beginning of the study to gather demographic information such as age, psychiatric history, use of medication/hormonal preparations and menstrual cycle length. The recruitment questionnaire acted as an initial screening tool to assess whether participants met the main inclusion/exclusion criteria.

7.2.9.2 PMS symptoms: The daily diary booklets, described in Section 3.1.1, containing seven copies of the DSR (Freeman et al., 1996) and one copy of the BDI (Beck et al., 1961), STAI (Spielberger, 1983b), BPAQ (Buss and Perry, 1992) and BIS-11 (Patton et al., 1995) were used throughout the study to assess the presence and severity of PMS and levels of depression, anxiety, aggression and impulsivity. These measures were also
the primary tools in the diagnosis of PMS (see section 3.5.2.1). Women were asked to complete the daily diary booklets for 10 cycles. Therefore, to reduce the time needed to complete the booklets, the trait measure of anxiety, STAIT, appeared only in the diaries of the first menstrual cycle, following Spielberger’s (1983b) recommendation that the STAIS should be given each time a measure is needed to assess change in anxiety over time.

7.2.9.3 Emotional Stroop Task: Participants were familiarised with the emotional Stroop task during their first study visit and were tested using it on all subsequent visits. See Section 3.4.4.4 for the emotional Stroop task procedure. These data are outside the scope of this thesis and will be reported elsewhere.

7.2.9.4 Blood samples: 10 blood samples (each 18mls) were taken during the course of the study to assess follicular and luteal hormone (FSH, LH, oestradiol, progesterone, prolactin and testosterone), serotonin (5-HT and 5-HIAA) and cytokine (IL-1β, IL-6, IL-8, IFN-γ and TNF-α) levels at baseline, placebo run-in, washout and during active and placebo treatment. Samples were collected as described in Section 3.4.4.3, and hormone and cytokine levels were assessed as described in Section 5.3.2.2.

7.2.10 Study visits
Follicular visits took place on cycle days 5-10, and luteal visits on days -1 to -6. The timing of these visits was estimated by the PI from the length of the woman’s first two cycles and from information about cycle lengths provided on the recruitment questionnaire. For example, for a usual cycle length of 30 days, visits were arranged as closely as possible to days 7-8 and 27-28. After a follicular visit had been booked, participants were asked to telephone the PI on the first day of their cycle to check that the appointment had been arranged during the correct phase of their cycle. If not, the appointment was re-scheduled.

On each visit to the HARU, the PI made a safety assessment to make sure that it was safe for the women to continue in the study, through the use of a semi-structured interview. This technique permitted the women to ask any questions they had and voice any concerns, whilst allowing the experimenter to cover specific questions, regarding potential side effects, the consumption of tablets and the health of the women (Appendix 8.3). Although the PI had an interview guide, with specific questions to
cover, issues that arose were discussed as appropriate. If there was any concern about the health of a participant, the responsible physician on the trial (Mr. Nigel Simpson) was contacted for advice, and the participant was offered the opportunity for a consultation with the physician. Furthermore, all participants had met with the study doctor (Dr Julie Ayres) before they were entered onto the trial, to ensure that they were in good physical and psychological health.

7.2.11 Procedure

Various recruitment methods were used to make women in the Yorkshire area aware of the trial, all of which directed attention to the study's website and gave the contact details of the PI (see section 3.4.1.1). If women contacted the PI, they were informed about what the study involved and were asked to fill in the recruitment questionnaire online if still interested. If they did not have internet access, they were sent a copy of the questionnaire by post. Women also reached this stage by accessing the website and completing the online recruitment questionnaire directly after seeing it advertised i.e. without contacting the PI.

If the recruitment questionnaire indicated that women did not live in the Leeds area, did not meet the inclusion criteria or met the exclusion criteria, they were contacted, informed that they could not take part in the clinical trial and were asked if they would like to take part in the three screening cycles by post. If they wanted to do this, they were sent an information sheet outlining what would be involved (Appendix 3.9) and a consent form (Appendix 3.10). On return of the consent form, participants were sent the first diary pack and asked to begin diary completion on the first day of their next period and to continue doing so for three full menstrual cycles. Once the participants had completed their symptom recording, the PI analysed their data and sent them a detailed symptom profile (for an example see Appendix 8.4).

Women who met the inclusion criteria for the clinical trial were sent the information sheet detailing the clinical trial (Appendix 3.8) and contacted to arrange their first study visit, which took place in the Rosalind Bolton PMS Clinic, Leeds General Infirmary. All subsequent study visits took place in the HARU. During the first study visit, participants were given the opportunity to ask any questions that had arisen from reading the information sheet. A medical history was then taken from the participants and their blood pressure was recorded. If the study doctor (Dr Julie Ayres) was satisfied
with their health, the admission form (Appendix 3.12) was completed and informed consent was taken (Appendix 3.13). Participants were then familiarised with the emotional Stroop task.

There were five phases to the clinical trial. A schematic representation of the study design is detailed in Figure 7.2.
### Phase 1 - Screening

<table>
<thead>
<tr>
<th>Cycle 1</th>
<th>Cycle 2</th>
<th>Cycle 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSR daily</td>
<td>DSR daily</td>
<td>DSR daily</td>
</tr>
<tr>
<td>BDI, STAI, BPAQ, BIS weekly</td>
<td>BDI, STAI, BPAQ, BIS weekly</td>
<td>BDI, STAI, BPAQ, BIS weekly</td>
</tr>
</tbody>
</table>

Follicular and luteal cognitive testing
Follicular and luteal blood sampling

### Phase 2 - Single-Blind Placebo Run-in Period

<table>
<thead>
<tr>
<th>Cycle 1</th>
<th>Cycle 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSR daily</td>
<td>DSR daily</td>
</tr>
<tr>
<td>BDI, STAI, BPAQ, BIS, AE reporting weekly</td>
<td>BDI, STAI, BPAQ, BIS, AE reporting weekly</td>
</tr>
</tbody>
</table>

Follicular and luteal cognitive testing
Follicular and luteal blood sampling

### Phase 3 - Treatment 1: A. Placebo Treatment

<table>
<thead>
<tr>
<th>Cycle 1</th>
<th>Cycle 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSR daily</td>
<td>DSR daily</td>
</tr>
<tr>
<td>BDI, STAI, BPAQ, BIS, AE reporting weekly</td>
<td>BDI, STAI, BPAQ, BIS, AE reporting weekly</td>
</tr>
</tbody>
</table>

Follicular and luteal cognitive testing
Follicular and luteal blood sampling

### Phase 3 - Treatment 1: B. SJW (900mg/day)

<table>
<thead>
<tr>
<th>Cycle 1</th>
<th>Cycle 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSR daily</td>
<td>DSR daily</td>
</tr>
<tr>
<td>BDI, STAI, BPAQ, BIS, AE reporting weekly</td>
<td>BDI, STAI, BPAQ, BIS, AE reporting weekly</td>
</tr>
</tbody>
</table>

Follicular and luteal cognitive testing
Follicular and luteal blood sampling

### Phase 4 - Washout

DSR daily
BDI, STAI, BPAQ, BIS, AE reporting weekly
Follicular and luteal cognitive testing
Follicular and luteal blood sampling

### Phase 5 - Treatment 2: A. Placebo Treatment

Cycles 1 & 2 as above

### Phase 5 - Treatment 2: B. SJW (900mg/day)

Cycles 1 & 2 as above

---

**Figure 7.2 Schematic representation of the study design and measures completed in each study phase**

**KEY:** DSR = Daily Symptom Report; BDI = Beck Depression Inventory (1961); STAI = State-Trait Anxiety Inventory (1983); BPAQ = Buss Perry Aggression Questionnaire (1992); BIS = Barratt Impulsiveness Scale (Patton et al., 1995); AE = Adverse event
7.2.11.1 Phase 1 - Screening (cycles 1-3)

During the first study visit, women were given their first set of daily diary booklets and asked to begin completing these on the first day of their next period and to carry on doing so on a daily basis until the end of the study, returning each diary booklet as soon after its completion as possible. Throughout the study, two diaries were sent to participants biweekly with Freepost envelopes for their return, with a letter reminding them of what they should be doing at that time and dates of forthcoming visits.

Participants visited the PI during the follicular and luteal phases of the third screening cycle to perform the emotional Stroop task and to have a blood sample taken. Eligibility to continue to the intervention phase was assessed from the analysis of the diary data from the first two screening cycles. If participants met the 30% increase criterion on the DSR (Freeman et al., 1996) in both cycles (see section 3.5.2.1), and did not meet criteria for anxiety or depression (see section 3.5.1), women were informed that they met PMS criteria and could continue with the study. These women were given tablets for cycles 4 and 5 and asked to begin taking tablets from the box labelled ‘cycle 4’ on the first day of their next period. The PI sent a letter to their GP to inform him/her of their participation in the study. Participants were requested to inform the PI immediately if they started new medication. If they did, the PI contacted the study doctor to check whether the participant could continue in the study.

If participants did not meet the 30% increase criterion on the DSR in cycles 1 and 2 (see section 3.5.2.1), or if they met criteria for anxiety or depression (see section 3.5.1), they were informed that they did not meet the PMS diagnostic criteria for this study and could not continue. They were given a detailed symptom profile (for an example see Appendix 8.4) and thanked for their time. If participants met the 30% increase criterion in only cycle 1 or 2, and did not meet depression or anxiety criteria, they were asked to undergo a fourth screening cycle to allow time for the diaries completed during the third cycle to be analysed. If participants were found to show the required increase in symptoms during cycle 3 (hence 2 out of 3 cycles), they were asked to visit the HARU to pick up tablets for cycles 4 and 5. Diaries that were completed during the fourth screening cycle were not analysed but were given to the participants to keep them in the cycle of completing diaries daily. The term ‘cycle 4’ refers to the first cycle of placebo run-in, although technically it was the fifth cycle in which these women had participated. Women were briefed in the same manner as those participants meeting
PMS criteria in cycles 1 and 2. If participants did not show the required increase in symptoms in the third screening cycle, they were informed of this by telephone and sent a detailed symptom profile by post.

7.2.11.2 Phase 2 - Single-blind placebo run-in (cycles 4-5)
Phase 2 was single blind, as the PI was aware that all participants were taking placebo tablets throughout this phase and the participating women were not. Women visited the PI during the follicular and luteal phases of cycle 5 to perform the emotional Stroop task and to have a blood sample taken. Any unused medication from cycle 4 was collected by the PI\(^2\), who returned these tablets to the pharmacy. In the luteal visit of cycle 5, women were given their tablets for cycles 6 and 7.

7.2.11.3 Phases 3 and 5 - Treatment 1 and 2 (cycles 6 & 7 and 9 & 10)
Phases 3 and 5 were double blind. Women took either 2 x 450mg tablets of placebo or SJW treatment daily throughout these phases, depending on which condition they had been assigned to by pharmacy. Participants visited the PI to perform the emotional Stroop task and to have a blood sample taken during the follicular and luteal phases of cycles 7 and 10. During their luteal visit of cycle 7, women were given their tablets for cycle 8. Unused medication from cycles 5 and 6 was returned to the PI in the luteal visit of cycle 7, while unused medication from cycles 8 and 9 was returned during the luteal visit of cycle 10. Unused medication from cycle 10 was returned to the PI in a Freepost envelope. The PI returned all unused medication to the hospital pharmacy.

7.2.11.4 Phase 4 - Washout (Cycle 8)
As in Phase 2, Phase 4 was single-blind. Phase 4 ensured that no active ingredients of SJW remained in those participants who had taken active treatment in phase 3 when allocated to the second treatment condition. Participants continued to complete the daily diary booklets, as discontinuity could have had a greater effect than continuation. Women visited the PI during their follicular and luteal phases of cycle 8 to have a blood sample taken and to perform the emotional Stroop task, partly to maintain regular contact with the participants. Tablets for cycles 9 and 10 were given to participants during their luteal visit and any unused medication from cycle 7 was collected by the PI, who then returned these tablets to the pharmacy.

\(^2\) Usually because of shorter cycle lengths rather than non compliance
7.2.12 Study completion
Following the final participant’s completion, the pharmacy unblinded the PI to the treatment conditions that the participants had been allocated to and destroyed all remaining tablets in accordance with the Trust’s pharmaceutical waste policy. Of the 34 participants who completed the study, 18 began taking SJW treatment followed by placebo, while 16 began taking placebo treatment followed by SJW. Participants were given an honorarium (£220) to compensate them for their time and travel costs. Women had not been informed of the honorarium until they had entered the placebo run-in, to prevent women from making false reports during the screening cycles for monetary reasons.

7.3 Data screening
DSR scores (Freeman et al., 1996) were explored in order to determine whether SJW was more beneficial than placebo treatment in relieving symptoms of PMS and to explore the nature of the placebo effect. Scores from the two factor DSR-20 (see Table 4.10) were assessed to determine whether the original DSR items, and additional three items ‘anger,’ ‘aggression’ and ‘impulsiveness,’ clustered together through the use of the factors derived in Chapter 4 (see section 4.5), and weighted by their factor score coefficients (see section 4.8.1), was a more sensitive tool to examine treatment and placebo effects. Scores from the STAIS, BDI, BIS-11 and BPAQ were examined to determine whether treatment and placebo effects operated on the core PMS symptoms anxiety, depression, impulsivity and aggression. Prior to data analysis, the dataset were screened for missing data, outliers and normality.

7.3.1 Missing data and protocol violations
7.3.1.1 Blood samples and the Emotional Stroop Task
As discussed in Section 5.4.3.2.1, some follicular phase data were missed due to participants failing to notify the PI when their period began, while some luteal phase data were missed due to unsuccessful prediction of participants’ cycle lengths. Moreover, some missing biological data arose from unsuccessful blood sampling.

7.3.1.2 Missing diary completions
Out of the 34 participants who completed the trial, five had a missing phase score for one study cycle on the DSR (i.e. one score from four phases of ten cycles (0.37% data)). 17 participants also had one or two missing phase scores for all weekly diary
questionnaires (28 missing in total (2.06% data)). Therefore, these values were estimated using the formula recommended by Tabachnick and Fidel (2007a).

7.3.2 Outliers
Outlying scores were not excluded from the analysis since the main concern was to examine how participant’s scores changed across time and in order to maintain the number of valid observations. Therefore, participants’ scores were retained, even if they were outliers (greater or less than 2sds from the mean) relative to the rest of the sample.

7.3.3 Normality
To check whether the data were normally distributed, both Z scores for skewness and the Kolmogorov-Smirnov (K-S) test were consulted for each group, for each DV (DSR and DSR-20 factor scores; STAI, BDI, BPAQ and BIS-11 total and factor scores), at each time-point. Where Z-scores for skewness were found to be greater than 2.58 (Field, 2005) and the K-S was significant, the data were considered to significantly deviate from a normal distribution. In repeated measures designs, transformation of one DV makes it necessary to transform all DVs making up that measure. Transformations were not considered necessary for the weekly questionnaire measures, as the majority of STAI, BDI BPAQ and BIS-11 DVs did not deviate from a normal distribution.

The majority of DVs comprising the DSR (Freeman et al., 1996) and the DSR-20 were found to be significantly positively skewed. The DSR total and subscale scores (Freeman et al., 1996) were transformed through the use of a square root transformation, which normalised the vast majority of DVs. Square root transformations could not be used for the DSR-20, as the factor scores comprising this subscale were calculated through the use of the regression method (see section 4.8.1), which resulted in some values having a negative weighting. Therefore, the DSR-20 factor scores were transformed through the use of a Log + 1 transformation. This resulted in most DVs becoming more normally distributed, although many DVs comprising the physical factor remained positively skewed. The results discussed for the DSR (Freeman et al., 1996) and the DSR-20 arise from the transformed scores. However, to be able to make meaningful comparison with the existing literature, discussion of mean values, and the values presented in tables and figures are based on untransformed data.
7.3.4 Sphericity

Sphericity was not a concern for the current study, since profile analysis was conducted, and this circumvents the need for the sphericity (Tabachnick and Fidell, 2007a). However, when simple effects and interaction contrast analyses were conducted (see section 5.4), the Greenhouse Geisser’s correction was applied when Mauchly’s (1940) test indicated that sphericity was not met.

7.4 Results

7.4.1 Sample characteristics

Sample characteristics are presented in Table 7.2. These were examined using independent samples t-tests and Pearson’s chi-square test, according to the order in which participants received active and placebo treatment. No significant differences or associations with treatment order were found between the groups on any characteristic.

Table 7.2 Characteristics of the clinical trial completers at study entry

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total (n=34)</th>
<th>Active 1st (n=18)</th>
<th>Placebo 1st (n=16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (sd)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>35.26 (5.88)</td>
<td>33.56 (6.49)</td>
<td>37.18 (4.58)</td>
</tr>
<tr>
<td>No. PMS years</td>
<td>12.67 (7.50)</td>
<td>13.41 (7.14)</td>
<td>11.77 (8.08)</td>
</tr>
<tr>
<td>BMI</td>
<td>25.13 (4.34)</td>
<td>25.66 (5.00)</td>
<td>24.54 (3.50)</td>
</tr>
<tr>
<td>N (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- No. births</td>
<td>21 (61.8)</td>
<td>9 (26.5)</td>
<td>12 (35.3)</td>
</tr>
<tr>
<td>- No. miscarriages</td>
<td>3 (8.8)</td>
<td>3 (8.8)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>- No. terminations</td>
<td>4 (11.8)</td>
<td>2 (5.9)</td>
<td>2 (5.9)</td>
</tr>
</tbody>
</table>

7.4.2 Baseline symptomatology

Table 7.3 displays the mean severity scores of the daily and weekly measures that were administered to the clinical trial participants during the three screening cycles. Cyclicity is evident for all symptoms. The highest mean symptom scores were evident during the luteal phase for all measures, with the exception of the behavioural and pain subscales of the DSR (Freeman et al., 1996) and the physical factor of the DSR-20, which were highest during bleeding.

---

3 Statistical analyses not performed, as cell frequency < 5
Table 7.3 Mean (sd) symptom severity during screening (average of study cycles 1-3)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Menstrual</th>
<th>Follicular</th>
<th>Rest</th>
<th>Luteal</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSR: Total</td>
<td>12.84 (9.30)</td>
<td>4.26 (4.46)</td>
<td>5.16 (4.83)</td>
<td>13.66 (8.28)</td>
</tr>
<tr>
<td>-Mood</td>
<td>5.00 (4.13)</td>
<td>1.98 (2.11)</td>
<td>2.31 (2.08)</td>
<td>5.72 (3.90)</td>
</tr>
<tr>
<td>-Behavioural</td>
<td>3.97 (3.33)</td>
<td>1.78 (1.74)</td>
<td>1.77 (1.74)</td>
<td>3.86 (2.97)</td>
</tr>
<tr>
<td>-Pain</td>
<td>2.38 (1.62)</td>
<td>0.28 (0.59)</td>
<td>0.54 (0.82)</td>
<td>2.02 (1.38)</td>
</tr>
<tr>
<td>-Physical</td>
<td>1.49 (1.06)</td>
<td>0.22 (0.34)</td>
<td>0.55 (0.49)</td>
<td>1.93 (1.07)</td>
</tr>
<tr>
<td>DSR-20</td>
<td>0.74 (0.80)</td>
<td>0.36 (0.37)</td>
<td>0.40 (0.35)</td>
<td>0.99 (0.82)</td>
</tr>
<tr>
<td>-Psychological</td>
<td>0.91 (0.80)</td>
<td>0.19 (0.32)</td>
<td>0.27 (0.41)</td>
<td>0.78 (0.65)</td>
</tr>
<tr>
<td>-Physical</td>
<td>0.74 (0.80)</td>
<td>0.36 (0.37)</td>
<td>0.40 (0.35)</td>
<td>0.99 (0.82)</td>
</tr>
<tr>
<td>BPAQ: Total</td>
<td>64.71 (17.34)</td>
<td>53.44 (16.14)</td>
<td>57.98 (16.00)</td>
<td>70.98 (21.12)</td>
</tr>
<tr>
<td>-Verbal</td>
<td>11.79 (3.80)</td>
<td>9.56 (3.37)</td>
<td>10.84 (4.20)</td>
<td>13.59 (4.70)</td>
</tr>
<tr>
<td>-Physical</td>
<td>16.21 (5.15)</td>
<td>14.41 (5.04)</td>
<td>14.42 (4.61)</td>
<td>16.97 (6.85)</td>
</tr>
<tr>
<td>-Anger</td>
<td>18.99 (4.87)</td>
<td>14.59 (4.38)</td>
<td>17.25 (4.97)</td>
<td>21.48 (5.86)</td>
</tr>
<tr>
<td>-Hostility</td>
<td>17.73 (6.09)</td>
<td>14.88 (5.51)</td>
<td>15.48 (5.18)</td>
<td>18.94 (7.30)</td>
</tr>
<tr>
<td>BIS: Total</td>
<td>64.09 (8.21)</td>
<td>57.52 (7.77)</td>
<td>61.32 (8.42)</td>
<td>67.36 (9.21)</td>
</tr>
<tr>
<td>-Attentional</td>
<td>15.12 (2.49)</td>
<td>12.59 (2.28)</td>
<td>14.32 (2.90)</td>
<td>16.16 (3.71)</td>
</tr>
<tr>
<td>-Motor</td>
<td>20.26 (2.88)</td>
<td>19.89 (2.41)</td>
<td>19.72 (2.40)</td>
<td>20.58 (2.84)</td>
</tr>
<tr>
<td>-Nonplanning</td>
<td>28.73 (4.80)</td>
<td>25.08 (5.02)</td>
<td>27.22 (5.00)</td>
<td>30.60 (5.06)</td>
</tr>
<tr>
<td>STAIS</td>
<td>46.29 (8.39)</td>
<td>35.08 (8.16)</td>
<td>41.34 (9.03)</td>
<td>50.14 (10.55)</td>
</tr>
<tr>
<td>BDI</td>
<td>10.30 (6.04)</td>
<td>4.65 (4.13)</td>
<td>7.31 (4.99)</td>
<td>12.72 (7.73)</td>
</tr>
</tbody>
</table>

The variability in symptom reporting over the four phases of the three screening cycles was explored through the use of repeated measures (RM) analysis of variance (ANOVA) and doubly multivariate analysis of variance models. RM ANOVAs were used to analyse measures comprising solely of a total scale score, while doubly multivariate ANOVAs were used to analyse the measures that were made up from separate subscales. Table 7.4 displays the multivariate/ univariate statistics for the main effects of phase and cycle, and for the cycle by phase interactions for all of the measures that were administered during the screening cycles.

Table 7.4 Baseline symptomatology: Multivariate/ univariate statistics for the main effects of phase and cycle, and for the cycle by phase interactions

<table>
<thead>
<tr>
<th>Measure</th>
<th>Phase</th>
<th>Cycle</th>
<th>Cycle by Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSR-Freeman</td>
<td>MV F(12, 22)=36.76**</td>
<td>MV F(8, 26)=1.80</td>
<td>MV F(24, 10)=0.58</td>
</tr>
<tr>
<td>DSR-20:</td>
<td>MV F(6, 26)=23.19**</td>
<td>MV F(4, 28)=0.90</td>
<td>MV F(12, 20)=1.01</td>
</tr>
<tr>
<td>Subscales</td>
<td>MV F(12, 22)=5.21**</td>
<td>MV F(8, 26)=3.40**</td>
<td>MV F(24, 10)=1.36</td>
</tr>
<tr>
<td>BPAQ</td>
<td>MV F(9, 25)=8.62**</td>
<td>MV F(6, 28)=5.04**</td>
<td>MV F(18, 16)=1.09</td>
</tr>
<tr>
<td>BIS-11</td>
<td>F(3, 99)=31.83**</td>
<td>F(2, 66)=2.21</td>
<td>F(6, 198)=0.95</td>
</tr>
<tr>
<td>STAIS</td>
<td>F(3, 99)=25.32**</td>
<td>F(2, 66)=4.94*</td>
<td>F(6, 198)=1.12</td>
</tr>
<tr>
<td>BDI</td>
<td>F(3, 99)=25.32**</td>
<td>F(2, 66)=4.94*</td>
<td>F(6, 198)=1.12</td>
</tr>
</tbody>
</table>

*p<0.05  **p<0.01  MV= Multivariate

Table 7.4 shows that there were highly significant main effects of phase for all of the measures that were administered, indicating cyclicity of symptom reporting. The cycle
by phase interactions were non-significant for all measures, indicating that women reported the same pattern of symptoms during each screening cycle. There was no main effect of cycle for the DSR, DSR-20 and STAIS, indicating that women reported the same level of symptoms during each screening cycle. However, there was a significant main effect of cycle for the BPAQ, BIS-11 and BDI. Inspection of the means revealed that women reported their lowest level of each of these symptoms during screening cycle three. Bonferroni pairwise comparisons revealed that women reported significantly lower levels of depression during the third screening cycle than during the first. This trend was also apparent for verbal (p=0.052) and physical (p=0.061) aggression. Women also reported significantly lower levels of hostility and anger during the second and third screening cycles than during the first. All women attended a study visit during the luteal phase of the third screening cycle, during which they were informed that they had met the study criteria and could continue to take part in the clinical trial. Therefore, this study visit may have influenced women’s symptom reporting during the luteal phase of screening cycle three.

7.4.3 Placebo effect

In order to investigate the placebo effect, two cycles of placebo treatment were administered to all women. If a placebo effect was present, a reduction in symptom reporting would be expected, particularly in the luteal phase. Because of the unknown effect of the study visit during the third screening cycle, the two non-visit screening cycles were used for comparison.

7.4.3.1 DSR (Freeman et al., 1996)

Figures 7.3 and 7.4 show the mean DSR total and subscale scores for the first five cycles during which participants reported their symptoms. Cycles 1-3 were the screening cycles, while during cycles 4 and 5, participants took placebo supplements.
A 2 x 2 x 4 RM ANOVA was performed on the mean total DSR scores to examine the placebo effect. The two study phases (screening, placebo run-in), two menstrual cycles and four cycle phases formed the within subjects IVs. There was no significant main effect of cycle (F(1, 33) = 1.58, p>0.05) and no significant study phase by cycle interaction (F(1, 33) = 0.95, p>0.05). Cyclicity of symptom reporting was indicated by the highly significant main effect of cycle phase (F(3, 31) = 56.20, p<0.001, partial $\eta^2$=0.85). Moreover, a placebo effect was indicated by the significant main effect of study phase (F(1, 33) = 12.16, p<0.01, partial $\eta^2$=0.27). Inspection of the means
revealed that women reported less severe symptoms during the placebo run-in (mean = 7.26), than during screening (mean = 9.09). There was a trend towards a study phase by cycle phase interaction (F(3, 32) = 2.38, p=0.089, partial η²=0.19). Figure 7.3 suggests that symptom reporting during the follicular and rest phases remained fairly similar between the screening and placebo run-in cycles, while symptom reporting appeared reduced during menses and the luteal phase. This was explored through a multivariate analysis of the DSR subscales.

A 2 (study phase) x 2 (cycle) x 4 (cycle phase) doubly multivariate ANOVA was performed on the DSR subscales. There was no significant main effect of cycle (multivariate F(4, 30) = 1.22, p>0.05) and no significant study phase by cycle interaction (multivariate F(4, 30) = 2.08, p>0.05). There was a significant main effect of cycle phase (multivariate F(12, 22) = 28.64, p<0.001, partial η²=0.94). A placebo effect was indicated by the significant main effect of study phase (multivariate F(4, 30) = 5.11, p<0.01, partial η²=0.41). Moreover, the significant study phase by cycle phase interaction revealed that the placebo effect was limited to specific cycle phases (multivariate F(12, 22) = 2.63, p<0.05, partial η²=0.59).

Given that there were some significant multivariate effects, the univariate tests of the main effect of study phase and of the study phase by cycle phase interaction were examined (Field, 2005; Tabachnick and Fidell, 2007a). These effects are summarized in Table 7.5.

Table 7.5 Multivariate and univariate statistics for the main effect of study phase, and for the study phase by cycle phase interaction for the DSR

<table>
<thead>
<tr>
<th>Study phase (Screening v placebo run-in)</th>
<th>Study phase by Cycle phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistic</td>
<td>Partial η²</td>
</tr>
<tr>
<td>Multivariate</td>
<td>F(4,30)=5.11**</td>
</tr>
<tr>
<td>Univariate:</td>
<td>0.41</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Statistic</th>
<th>Partial η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univariate:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Mood</td>
<td>F(1,33)=18.06**</td>
<td>0.35</td>
</tr>
<tr>
<td>-Behavioural</td>
<td>F(1,33)=3.07</td>
<td>0.085</td>
</tr>
<tr>
<td>-Pain</td>
<td>F(1,33)=3.90</td>
<td>0.11</td>
</tr>
<tr>
<td>-Physical</td>
<td>F(1,33)=4.19*</td>
<td>0.11</td>
</tr>
</tbody>
</table>

NB. *p<0.05 **p<0.01
Table 7.5 shows that there were significant univariate main effects of study phase on the mood and physical DSR subscales. The study phase by cycle phase interactions were significant for the pain and physical subscales. These interaction effects were explored by assessing the effect of study phase, separately at each cycle phase. This revealed that luteal phase reporting of pain symptoms was significantly lower during the placebo run-in than screening (t (33)=2.94, p<0.01), while physical symptoms were lower during the rest and luteal phases (smallest t (33)=3.49, p<0.01). Therefore, the placebo effect operates on mood symptoms throughout the cycle, but affects pain symptoms only premenstrually and physical symptoms during the second half of the cycle.

7.4.3.2 DSR-20

Figure 7.5 displays the mean DSR-20 subscale scores for the first five cycles during which participants reported their symptoms.

![Figure 7.5 Mean DSR-20 subscale scores during study cycles 1-5](image)

A 2 (study phase) x 2 (cycle) x 4 (cycle phase) doubly multivariate ANOVA was performed on the DSR20 subscale scores. There was no significant main effect of cycle (multivariate F(2, 32) = 2.48, p>0.05) and no significant study phase by cycle interaction (multivariate F(2, 32) = 0.63, p>0.05). There was a significant main effect of cycle phase (multivariate F(6, 28) = 27.66, p<0.001, partial η²=0.86). A placebo effect was indicated by the significant main effect of study phase (multivariate F(2, 32) = 6.71, p<0.01, partial η²=0.30). Moreover, the significant study phase by cycle phase interaction revealed that the placebo effect was limited to specific cycle phases (multivariate F(6, 28) = 3.61, p<0.01, partial η²=0.44). The univariate tests for the main
effect of study phase and for the study phase by cycle phase interaction are summarized in Table 7.6.

Table 7.6 Multivariate and univariate statistics for the main effect of study phase, and for the study phase by cycle phase interaction for the DSR-20

<table>
<thead>
<tr>
<th>Study phase by Cycle phase</th>
<th>Study phase (Screening v placebo run-in)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistic</td>
<td>Partial κ²</td>
</tr>
<tr>
<td>Multivariate</td>
<td>F(2, 32)=6.71</td>
</tr>
<tr>
<td>Univariate:</td>
<td></td>
</tr>
<tr>
<td>-Psychological</td>
<td>F(1, 33)=13.73**</td>
</tr>
<tr>
<td>-Physical</td>
<td>F(1, 33)=0.17</td>
</tr>
</tbody>
</table>

NB. *p<0.05 **p<0.01

Table 7.6 shows that there was a significant univariate main effect of study phase on the psychological subscale. The study phase by cycle phase interactions were significant for both subscales. A simple effects analysis (see section 5.4) revealed that women reported significantly lower psychological symptoms during the placebo run-in than during screening in the menstrual and luteal cycle phases (smallest t (33)=3.06, p<0.01). However, no statistically significant differences between these study phases were found for physical symptoms when familywise error (see section 5.4) was controlled for (t (33)=2.26, p=0.03).
7.4.3.3 State anxiety (STAIS)

Figure 7.6 displays the mean state anxiety scores that were reported during the first five study cycles.

![Figure 7.6 Mean STAIS scores during study cycles 1-5](image)

A 2 (study phase) x 2 (cycle) x 4 (cycle phase) RM ANOVA was performed on the mean STAIS scores. There was a significant main effect of cycle (F(1, 33) = 7.35, p<0.05, partial $\eta^2=0.18$). However, the study phase by cycle interaction was non-significant (F(1, 33) = 1.57, p>0.05). There was a highly significant main effect of cycle phase (F(3, 99) = 28.07, p<0.001, partial $\eta^2=0.46$). There was a trend towards an effect of study phase (F(1, 33) = 3.35, p=0.076, partial $\eta^2=0.092$). Inspection of the means revealed that women reported slightly lower anxiety levels during placebo run-in (mean = 41.40) than during screening (mean = 43.94). There was a significant study phase by cycle phase interaction (F(3, 99) = 9.44, p<0.001, partial $\eta^2=0.22$), indicating the presence of a placebo effect during specific cycle phases. Figure 7.6 suggests that symptom reporting during the follicular, rest and luteal phases remained fairly similar between the screening and placebo run-in cycles, while symptom reporting appeared reduced during menses. This was confirmed through a simple effects analysis, which showed that symptom reporting was only reduced from screening to the placebo run-in during menses (t (33)=4.36, p<0.001).
7.4.3.4 Depression (BDI)

Figure 7.7 displays the mean BDI scores that were reported during the first five study cycles.

![Mean BDI scores during study cycles 1-5](image)

A 2 (study phase) x 2 (cycle) x 4 (cycle phase) RM ANOVA was performed on the mean total BDI scores. There was a significant main effect of cycle (F(1, 33) = 8.26, p<0.01, partial η²=0.20). However, there was a non-significant study phase by cycle interaction (F(1, 33) = 0.18, p>0.05). There was a highly significant main effect of cycle phase (F(3, 99) = 21.32, p<0.001, partial η²=0.39). There was a non-significant main effect of study phase (F(1, 33) = 2.43, p>0.05). A placebo effect was indicated during specific cycle phases by the significant study phase by cycle phase interaction (F(3, 99) = 7.01, p<0.01, partial η²=0.18). Figure 7.7 suggests that symptom reporting during the follicular, rest and luteal phases remained fairly similar between the screening and placebo run-in cycles, while symptom reporting appeared reduced during menses. This was confirmed through a simple effects analysis, which revealed that symptom reporting was only reduced from screening to the placebo run-in during menses (t(33)=3.49, p=0.001).
7.4.3.5 Aggression (BPAQ)

Figures 7.8 and 7.9 show the mean BPAQ total and subscale scores for the first five cycles during which participants reported their symptoms.

Figure 7.8 Mean BPAQ scores during study cycles 1-5

A 2 (study phase) x 2 (cycle) x 4 (cycle phase) RM ANOVA was performed on the mean BPAQ total scores. There was a significant main effect of cycle ($F(1, 33) = 23.24$, $p<0.001$, partial $\eta^2=0.41$). However, the study phase by cycle interaction was non-significant ($F(1, 33) =0.34$, $p>0.05$). Cyclicity was demonstrated by the highly significant main effect of cycle phase ($F(3, 99) = 15.61$, $p<0.001$, partial $\eta^2=0.32$). A placebo effect was indicated by the significant main effect of study phase ($1, 33) = 7.80,
p<0.01. partial η²=0.19). Inspection of the means revealed that women reported lower aggression levels during the placebo run-in (mean = 57.97) than during screening (mean = 63.62). Moreover, the significant study phase by cycle phase interaction revealed that the placebo effect was specific to certain cycle phases (F(3, 99) = 4.34, p<0.01, partial η²=0.12). Figure 7.8 suggests that symptom reporting was most similar between screening and the placebo run-in during the rest and luteal phases, while symptom reporting appeared reduced during menses. This was explored through a multivariate analysis of the BPAQ subscales.

A 2 (study phase) x 2 (cycle) x 4 (cycle phase) doubly multivariate ANOVA was performed on the BPAQ subscales. There was a significant main effect of cycle (multivariate F(4, 30) = 5.77, p<0.01. partial η²=0.44). However, there was a non-significant study phase by cycle interaction (multivariate F(4, 30) = 0.14, p>0.05). There was a significant main effect of cycle phase (F(multivariate F(12, 22) = 6.64, p<0.001, partial η²=0.78). A placebo effect was indicated by the significant main effect of study phase (multivariate F(4, 30) = 3.70, p<0.05, partial η²=0.33). The non-significant study phase by cycle phase interaction revealed that the placebo effect was not specific to particular cycle phases for the linear combination of the BPAQ subscales (multivariate F(12, 22) = 1.86, p>0.05). The univariate tests for the main effect of study phase and for the study phase by cycle phase interaction are summarized in Table 7.7.

<table>
<thead>
<tr>
<th>Study phase</th>
<th>Study phase by Cycle phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistic</td>
<td>Partial η²</td>
</tr>
<tr>
<td>Multivariate</td>
<td>F(4, 40)=3.70*</td>
</tr>
<tr>
<td>Univariate:</td>
<td></td>
</tr>
<tr>
<td>- Verbal</td>
<td>F(1, 33)=7.43*</td>
</tr>
<tr>
<td>- Physical</td>
<td>F(1, 33)=1.06</td>
</tr>
<tr>
<td>- Hostility</td>
<td>F(1, 33)=11.68**</td>
</tr>
<tr>
<td>- Anger</td>
<td>F(1, 33)=6.02*</td>
</tr>
</tbody>
</table>

NB. *p<0.05 **p<0.01

Table 7.7 shows that there were significant univariate main effects of study phase on the verbal, hostility and anger BPAQ subscales. The study phase by cycle phase interactions were also significant for these subscales. A simple effects analysis revealed that the placebo effect operated on verbal aggression and anger during menses (smallest
t(33)=3.32, p=0.002) and on hostility during the bleed, follicular and luteal phases (smallest t(33)=2.81, p=0.008).

7.4.3.6 Impulsivity (BIS-11)

Figures 7.10 and 7.11 display the mean BIS-11 total and subscale scores reported during study cycles one to five.

A 2 (study phase) x 2 (cycle) x 4 (cycle phase) RM ANOVA was performed on the mean BIS-11 total scores. There was a significant main effect of cycle (F(1, 33) = 7.73, p<0.01, partial η²=0.19). However, there was a non-significant study phase by cycle
interaction \( F(1, 33) = 0.35, p>0.05 \). There was a significant main effect of cycle phase \( F(3, 99) = 18.67, p<0.001, \text{partial } \eta^2=0.36 \). There was a trend towards a main effect of study phase \( F(1, 33) = 3.47, p=0.072, \text{partial } \eta^2=0.095 \). Inspection of the means revealed that women reported slightly lower impulsivity levels during the placebo run-in (mean = 61.10) than during screening (mean = 62.78). A placebo effect was indicated during specific cycle phases by the significant study phase by cycle phase interaction \( F(3, 99) = 5.98, p<0.01, \text{partial } \eta^2=0.15 \). Figure 7.10 suggests that symptom reporting during the follicular, rest and luteal phases remained fairly similar between the screening and placebo run-in cycles, but was reduced during menses. This was explored through a multivariate analysis of the BIS-11 subscales.

A 2 (study phase) x 2 (cycle) x 4 (cycle phase) doubly multivariate ANOVA was performed on the BIS-11 subscale scores. There was a significant main effect of cycle (multivariate \( F(3, 31) = 6.27, p<0.01, \text{partial } \eta^2=0.38 \)) and a significant study phase by cycle interaction (multivariate \( F(3, 31) = 3.03, p<0.05, \text{partial } \eta^2=0.23 \)). Cyclicity was demonstrated by the significant main effect of cycle phase (multivariate \( F(9, 25) = 5.23, p<0.001, \text{partial } \eta^2=0.65 \)). There was not a significant main effect of study phase (multivariate \( F(3, 31) = 1.27, p>0.05 \)). However, a placebo effect was indicated during specific cycle phases by the significant study phase by cycle phase interaction (multivariate \( F(9, 25) = 2.72, p<0.05, \text{partial } \eta^2=0.50 \)). The univariate tests of the main effect of study phase and of the study phase by cycle phase interaction are summarized in Table 7.8.

### Table 7.8 Multivariate and univariate statistics for the main effect of study phase, and for the study phase by cycle phase interaction for the BIS-11 subscales

<table>
<thead>
<tr>
<th>Study phase</th>
<th>Statistic</th>
<th>Partial ( \eta^2 )</th>
<th>Study phase by Cycle phase</th>
<th>Statistic</th>
<th>Partial ( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multivariate</td>
<td>( F(3, 31) = 1.27 )</td>
<td>0.11</td>
<td>( F(9, 25) = 2.72^* )</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>Univariate:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Attentional</td>
<td>( F(1, 33) = 1.98 )</td>
<td>0.057</td>
<td>( F(3, 99) = 2.92^* )</td>
<td>0.081</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( F(1, 33) = 3.22 )</td>
<td>0.089</td>
<td>( F(3, 99) = 1.06 )</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td>- Motor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( F(1, 33) = 1.89 )</td>
<td>0.054</td>
<td>( F(3, 99) = 7.25^{**} )</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>- Nonplanning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NB. \(^*p<0.05\)^ \(^{**}p<0.01\)

Table 7.8 shows that there were no significant univariate main effects of study phase on any BIS-11 subscale. However, the study phase by cycle phase interactions were significant for the attentional and nonplanning subscales. A simple effects analysis revealed that a placebo effect operated on nonplanning impulsivity during menses.
Although this trend was also apparent for attentional impulsivity, this did not reach significance when familywise type I error rate (see section 5.4) was controlled for ($t(33)=2.59, p=0.014$).

### 7.4.3.7 Summary

Table 7.9 displays the univariate/multivariate statistics for the main effects of cycle phase and study phase, and for the study phase by cycle phase interactions for each of the measures that were administered to clinical trial participants during study cycles 1-5.

#### Table 7.9 Summary table of placebo effects

<table>
<thead>
<tr>
<th>Measure</th>
<th>Cycle Phase</th>
<th>Study Phase</th>
<th>Study Phase by Cycle Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSR: Total</td>
<td>$F(3, 31)=56.20^{**}$</td>
<td>$F(1, 33)=12.16^{**}$</td>
<td>$F(3, 32)=2.38$</td>
</tr>
<tr>
<td>Subscales (MV)</td>
<td>$F(12, 22)=28.64^{**}$</td>
<td>$F(4, 30)=5.11^{**}$</td>
<td>$F(12, 22)=2.63^*$</td>
</tr>
<tr>
<td>Univariate:</td>
<td>$F(3, 99)=38.28^{**}$</td>
<td>$F(1, 33)=18.06^{**}$</td>
<td>$F(3, 99)=2.39$</td>
</tr>
<tr>
<td>-Mood</td>
<td>$F(3, 99)=32.54^{**}$</td>
<td>$F(1, 33)=3.07$</td>
<td>$F(3, 99)=0.86$</td>
</tr>
<tr>
<td>-Behavioural</td>
<td>$F(3, 99)=82.68^{**}$</td>
<td>$F(1, 33)=3.90$</td>
<td>$F(3, 99)=3.50^*$</td>
</tr>
<tr>
<td>-Physical</td>
<td>$F(3, 99)=94.28^{**}$</td>
<td>$F(1, 33)=4.19^*$</td>
<td>$F(3, 99)=7.02^{**}$</td>
</tr>
<tr>
<td>DSR-20 (MV)</td>
<td>$F(6, 28)=27.66^{**}$</td>
<td>$F(2, 32)=6.71^{**}$</td>
<td>$F(6, 28)=3.61^{**}$</td>
</tr>
<tr>
<td>Univariate:</td>
<td>$F(3, 99)=15.20^{**}$</td>
<td>$F(1, 33)=13.73^{**}$</td>
<td>$F(3, 99)=3.32^*$</td>
</tr>
<tr>
<td>-Psychological</td>
<td>$F(3, 99)=53.07^{**}$</td>
<td>$F(1, 33)=0.17$</td>
<td>$F(3, 99)=3.14^*$</td>
</tr>
<tr>
<td>-Physical</td>
<td>$F(3, 99)=94.28^{**}$</td>
<td>$F(1, 33)=4.19^*$</td>
<td>$F(3, 99)=7.02^{**}$</td>
</tr>
<tr>
<td>BPAQ: Total</td>
<td>$F(3, 99)=15.61^{**}$</td>
<td>$F(1, 33)=7.80^{**}$</td>
<td>$F(3, 99)=4.34^{**}$</td>
</tr>
<tr>
<td>Subscales (MV):</td>
<td>$F(12, 22)=6.64^{**}$</td>
<td>$F(4, 30)=3.70^*$</td>
<td>$F(12, 22)=1.86$</td>
</tr>
<tr>
<td>Univariate:</td>
<td>$F(3, 99)=14.51^{**}$</td>
<td>$F(1, 33)=7.43^*$</td>
<td>$F(3, 99)=4.12^{**}$</td>
</tr>
<tr>
<td>-Verbal</td>
<td>$F(3, 99)=7.24^{**}$</td>
<td>$F(1, 33)=1.06$</td>
<td>$F(3, 99)=2.71$</td>
</tr>
<tr>
<td>-Physical</td>
<td>$F(3, 99)=9.43^{**}$</td>
<td>$F(1, 33)=11.68^{**}$</td>
<td>$F(3, 99)=3.47^*$</td>
</tr>
<tr>
<td>-Hostility</td>
<td>$F(3, 99)=21.29^{**}$</td>
<td>$F(1, 33)=6.02^{**}$</td>
<td>$F(3, 99)=4.03^{**}$</td>
</tr>
<tr>
<td>BIS-11: Total</td>
<td>$F(3, 99)=18.67^{**}$</td>
<td>$F(1, 33)=3.47^*$</td>
<td>$F(3, 99)=5.98^{**}$</td>
</tr>
<tr>
<td>Subscales (MV)</td>
<td>$F(9.25)=5.23^{**}$</td>
<td>$F(3, 31)=1.27$</td>
<td>$F(9.25)=2.72^*$</td>
</tr>
<tr>
<td>Univariate:</td>
<td>$F(3, 99)=20.09^{**}$</td>
<td>$F(1, 33)=1.98$</td>
<td>$F(3, 99)=2.92^*$</td>
</tr>
<tr>
<td>-Attentional</td>
<td>$F(3, 99)=1.49$</td>
<td>$F(1, 33)=3.22$</td>
<td>$F(3, 99)=1.06$</td>
</tr>
<tr>
<td>-Motor</td>
<td>$F(3, 99)=18.52^{**}$</td>
<td>$F(1, 33)=1.89$</td>
<td>$F(3, 99)=7.25^{**}$</td>
</tr>
<tr>
<td>-Non-planning</td>
<td>$F(3, 99)=28.07^{**}$</td>
<td>$F(1, 33)=3.35^*$</td>
<td>$F(3, 99)=9.44^{**}$</td>
</tr>
<tr>
<td>STAIS</td>
<td>$F(3, 99)=21.32^{**}$</td>
<td>$F(1, 33)=2.43$</td>
<td>$F(3, 99)=7.01^{**}$</td>
</tr>
</tbody>
</table>

*p<0.05 **p<0.01 `trend (p<0.08)
7.4.4 How long do placebo effects last?

Figures 7.3 to 7.11 suggest that symptom reporting had not returned to screening levels after two cycles of placebo treatment. In approximately half of the sample, who received treatment order two (n=16), a further two cycles of placebo followed the placebo run-in, providing a total of four continuous placebo treated cycles. Figures 7.12 to 7.17 display the symptom profiles that treatment group two reported during the first seven study cycles. Cycles 1-3 were the screening cycles, cycles 4 and 5 the placebo run-in cycles, and cycles 6 and 7 the placebo treatment cycles.

![Figure 7.12 Mean DSR scores reported by treatment group two during study cycles 1-7](image1)

![Figure 7.13 Mean DSR-20 scores reported by treatment group two during study cycles 1-7](image2)
Figure 7.14 Mean STAIS scores reported by treatment group two during study cycles 1-7

Figure 7.15 Mean BDI scores reported by treatment group two during study cycles 1-7
Figure 7.16 Mean BPAQ scores reported by treatment group two during study cycles 1-7.

Figure 7.17 Mean BIS-11 scores reported by treatment group two during study cycles 1-7.
In order to assess whether symptom reporting had returned to screening levels during study cycles 6 and 7, the symptom profiles treatment group two reported during placebo treatment were compared with those that they reported during screening. Total scale scores were analysed through the use of 2 (study phase) x 2 (cycle) x 4 (phase) RM ANOVAs, while 2 (study phase) x 2 (cycle) x 4 (phase) doubly multivariate ANOVAs were used to analyse subscale scores. Table 7.10 displays the multivariate and univariate statistics for the main effect of study phase and for the study phase by cycle phase and study phase by cycle interactions, for each of the measures that were administered.

Table 7.10 The placebo effect during study cycles 6-7: Multivariate and univariate statistics for treatment group 2 (n=16)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Placebo Treatment Vs. Screening</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Study Phase</td>
</tr>
<tr>
<td>DSR (Total)</td>
<td>F(1,15)=5.26</td>
</tr>
<tr>
<td>Subscales</td>
<td>MV F(4, 12)=2.24</td>
</tr>
<tr>
<td>-Mood</td>
<td>F(1,15)=9.50***</td>
</tr>
<tr>
<td>-Behavioural</td>
<td>F(1,15)=1.31</td>
</tr>
<tr>
<td>-Pain</td>
<td>F(1,15)=1.72</td>
</tr>
<tr>
<td>-Physical</td>
<td>F(1,15)=4.04</td>
</tr>
<tr>
<td>DSR-20 subscales</td>
<td>MV F(2, 14)=5.10*</td>
</tr>
<tr>
<td>-Psychological</td>
<td>F(1,15)=9.95**</td>
</tr>
<tr>
<td>-Physical</td>
<td>F(1,15)=0.063</td>
</tr>
<tr>
<td>BPAQ (Total)</td>
<td>F(1,15)=12.32**</td>
</tr>
<tr>
<td>Subscales</td>
<td>MV F(4, 12)=3.46*</td>
</tr>
<tr>
<td>-Verbal</td>
<td>F(1,15)=3.88</td>
</tr>
<tr>
<td>-Physical</td>
<td>F(1,15)=6.26*</td>
</tr>
<tr>
<td>-Hostility</td>
<td>F(1,15)=11.60**</td>
</tr>
<tr>
<td>-Anger</td>
<td>F(1,15)=5.59*</td>
</tr>
<tr>
<td>BIS-11 (Total)</td>
<td>F(1,15)=0.26</td>
</tr>
<tr>
<td>Subscales</td>
<td>MV F(3, 13)=1.23</td>
</tr>
<tr>
<td>-Attentional</td>
<td>F(1,15)=0.96</td>
</tr>
<tr>
<td>-Non-planning</td>
<td>F(1,15)=0.24</td>
</tr>
<tr>
<td>-Motor</td>
<td>F(1,15)=0.74</td>
</tr>
<tr>
<td>STAIS</td>
<td>F(1,15)=2.68</td>
</tr>
<tr>
<td>BDI</td>
<td>F(1,15)=2.80</td>
</tr>
</tbody>
</table>

7.4.4.1 DSR

Table 7.10 indicates that treatment group two reported significantly fewer mood symptoms during placebo treatment (cycles 6-7) than during screening (F(1, 15)= 9.50, p<0.01, partial \( \eta^2=0.39 \)), while this trend was shown for physical symptoms (F(1, 15) = 4.04, p=0.063, partial \( \eta^2=0.21 \)). These effects were limited to particular cycle phases (smallest F(3, 45) = 2.93, p<0.05, partial \( \eta^2=0.16 \)). A simple effects analysis revealed...
that women reported significantly fewer mood symptoms during placebo treatment than during screening in the menstrual and luteal phases (smallest $t(15)=3.40$, $p=0.004$), and showed a trend towards reporting fewer physical symptoms during the rest and luteal phases (largest $t(15)=3.04$, $p=0.008$). Therefore, the placebo effects that were observed for mood and physical symptoms (see section 7.4.3.1) remained after four placebo treatment cycles, while the placebo effect operating on pain symptoms had abated by the time women entered the placebo treatment cycles.

7.4.4.2 DSR-20

The placebo effect was found to affect reports of psychological symptoms during the menstrual and luteal phases (see section 7.4.3.2). The DSR-20 profiles reported by treatment group two suggest that these effects remained after four cycles of placebo treatment (Figure 7.13). As Table 7.10 indicates, this treatment group reported significantly fewer psychological symptoms during placebo treatment than during screening ($F(1, 15)=9.95$, $p<0.01$, partial $\eta^2=0.40$). This effect was limited to particular cycle phases ($F(3, 45)=5.20$, $p<0.01$, partial $\eta^2=0.26$). A simple effects analysis revealed that psychological symptom reporting remained reduced during the menstrual and luteal phases (smallest $t(15)=3.62$, $p=0.003$).

7.4.4.3 STAIS

The placebo effect was found to affect reports of anxiety during menses (see section 7.4.3.3). The STAIS symptom profiles reported by treatment group two suggest that menstrual anxiety levels had returned to screening levels by the time women entered the placebo treatment phase (Figure 7.14). This was confirmed through the non-significant effect of study phase ($F(1, 15)=2.68$, $p>0.05$, partial $\eta^2=0.15$) and through the non-significant study phase by cycle phase interaction ($F(3, 45)=2.16$, $p>0.05$, partial $\eta^2=0.13$). The significant study phase by cycle interaction ($F(1, 15)=5.86$, $p<0.05$, partial $\eta^2=0.28$) may be explained by the lower anxiety levels that were reported during the third placebo treatment cycle (Figure 7.14).

7.4.4.4 BDI

The placebo effect was found to operate on depression during menses (see section 7.4.3.4). Figure 7.15 suggests that this effect had subsided by the time women entered the placebo treatment phase. This was confirmed through the non-significant effect of study phase ($F(1, 15)=2.80$, $p>0.05$, partial $\eta^2=0.16$), and through the non-significant
study phase by cycle phase interaction ($F(3, 45) = 1.64, p>0.05$, partial $\eta^2=0.098$). The significant study phase by cycle interaction ($F(1, 15) = 6.66, p<0.05$, partial $\eta^2=0.31$) may be explained by the lower depression levels that were reported during the third placebo treatment cycle (Figure 7.15).

### 7.4.4.5 BPAQ

The placebo effect was found to affect reports of verbal aggression and anger during menses and hostility during the menstrual, follicular and luteal phases (see section 7.4.3.5). The BPAQ profiles reported by treatment group two suggest that their symptom reporting of verbal aggression, anger and hostility was reduced across the third placebo treatment cycle, but had returned to screening levels during the fourth (Figure 7.16). As Table 7.10 indicates, this treatment group reported significantly lower levels of hostility and anger during placebo treatment than during screening (smallest $F(1, 15) = 5.59, p<0.05$, partial $\eta^2=0.27$), and showed this trend for verbal aggression ($F(1, 15)= 3.88, p=0.068$, partial $\eta^2=0.21$). These effects were not limited to specific cycle phases (largest $F(3, 45)=1.72, p>0.05$, partial $\eta^2=0.10$). There was a significant study phase by cycle interaction for verbal aggression, hostility and anger. A simple effects analysis revealed that women reported significantly lower levels of verbal aggression, hostility and anger during placebo treatment than screening during the third placebo treatment cycle (smallest $F(1, 15)=10.66, p<0.01$, partial $\eta^2=0.42$), but not the fourth (largest $F(1, 15)=2.38, p>0.05$, partial $\eta^2=0.14$).

### 7.4.4.6 BIS-11

The placebo effect was found to operate on non-planning impulsivity during menses. This trend was also shown for attentional impulsivity (see section 7.4.3.6). The BIS-11 profiles reported by treatment group two suggest that these effects had subsided by the fourth treatment cycle (Figure 7.17). As Table 7.10 indicates, there was a non-significant effect of study phase (largest $F(1, 15) = 0.96, p>0.05$) and a non-significant study phase by cycle phase interaction (largest $F(3, 45) = 1.45, p>0.05$) for each subscale. However, there was a significant study phase by cycle interaction for the attentional subscale ($F(1, 15) = 14.15, p<0.01$, partial $\eta^2=0.49$). A simple effects analysis revealed that women reported significantly lower levels of attentional impulsivity during placebo treatment than screening during the third placebo treatment cycle ($F(1, 15) = 7.70, p=0.014$, partial $\eta^2=0.34$), but not the fourth ($F(1, 15) = 0.33, p=0.57$, partial $\eta^2=0.022$).
7.4.5 Placebo responders
The placebo effect could be the result of particular women responding more strongly than others to taking placebo supplements. Researchers advocate the importance of eliminating these women from treatment analyses (Endicott, 1995; Halbreich and Endicott, 1985; Rapkin, 2005). Therefore, some researchers (e.g. Freeman et al., 1999; Freeman and Rickels, 1999; Halbreich et al., 2002; Pearlstein et al., 1997) have used one cycle of placebo run-in to eliminate women from their treatment analysis on the basis that they no longer meet their PMS diagnostic entry criteria during the placebo run-in cycle. Two studies have used two cycles of placebo run-in and also eliminated women on this basis (Freeman et al., 1990; Steiner et al., 1995). However, these researchers did not specify whether women were excluded if they no longer met their entry criteria during only one placebo run-in cycle or whether women were required to no longer meet the criteria in both placebo run-in cycles. In this study, two women no longer met the 30% increase criterion (see section 3.5.2.1) during both placebo run-in cycles. Therefore, these women were excluded from the treatment analyses. This resulted in a similar percentage (5.88%) of exclusions as has been reported in previous research (e.g. Freeman et al., 1999; Pearlstein et al., 1997; Steiner et al., 1995).

7.4.6 Success of randomisation
If women who were assigned a particular treatment order (e.g. SJW, placebo treatment) differed in their symptom profiles during screening from those women who were assigned to the other treatment order (e.g. placebo treatment, SJW), this could impact on the outcome of the analysis of the effectiveness of SJW. Hence, it was important to examine whether there were any differences in symptom profiles during the screening cycles between the treatment order groups.

Women who were assigned to the first treatment order (SJW, placebo) were compared to those assigned to the second (placebo, SJW) on all measures, through the use of 4 (phase) x 2 (order) RM ANOVAs and doubly multivariate ANOVAs. Table 7.11 displays the multivariate and univariate statistics for the main effects of phase and treatment order, and for the phase by treatment order interactions that were produced.
Table 7.11 Success of randomisation: Multivariate and univariate statistics for the main effects of phase and treatment order, and for the phase by treatment order interactions

<table>
<thead>
<tr>
<th>Measure</th>
<th>Phase Statistic</th>
<th>Partial η²</th>
<th>Order Statistic</th>
<th>Partial η²</th>
<th>Phase by Order Statistic</th>
<th>Partial η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSR: Total</td>
<td>F(3, 90)=78.28**</td>
<td>0.72</td>
<td>F(1, 30)=0.90</td>
<td>0.029</td>
<td>F(3, 90)=2.65</td>
<td>0.081</td>
</tr>
<tr>
<td>DSR: Subscales</td>
<td>F(12, 19)=33.76**</td>
<td>0.96</td>
<td>F(4, 27)=1.10</td>
<td>0.14</td>
<td>F(12, 19)=1.33</td>
<td>0.46</td>
</tr>
<tr>
<td>DSR-20</td>
<td>F(6, 25)=22.23**</td>
<td>0.84</td>
<td>F(2, 29)=0.41</td>
<td>0.027</td>
<td>F(6, 25)=1.49</td>
<td>0.26</td>
</tr>
<tr>
<td>STAIS</td>
<td>F(3, 90)=28.91**</td>
<td>0.49</td>
<td>F(1, 30)=0.014</td>
<td>&lt;0.001</td>
<td>F(3, 90)=0.17</td>
<td>0.006</td>
</tr>
<tr>
<td>BDI</td>
<td>F(3, 90)=23.17**</td>
<td>0.44</td>
<td>F(1, 30)=0.22</td>
<td>0.007</td>
<td>F(3, 90)=0.14</td>
<td>0.005</td>
</tr>
<tr>
<td>BPAQ: Total</td>
<td>F(3, 90)=15.86**</td>
<td>0.35</td>
<td>F(1, 30)=0.13</td>
<td>0.004</td>
<td>F(3, 90)=0.43</td>
<td>0.014</td>
</tr>
<tr>
<td>BPAQ: Subscales</td>
<td>F(12, 19)=4.46**</td>
<td>0.74</td>
<td>F(4, 27)=0.073</td>
<td>0.011</td>
<td>F(12, 19)=0.86</td>
<td>0.35</td>
</tr>
<tr>
<td>BIS: Total</td>
<td>F(3, 90)=22.84**</td>
<td>0.43</td>
<td>F(1, 30)=0.004</td>
<td>&lt;0.001</td>
<td>F(3, 90)=0.11</td>
<td>0.004</td>
</tr>
<tr>
<td>BIS: Subscales</td>
<td>F(9, 22)=8.04***</td>
<td>0.77</td>
<td>F(3, 28)=0.25</td>
<td>0.026</td>
<td>F(9, 22)=0.53</td>
<td>0.18</td>
</tr>
</tbody>
</table>

* p<0.01  **p<0.001

Table 7.11 shows that there was a significant main effect of phase for each measure, indicating cyclicity of symptom reporting. However, all main effects of treatment order and phase by treatment order interactions were non-significant, indicating that randomisation to treatment order did not result in women with different profiles at screening receiving different treatment orders.

**7.4.7 Effectiveness of SJW for PMS**

The effectiveness of SJW for PMS was investigated through the comparison of the DSR (Freeman et al., 1996) data collected during SJW and placebo treatment, during the four cycle phases of the two treatment cycles. The total scale scores were analysed through the use of a 2 (treatment) x 2 (cycle) x 4 (phase) x 2 (treatment order) repeated measures ANOVA, while the subscale scores (mood, behavioural, pain and physical) were analysed through the use of a 2 (treatment) x 2 (cycle) x 4 (phase) x 2 (treatment order) doubly multivariate ANOVA.

Figures 7.18 to 7.22 display the mean DSR total and subscale scores that were reported during the two cycles of SJW and placebo treatment.
Figure 7.18 Effect of treatment on total DSR symptoms during each treatment cycle
The 2 (treatment) x 2 (cycle) x 4 (phase) x 2 (treatment order) RM ANOVA that was conducted on the total DSR scores revealed a non-significant main effect of treatment order (F(1, 30) = 0.96, p>0.05), and a non-significant treatment by treatment order interaction (F(1, 30) = 1.01, p>0.05). There was a significant main effect of treatment (F(1, 30) = 4.82, p<0.05, partial $\eta^2=0.14$), such that women reported significantly fewer symptoms during SJW (mean = 5.80) than placebo (mean = 6.58) treatment. The treatment by cycle (F(1, 30) = 0.73, p>0.05) and treatment by cycle phase (F(3, 90) = 1.16, p>0.05) interactions were non-significant.
The 2 (treatment) x 2 (cycle) x 4 (phase) x 2 (treatment order) doubly multivariate ANOVA that was performed on the DSR subscale scores revealed a non-significant main effect of treatment order (multivariate $F(4, 27) = 0.33, p>0.05$), and a non-significant treatment by treatment order interaction (multivariate $F(4, 27) = 2.83, p>0.05$). There was a non-significant main effect of treatment (multivariate $F(4, 27) = 1.99, p>0.05$) and non-significant treatment by cycle (multivariate $F(4, 27) = 1.27, p>0.05$) and treatment by cycle phase (multivariate $F(12, 19) = 0.66, p>0.05$) interactions.

When a multivariate analysis is conducted, univariate tests are produced for each of the separate subscales. It is recommended that these univariate tests should only be consulted in the presence of a significant multivariate effect (Tabachnick and Fidell, 2007a). However, there is a lack of previous research that has assessed the efficacy of SJW for PMS and it may be reasonable to hypothesize that SJW affects specific components of PMS, rather than total or global scores. Hence it was considered appropriate to explore the univariate tests of the main effect of treatment and the treatment by cycle and treatment by cycle phase interactions (Table 7.12).

Table 7.12 Multivariate and univariate statistics for the main effect of treatment, and for the treatment by cycle and treatment by cycle phase interactions for the DSR treatment analysis

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Treatment by cycle</th>
<th>Treatment by cycle phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistic</td>
<td>Partial $\eta^2$</td>
<td>Statistic</td>
</tr>
<tr>
<td>Actual Scores Multivariate</td>
<td>$F(4, 27)=1.99$</td>
<td>0.23</td>
</tr>
<tr>
<td>Univariate:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Mood</td>
<td>$F(1, 30)=2.33$</td>
<td>0.072</td>
</tr>
<tr>
<td>-Behavioural</td>
<td>$F(1, 30)=5.07^*$</td>
<td>0.15</td>
</tr>
<tr>
<td>-Pain</td>
<td>$F(1, 30)=0.20$</td>
<td>0.006</td>
</tr>
<tr>
<td>-Physical</td>
<td>$F(1, 30)=7.00^*$</td>
<td>0.19</td>
</tr>
</tbody>
</table>

NB. $**p<0.01$ *$p<0.05$

Table 7.12 shows that there were significant univariate main effects of treatment on the behavioural and physical DSR subscales, such that women reported significantly fewer symptoms on SJW than placebo treatment. The treatment by cycle and treatment by cycle phase interactions were non-significant for each subscale.
7.4.8 Secondary analyses

7.4.8.1 Other analytical strategies applied to the DSR (Freeman et al., 1996)

Researchers conducting placebo controlled PMS trials determine efficacy through the comparison of the symptom profiles reported on active treatment with those reported during placebo treatment (e.g. Collins et al., 1993; De Souza et al., 2000; Doll et al., 1989; Hagen et al., 1985; Halbreich et al., 2002; Hicks et al., 2004; Schellenberg, 2001). However, a variety of outcome measures are used for this purpose. Some assess premenstrual symptom scores (e.g. Halbreich et al., 2002; Walker et al., 1998), while others assess follicular to luteal change scores (e.g. Doll et al., 1989). Some use both forms of analysis (e.g. Collins et al., 1993; De Souza et al., 2000; Freeman et al., 1990). However, many simply analyse the change in symptoms that are reported during the luteal phase of the last treatment cycle from those reported during the luteal phase at baseline (e.g. Atmaca et al., 2003; Hagen et al., 1985; Hicks et al., 2004; Schellenberg, 2001). The following, commonly used analytical approaches were performed on the DSR scores, and compared with each other and with the primary analysis presented in Section 7.4.7 (actual scores), to investigate how the outcome from a PMS treatment trial might be interpreted differently by manipulating the data in various ways.

i. **Premenstrual scores**: This involved the comparison of the data collected during the premenstrual phases of SJW and placebo treatment during each treatment cycle. Total scale scores in the premenstrual phases were analysed through the use of 2 (treatment) x 2 (cycle) x 2 (treatment order) RM ANOVAs, while subscale scores were analysed through the use of 2 (treatment) x 2 (cycle) x 2 (treatment order) doubly multivariate ANOVAs.

ii. **Follicular to luteal change scores**: This involved the comparison of the change in symptoms reported between the follicular and luteal phases during each of the SJW and placebo treatment cycles. Total scale scores were analysed through the use of 2 (treatment) x 2 (cycle) x 2 (treatment order) RM ANOVAs, while subscale scores were analysed through the use of 2 (treatment) x 2 (cycle) x 2 (treatment order) doubly multivariate ANOVAs.

iii. **Follicular to luteal change scores (2nd treatment cycle)**: Researchers have shown that the beneficial effect of SSRIs for PMS treatment occurs rapidly (Dimmock et al., 2000; Landen and Eriksson, 2003; Yonkers et al., 2008), with
efficacy demonstrated during the luteal phase of the first treatment cycle when
dosing commenced (Halbreich et al., 2002; Pearlstein et al., 1997; Steiner et al.,
1995). There is some evidence to suggest that this should also be the case for
SJW (Hicks et al., 2004), but this is unlikely to be detected in the follicular
phase of the first treatment cycle when symptom levels are low. Therefore, the
follicular to luteal change scores from the second treatment cycle were assessed
through the use of 2 (treatment) x 2 (treatment order) RM ANOVAs and doubly
multivariate ANOVAs.

iv. Change from Baseline (premenstrual scores): Change from baseline scores
were computed for SJW and placebo treatment by deducting the treatment score
reported during the premenstrual phase of the second treatment cycle from the
average premenstrual screening score. The change from baseline scores
reported on SJW and placebo treatment were compared through the use of 2
(treatment) x 2 (treatment order) RM ANOVAs and doubly multivariate
ANOVAs.

7.4.8.1.1 Premenstrual scores
A 2 (treatment) x 2 (cycle) x 2 (treatment order) RM ANOVA was conducted on the
total DSR scores that were reported during the premenstrual phase. There was a non-
significant main effect of treatment order (F(1, 30) = 1.14, p>0.05) and a non-significant
treatment by treatment order interaction (F(1, 30) = 0.25, p>0.05). There was a
significant main effect of treatment (F(1, 30) = 6.18, p<0.05, partial η²=0.17).
Moreover, there was a significant treatment by cycle interaction (F(1, 30) = 6.59,
p<0.05, partial η²=0.18). Although women reported fewer symptoms on SJW than on
placebo treatment during the luteal phase of each treatment cycle, there appeared a
much larger difference during treatment cycle two (Figure 7.18). This was confirmed
through a simple effects analysis, which revealed that there was only a significant effect
of treatment during the luteal phase of the second treatment cycle (t(31)= -3.18,
p=0.003).

A 2 (treatment) x 2 (cycle) x 2 (treatment order) doubly multivariate ANOVA was
conducted on the premenstrual DSR subscale scores. There was a non-significant main
effect of treatment order (multivariate F(4, 27) = 1.19, p>0.05) and a non-significant
treatment by treatment order interaction (multivariate F(4, 27) = 1.37, p>0.05). There
was a significant main effect of treatment (multivariate $F(4, 27) = 3.18$, $p<0.05$, partial $\eta^2=0.32$). Moreover, there was a trend towards a treatment by cycle interaction (multivariate $F(4, 27) = 2.66$, $p=0.054$, partial $\eta^2=0.28$). The univariate tests of the main effect of treatment and the treatment by cycle interaction are displayed in Table 7.13.

Table 7.13 Multivariate and univariate statistics for the main effect of treatment and treatment by cycle interaction for the DSR analysis of premenstrual scores

<table>
<thead>
<tr>
<th></th>
<th>Treatment Statistic</th>
<th>Partial $\eta^2$</th>
<th>Treatment by cycle Statistic</th>
<th>Partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Premenstrual</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariate</td>
<td>$F(4, 27)=3.18^*$</td>
<td>0.17</td>
<td>$F(4, 27)=2.66^*$</td>
<td>0.28</td>
</tr>
<tr>
<td>Univariate:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Mood</td>
<td>$F(1, 30)=0.54$</td>
<td>0.018</td>
<td>$F(1, 30)=8.79^{**}$</td>
<td>0.23</td>
</tr>
<tr>
<td>-Behavioural</td>
<td>$F(1, 30)=5.97^*$</td>
<td>0.17</td>
<td>$F(1, 30)=5.64^*$</td>
<td>0.16</td>
</tr>
<tr>
<td>-Pain</td>
<td>$F(1, 30)=0.99$</td>
<td>0.032</td>
<td>$F(1, 30)=3.68^*$</td>
<td>0.11</td>
</tr>
<tr>
<td>-Physical</td>
<td>$F(1, 30)=4.19^*$</td>
<td>0.12</td>
<td>$F(1, 30)=0.075$</td>
<td>0.002</td>
</tr>
</tbody>
</table>

NB. **$p<0.01$ *$p<0.05$ ~$p<0.08

Table 7.13 shows that there were significant univariate main effects of treatment on the behavioural and physical DSR subscales, such that women reported significantly fewer symptoms on SJW than placebo treatment. The treatment by cycle interactions were significant for the mood and behavioural subscales. Mood and behavioural symptom reporting during the luteal phase of the first treatment cycle appeared similar on SJW and placebo treatment, with greater separation occurring during the second cycle (Figure 7.19 and 7.20). A simple effects analysis revealed that there was not a significant effect of treatment during the first treatment cycle on either of these subscales (largest $t(31)=1.19$, $p=0.24$), but there was a significant treatment effect on mood and behavioural symptoms during the second cycle (smallest $t(31)=-2.45$, $p=0.019$).

7.4.8.1.2 **Follicular to luteal difference scores**

A $2 \times 2 \times 2$ (treatment) x (cycle) x (treatment order) RM ANOVA was conducted on the follicular to luteal DSR change scores. The main effect of treatment order ($F(1, 30) = 0.27$, $p>0.05$) and the treatment by treatment order interaction ($F(1, 30) = 0.51$, $p>0.05$) were not significant. There was a non-significant main effect of treatment ($F(1, 30) = 0.50$, $p>0.05$). Moreover, there was a non-significant treatment by cycle interaction ($F(1, 30) = 2.08$, $p>0.05$).
A 2 (treatment) x 2 (cycle) x 2 (treatment order) doubly multivariate ANOVA was conducted on the follicular to luteal DSR subscale change scores. There was a non-significant main effect of treatment order (multivariate F(4, 27) = 1.56, p>0.05) and a non-significant treatment by treatment order interaction (multivariate F(4, 27) = 0.28, p>0.05). There was a non-significant main effect of treatment (multivariate F(4, 27) = 0.95, p>0.05). Moreover, there was a non-significant treatment by cycle interaction (multivariate F(4, 27) = 1.01, p>0.05). The univariate tests of the main effect of treatment and the treatment by cycle interaction are displayed in Table 7.14.

**Table 7.14 Multivariate and univariate statistics for the main effect of treatment and treatment by cycle interaction for the DSR analysis of follicular to luteal change scores**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Partial ( \eta^2 )</th>
<th>Treatment by cycle</th>
<th>Partial ( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Follicular to luteal change Multivariate</td>
<td>F(4, 27)=0.95</td>
<td>0.12</td>
<td>F(4, 27)=1.01</td>
</tr>
<tr>
<td>Univariate: Mood</td>
<td>F(1, 30)=0.49</td>
<td>0.016</td>
<td>F(1, 30)=3.28</td>
</tr>
<tr>
<td>Univariate: Behavioural</td>
<td>F(1, 30)=1.07</td>
<td>0.034</td>
<td>F(1, 30)=2.25</td>
</tr>
<tr>
<td>Univariate: Pain</td>
<td>F(1, 30)=0.43</td>
<td>0.014</td>
<td>F(1, 30)=1.94</td>
</tr>
<tr>
<td>Univariate: Physical</td>
<td>F(1, 30)=0.84</td>
<td>0.027</td>
<td>F(1, 30)=0.31</td>
</tr>
</tbody>
</table>

NB. **p<0.01 *p<0.05 -p<0.08

Inspection of the univariate effects revealed that there was a non-significant main effect of treatment for each subscale. Moreover, the treatment by cycle interactions were also non-significant (Table 7.14).

**7.4.8.1.3 Follicular to luteal difference scores (2\textsuperscript{nd} treatment cycle)**

A 2 (treatment) x 2 (treatment order) RM ANOVA was conducted on the follicular to luteal DSR change scores from the second treatment cycle. There was a non-significant main effect of treatment order (F(1, 30) = 1.08, p>0.05). However, there was a significant treatment by treatment order interaction (F(1, 30) = 5.56, p<0.05, partial \( \eta^2=0.16 \)). There was a trend towards a main effect of treatment (F(1, 30) = 3.14, p=0.086, partial \( \eta^2=0.095 \)). Figure 7.18 suggests that women showed a greater increase in symptom reporting between the follicular and luteal phases on placebo treatment than on SJW. A simple effects analysis revealed that this difference was statistically significant for those women who were administered SJW as their second treatment (t(14)= -2.55, p=0.023) but not for those who were administered it first (t(16)= -1.90, p=0.076).
A 2 (treatment) x 2 (treatment order) doubly multivariate ANOVA was conducted on the follicular to luteal DSR subscale change scores from the second treatment cycle. There was a non-significant main effect of treatment order (multivariate F(4, 27) = 2.69, p>0.05, partial $\eta^2=0.29$) and a non-significant treatment by treatment order interaction (multivariate F(4, 27) = 0.82, p>0.05, partial $\eta^2=0.11$). There was a non-significant main effect of treatment (multivariate F(4, 27) = 0.83, p>0.05, partial $\eta^2=0.11$). The univariate tests of the main effect of treatment are displayed in Table 7.15.

Table 7.15 Multivariate and univariate statistics for the main effect of treatment for the DSR analysis of follicular to luteal change scores (2nd cycle)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multivariate</td>
<td>F(4, 27)=0.83</td>
</tr>
<tr>
<td>Univariate:</td>
<td></td>
</tr>
<tr>
<td>- Mood</td>
<td>F(1, 30)=1.01</td>
</tr>
<tr>
<td>- Behavioural</td>
<td>F(1, 30)=3.12*</td>
</tr>
<tr>
<td>- Pain</td>
<td>F(1, 30)=2.04</td>
</tr>
<tr>
<td>- Physical</td>
<td>F(1, 30)=1.03</td>
</tr>
</tbody>
</table>

NB. **p<0.01 *p<0.05 -p<0.08

Table 7.15 shows that there was a non-significant effect of treatment on the mood, pain and physical DSR subscales (largest F(1, 30)=2.04, p=0.16, partial $\eta^2=0.064$). However, there was a trend for an effect of treatment on the behavioural subscale (F(1, 30) = 3.12, p=0.088 partial $\eta^2=0.094$).

7.4.8.1.4 Change from baseline scores

A 2 (treatment) x 2 (treatment order) RM ANOVA was conducted on the change from baseline DSR scores. There was a non-significant main effect of treatment order (F(1, 30) = 2.60, p>0.05) and a non-significant treatment by treatment order interaction (F(1, 30) = 0.47, p>0.05). There was a significant main effect of treatment (F(1, 30) = 10.15, p<0.01, partial $\eta^2=0.25$). Inspection of the means revealed that women showed a greater change from baseline score on SJW treatment (mean = 5.88) than on placebo (mean = 2.75).

A 2 (treatment) x 2 (treatment order) doubly multivariate ANOVA was conducted on the change from baseline DSR subscale scores. There was a non-significant main effect of treatment order (multivariate F(4, 27) = 1.21, p>0.05) and a non-significant treatment by treatment order interaction (multivariate F(4, 27) = 0.43, p>0.05). There was a non-
significant main effect of treatment (multivariate F(4, 27) = 2.03, p>0.05). The univariate tests of the main effect of treatment are displayed in Table 7.16.

Table 7.16 Multivariate and univariate statistics for the main effect of treatment for the DSR analysis of change from baseline scores

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Partial η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multivariate</td>
<td>F(4, 27)=2.03</td>
</tr>
<tr>
<td>Univariate:</td>
<td></td>
</tr>
<tr>
<td>-Mood</td>
<td>F(1, 30)=5.84*</td>
</tr>
<tr>
<td>-Behavioural</td>
<td>F(1, 30)=8.17**</td>
</tr>
<tr>
<td>-Pain</td>
<td>F(1, 30)=3.86~</td>
</tr>
<tr>
<td>-Physical</td>
<td>F(1, 30)=2.80</td>
</tr>
</tbody>
</table>

NB. **p<0.01 *p<0.05 ~p<0.08

Inspection of the univariate effects revealed that there was a significant effect of treatment on the mood and behavioural subscales (smallest F(1, 30) = 5.84, p=0.022, partial η²=0.16). This trend was also shown for the pain subscale (F(1. 30) = 3.86, p=0.059, partial η²=0.11). Inspection of the means revealed that the change scores were larger on SJW than placebo treatment for all subscales.

7.4.8.2 Analytical strategies applied to the DSR-20

All forms of analysis that were applied to the DSR (Freeman et al., 1996) (see sections 7.4.7 to 7.4.8.1.4) were then applied to the DSR-20. This was done to evaluate whether the DSR-20, comprising the original DSR items, and the additional items ‘anger,’ ‘aggression’ and ‘impulsiveness,’ clustered together through the use of the factors derived in Chapter 4 (see section 4.5), and weighted by their factor score coefficients (see section 4.8.1), was a more sensitive tool to examine treatment effects than the original DSR (Freeman et al., 1996).

Figures 7.23 and 7.24 show the mean DSR-20 psychological and physical subscale scores that were reported during the two cycles of SJW and placebo treatment.
7.4.8.2.1 Actual scores

A 2 (treatment) x 2 (cycle) x 4 (phase) x 2 (treatment order) doubly multivariate ANOVA was performed on the DSR-20 subscale scores. There was a non-significant main effect of treatment order (multivariate F(2, 29) = 0.48, p>0.05) and a non-significant treatment by treatment order interaction (multivariate F(2, 29) = 1.07, p>0.05). There was a non-significant main effect of treatment (multivariate F(2, 29) = 2.24, p>0.05) and a non-significant treatment by cycle (multivariate F(2, 29) = 0.40, p>0.05) interaction. There was a trend towards a treatment by cycle phase interaction (multivariate F (6, 25) = 2.09, p=0.091, partial \( \eta^2 = 0.33 \)). The univariate tests of the main effect of treatment and the treatment by cycle and treatment by cycle phase interactions are displayed in Table 7.17.

### Table 7.17 Multivariate and univariate statistics for the main effect of treatment, and treatment by cycle and treatment by cycle phase interactions, for the DSR-20 treatment analysis of actual scores

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Treatment</th>
<th>Treatment by cycle</th>
<th>Treatment by cycle phase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
<td>Partial ( \eta^2 )</td>
<td>Statistic</td>
</tr>
<tr>
<td>Actual scores</td>
<td>F(2, 29)=2.24</td>
<td>0.13</td>
<td>F(2, 29)=0.40</td>
</tr>
<tr>
<td>Multivariate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Univariate:</td>
<td>F(1, 30)=0.87</td>
<td>0.028</td>
<td>F(1, 30)=0.21</td>
</tr>
<tr>
<td>-Psychological</td>
<td>F(1, 30)=4.31*</td>
<td>0.13</td>
<td>F(1, 30)=0.70</td>
</tr>
</tbody>
</table>

NB. **p<0.01  *p<0.05**
Inspection of the univariate effects revealed that there was a significant main effect of treatment for the physical subscale (Table 7.17), such that women reported significantly fewer symptoms on SJW (mean = 0.39) than placebo (mean = 0.46) treatment. The treatment by cycle interaction was non-significant for each subscale. There was a significant treatment by cycle phase interaction for physical symptoms. The symptoms reported during the bleed, follicular and rest phases appeared similar, with the greatest separation occurring premenstrually, especially during treatment cycle two (Figure 7.24). This was confirmed through a simple effects analysis, which showed that there was a non-significant effect of treatment at the bleed, follicular and rest cycle phases (largest t(31)= -1.49, p=0.15), while there was a significant treatment effect premenstrually (t(31)= -3.37, p=0.002).

7.4.8.2.2 Premenstrual scores
A 2 (treatment) x 2 (cycle) x 2 (treatment order) doubly multivariate ANOVA was conducted on the premenstrual DSR-20 subscale scores. There was a non-significant main effect of treatment order (multivariate F(2, 29) = 0.16 p>0.05) and a non-significant treatment by treatment order interaction (multivariate F(2, 29) = 1.22, p>0.05). There was a significant main effect of treatment (multivariate F(2, 29) = 5.53, p<0.01, partial η²=0.28). Moreover, there was a significant treatment by cycle interaction (multivariate F(2, 29) = 3.72, p<0.05, partial η²=0.20). The univariate tests of the main effect of treatment and the treatment by cycle interaction are displayed in Table 7.18

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Statistic</th>
<th>Partial η²</th>
<th>Treatment by cycle</th>
<th>Statistic</th>
<th>Partial η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premenstrual Multivariate</td>
<td>F(2, 29)=5.53**</td>
<td>0.28</td>
<td>F(2, 29)=3.72*</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>Univariate: -Psychological</td>
<td>F(1, 30)=0.13</td>
<td>0.004</td>
<td>F(1, 30)=4.91*</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>Physical</td>
<td>F(1, 30)=10.82**</td>
<td>0.27</td>
<td>F(1, 30)=1.48*</td>
<td>0.047</td>
<td></td>
</tr>
</tbody>
</table>

NB. **p<0.01 *p<0.05

Inspection of the univariate effects revealed that there was a significant main effect of treatment on the physical subscale, such that women reported significantly fewer symptoms on SJW (mean = 0.48) than placebo (mean = 0.73) treatment. The treatment by cycle interaction was significant for the psychological subscale. Women reported
slightly greater levels of psychological symptoms on SJW than placebo treatment during the luteal phase of the first treatment cycle and reported lower levels on SJW during the luteal phase of the second (Figure 7.23). A simple effects analysis revealed that there was a non-significant effect of treatment on psychological symptoms during both treatment cycles (largest t(31)= -1.85, p=0.075).

7.4.8.2.3 Follicular to luteal difference scores
A 2 (treatment) x 2 (cycle) x 2 (treatment order) doubly multivariate ANOVA was conducted on the follicular to luteal DSR-20 subscale change scores. There was a non-significant main effect of treatment order (multivariate F(2, 29) = 0.57, p>0.05) and a non-significant treatment by treatment order interaction (multivariate F(2, 29) = 0.45, p>0.05). There was a significant main effect of treatment (multivariate F(2, 29) = 3.44, p<0.05, partial η²=0.19). There was a non-significant treatment by cycle interaction (multivariate F(2, 29) = 1.59, p>0.05). The univariate tests of the main effect of treatment and the treatment by cycle interaction are displayed in Table 7.19.

Table 7.19 Multivariate and univariate statistics for the main effect of treatment and treatment by cycle interaction for the DSR-20 analysis of follicular to luteal change scores

<table>
<thead>
<tr>
<th></th>
<th>Treatment</th>
<th>Treatment by cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
<td>Partial η²</td>
</tr>
<tr>
<td>Follicular to luteal change Multivariate</td>
<td>F(2, 29)=3.44*</td>
<td>0.19</td>
</tr>
<tr>
<td>Univariate:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Psychological</td>
<td>F(1, 30)=0.24</td>
<td>0.008</td>
</tr>
<tr>
<td>-Physical</td>
<td>F(1, 30)=6.63*</td>
<td>0.18</td>
</tr>
</tbody>
</table>

NB. **p<0.01 *p<0.05

Inspection of the univariate effects revealed that there was a significant main effect of treatment on the physical subscale, such that women reported a significantly smaller symptom increase between the follicular and luteal phases on SJW (mean = 0.32) than placebo (mean = 0.57) treatment. The treatment by cycle interactions were non-significant for both subscales.

7.4.8.2.4 Follicular to luteal difference scores (2nd treatment cycle)
A 2 (treatment) x 2 (treatment order) doubly multivariate ANOVA was conducted on the follicular to luteal DSR-20 subscale change scores from the second treatment cycle. There was a non-significant main effect of treatment order (multivariate F(2, 29) = 1.09,
p>0.05) and a non-significant treatment by treatment order interaction (F(2, 29) = 1.75, p>0.05). There was a trend towards a main effect of treatment (multivariate F(2, 29) = 3.19, p=0.056, partial η²=0.18). The univariate tests of the main effect of treatment are displayed in Table 7.20.

Table 7.20 Multivariate and univariate statistics for the main effect of treatment for the DSR-20 analysis of follicular to luteal change scores (2nd cycle)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Partial η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multivariate</td>
<td>F(2, 29)=3.19*</td>
</tr>
<tr>
<td>Univariate: Psychological</td>
<td>F(1, 30)=0.76</td>
</tr>
<tr>
<td>Physical</td>
<td>F(1, 30)=5.87*</td>
</tr>
</tbody>
</table>

NB. **p<0.01 *p<0.05 ~p<0.08

Inspection of the univariate effects revealed that there was a non-significant effect of treatment on the psychological subscale (F(1, 30) = 0.76, p>0.05), but there was a significant treatment effect on physical symptoms (F(1, 30) = 5.87, p<0.05, partial η²=0.16).

7.4.8.2.5 Change from baseline scores

A 2 (treatment) x 2 (treatment order) doubly multivariate ANOVA was conducted on the change from baseline DSR-20 subscale scores. There was a non-significant main effect of treatment order (multivariate F(2, 29) = 1.40, p>0.05) and a non-significant treatment by treatment order interaction (F(2, 29) = 3.02, p>0.05). There was a significant main effect of treatment (multivariate F(2, 29) = 4.93, p<0.05, partial η²=0.25). The univariate tests of the main effect of treatment are displayed in Table 7.21.

Table 7.21 Multivariate and univariate statistics for the main effect of treatment for the DSR-20 analysis of change from baseline scores

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Partial η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multivariate</td>
<td>F(2, 29)=4.93*</td>
</tr>
<tr>
<td>Univariate: Psychological</td>
<td>F(1, 30)=3.31*</td>
</tr>
<tr>
<td>Physical</td>
<td>F(1, 30)=7.81*</td>
</tr>
</tbody>
</table>

NB. **p<0.01 *p<0.05 ~p<0.08
Inspection of the univariate effects revealed that there was a significant effect of treatment on the physical subscale \( (F(1, 30) = 7.81, p<0.01, \text{ partial } \eta^2=0.21) \) and a trend towards a treatment effect of psychological symptoms \( (F(1, 30) = 3.31, p=0.079, \text{ partial } \eta^2=0.099) \). Inspection of the means revealed that the change scores were larger on SJW \( (\text{mean} = 0.54) \) than placebo \( (\text{mean} = 0.37) \) treatment for both subscales.

### 7.4.8.3 Aggression, impulsivity, anxiety and depression in response to SJW

Scores from the BPAQ, BIS-11, STAI-S and BDI were then examined to examine whether a treatment effect operated on the PMS symptoms aggression, impulsivity, anxiety and depression.

#### 7.4.8.3.1 BPAQ

Figures 7.25 to 7.29 show the mean BPAQ total and subscale scores that were reported during the two cycles of SJW and placebo treatment.

![Figure 7.25 Effect of treatment on total aggression scores during each treatment cycle](image)
The aggression profiles that were reported during SJW and placebo treatment appear very similar (Figures 7.25 to 7.29). As Table 7.22 indicates, the treatment analysis that was conducted on the actual BPAQ scores produced a non-significant main effect of treatment, and non-significant treatment by cycle and treatment by cycle phase interactions when the BPAQ total scores were analysed, and when the subscale scores were assessed through a doubly multivariate ANOVA. The RM ANOVA and the
doubly multivariate ANOVA that were conducted on the premenstrual total and subscale scores also produced non-significant main effects of treatment and non-significant treatment by cycle interactions. When the change from baseline scores were considered, non-significant effects of treatment were produced for the BPAQ total and subscale scores. The analysis of the follicular to luteal change scores from both treatment cycles produced a trend towards an effect of treatment on the physical aggression subscale. This effect became significant when only the second treatment cycle was considered. Women reported a smaller follicular to luteal increase in physical aggression on SJW (mean = 0.74) than on placebo (mean = 2.65) treatment, partly due to reporting slightly greater levels of physical aggression on SJW than placebo treatment in the follicular phase, and partly due to reporting slightly lower levels premenstrually (Figure 7.27).

7.4.8.3.2 BIS-11

Figures 7.30 to 7.33 show the mean BIS-11 total and subscale scores that were reported during the two cycles of SJW and placebo treatment.
The impulsivity profiles that were reported on SJW appear extremely similar to those reported on placebo treatment (Figures 7.30 to 7.33). This was confirmed through the multivariate and univariate treatment analyses. As Table 7.22 indicates, the RM ANOVAs and the doubly multivariate ANOVAs that were conducted on the actual BIS-11 scores, the premenstrual scores, the follicular to luteal change scores and the change from baseline scores all produced a non-significant main effect of treatment. All treatment by cycle and treatment by cycle phase interactions were also found to be non-significant (p>0.05).
7.4.8.3.3 STAIS and BDI

Figures 7.34 and 7.35 show the mean STAIS and BDI total scores that were reported during the two cycles of SJW and placebo treatment.

The anxiety and depression profiles that were reported on SJW and placebo treatment appear similar, although women showed the tendency to report slightly lower depression levels on SJW treatment (Figures 7.34 and 7.35). The RM ANOVAs that were conducted on the actual STAIS and BDI scores, their premenstrual scores, their follicular to luteal change scores and their change from baseline scores all produced a non-significant main effect of treatment. All treatment by cycle and treatment by cycle phase interactions were also found to be non-significant (Table 7.22).
<table>
<thead>
<tr>
<th></th>
<th>Actual Scores</th>
<th>Premenstrual Scores</th>
<th>Post-Pre change scores</th>
<th>Post-Pre (2nd cycle)</th>
<th>Change from baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment</td>
<td>Treatment by cycle</td>
<td>Treatment by cycle phase</td>
<td>Treatment by cycle</td>
<td>Treatment by cycle</td>
</tr>
<tr>
<td>DSRT: Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>subscales (MV)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Univariate:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mood</td>
<td>F(1, 30)=4.82*</td>
<td>F(1, 30)=1.99</td>
<td>F(1, 30)=1.16</td>
<td>F(1, 30)=6.18*</td>
<td>F(1, 30)=10.15**</td>
</tr>
<tr>
<td>Behavioural</td>
<td>F(1, 30)=2.33</td>
<td>F(1, 30)=3.07</td>
<td>F(1, 30)=1.18</td>
<td>F(1, 30)=8.79**</td>
<td>F(1, 30)=5.84*</td>
</tr>
<tr>
<td>Physical</td>
<td>F(1, 30)=0.20</td>
<td>F(1, 30)=4.19*</td>
<td>F(1, 30)=0.99</td>
<td>F(1, 30)=3.28*</td>
<td>F(1, 30)=3.86*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSRT-20 (MV)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Univariate:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychological</td>
<td>F(2, 29)=2.24</td>
<td>F(2, 29)=0.40</td>
<td>F(2, 29)=2.09</td>
<td>F(2, 29)=3.53**</td>
<td>F(2, 29)=4.93*</td>
</tr>
<tr>
<td>Physical</td>
<td>F(1, 30)=0.87</td>
<td>F(1, 30)=0.21</td>
<td>F(1, 30)=0.26</td>
<td>F(1, 30)=3.72*</td>
<td>F(2, 29)=3.19*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPAQ: Total</td>
<td>F(1, 30)=0.45</td>
<td>F(1, 30)=0.066</td>
<td>F(1, 30)=0.42</td>
<td>F(1, 30)=0.26</td>
<td>F(1, 30)=0.93</td>
</tr>
<tr>
<td>Subscales (MV):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Univariate:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal</td>
<td>F(4, 27)=0.31</td>
<td>F(4, 27)=1.03</td>
<td>F(4, 27)=1.09</td>
<td>F(4, 27)=1.99</td>
<td>F(4, 27)=1.24</td>
</tr>
<tr>
<td>Physical</td>
<td>F(1, 30)=0.21</td>
<td>F(1, 30)=0.38</td>
<td>F(1, 30)=0.24</td>
<td>F(1, 30)=0.072</td>
<td>F(1, 30)=0.26</td>
</tr>
<tr>
<td>Hostility</td>
<td>F(1, 30)=0.11</td>
<td>F(1, 30)=0.037</td>
<td>F(1, 30)=0.51</td>
<td>F(1, 30)=0.18</td>
<td>F(1, 30)=0.43</td>
</tr>
<tr>
<td>Anger</td>
<td>F(1, 30)=0.93</td>
<td>F(1, 30)=0.07</td>
<td>F(1, 30)=0.42</td>
<td>F(1, 30)=0.084</td>
<td>F(1, 30)=0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIS-11: Total</td>
<td>F(1, 30)=0.89</td>
<td>F(1, 30)=0.016</td>
<td>F(1, 30)=0.59</td>
<td>F(1, 30)=0.082</td>
<td>F(1, 30)=0.91</td>
</tr>
<tr>
<td>Subscales (MV):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Univariate:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attentional</td>
<td>F(3, 28)=0.44</td>
<td>F(3, 28)=0.76</td>
<td>F(3, 28)=0.98</td>
<td>F(3, 28)=0.08</td>
<td>F(3, 28)=1.18</td>
</tr>
<tr>
<td>Motor</td>
<td>F(1, 30)=1.99</td>
<td>F(1, 30)=1.37</td>
<td>F(1, 30)=0.84</td>
<td>F(1, 30)=0.23</td>
<td>F(3, 28)=0.50</td>
</tr>
<tr>
<td>Non-planning</td>
<td>F(1, 30)=0.43</td>
<td>F(1, 30)=0.15</td>
<td>F(1, 30)=0.21</td>
<td>F(1, 30)=0.029</td>
<td>F(3, 28)=0.18</td>
</tr>
<tr>
<td>STAI-S</td>
<td>F(1, 30)=0.32</td>
<td>F(1, 30)=1.14</td>
<td>F(1, 30)=0.26</td>
<td>F(1, 30)=0.032</td>
<td>F(1, 30)=0.14</td>
</tr>
<tr>
<td>BDI</td>
<td>F(1, 30)=2.13</td>
<td>F(1, 30)=0.12</td>
<td>F(1, 30)=1.22</td>
<td>F(1, 30)=0.05</td>
<td>F(3, 28)=1.18</td>
</tr>
</tbody>
</table>

*p<0.05  **p<0.01  *trend (p<0.08)
7.4.9 Biochemical analysis

Measures of the steroid hormones oestradiol, progesterone, testosterone, LH, FSH, and PRL, and the cytokines IL-1β, IL-6, IL-8, IFN-γ and TNF-α, were taken during the follicular and luteal phases of the second cycle of placebo and SJW treatment. These were examined to determine whether SJW administration had an effect on levels of their excretion. In total, five RM ANOVAs were conducted on the cytokine data and six RM ANOVAs were conducted on the steroid hormone data. Univariate analyses were conducted for the steroid hormone and cytokine analysis (see section 5.4.2).
7.4.9.1 Steroid hormones

Figures 7.36 to 7.41 display the mean steroid hormone concentrations during the follicular and luteal cycle phases of SJW and placebo treatment.

Figure 7.36 Mean oestradiol concentration
Figure 7.37 Mean progesterone concentration

Figure 7.38 Mean testosterone concentration
Figure 7.39 Mean LH concentration

Figure 7.40 Mean FSH concentration
Figure 7.41 Mean PRL concentration
The luteal levels of oestradiol and progesterone were higher on SJW than placebo treatment (see Figures 7.36 and 7.37). As levels of oestradiol and progesterone decrease towards the end of the menstrual cycle (see section 1.1.4), it was therefore important to ensure that the women had not had their blood sampled earlier in the luteal phase on SJW than placebo treatment. A paired samples t-test revealed that there was no significant difference between the cycle days on which blood sampling took place during the luteal phase of SJW (mean = -2.94) and placebo (mean = -3.50) treatment (t(17)=1.57, p=0.14). Moreover, there was no significant difference between the cycle days on which blood sampling took place during the follicular phase on SJW (mean = 7.19) and placebo (mean = 7.07) treatment (t(26)=0.21, p=0.84). Therefore the steroid hormone and cytokine levels that were exhibited during the follicular and luteal phases on SJW and placebo treatment could be compared.

The univariate tests of the main effects of treatment and cycle phase and of the treatment by cycle phase interactions of the steroid hormone analyses can be found in Table 7.23.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Cycle phase</th>
<th>Treatment by cycle phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistic</td>
<td>Partial $\eta^2$</td>
<td>Statistic</td>
</tr>
<tr>
<td>Oestradiol</td>
<td>F(1, 13)=0.51</td>
<td>0.038</td>
</tr>
<tr>
<td>Progesterone</td>
<td>F(1, 13)=3.89</td>
<td>0.23</td>
</tr>
<tr>
<td>Testosterone</td>
<td>F(1, 13)=3.33</td>
<td>0.20</td>
</tr>
<tr>
<td>LH</td>
<td>F(1, 13)=0.01</td>
<td>0.001</td>
</tr>
<tr>
<td>FSH</td>
<td>F(1, 13)=0.53</td>
<td>0.039</td>
</tr>
<tr>
<td>PRL</td>
<td>F(1, 13)=0.19</td>
<td>0.014</td>
</tr>
</tbody>
</table>

** $p<0.001$  * $p<0.01$

Table 7.23 shows that there was a significant main effect of cycle phase for progesterone and FSH, such that women exhibited significantly greater levels of FSH during the follicular phase and significantly greater levels of progesterone during the luteal phase. However, the main effect of treatment and treatment by cycle phase interactions were non-significant for each steroid hormone.
7.4.9.2 Cytokines

Figures 7.32 to 7.36 display the mean cytokine concentrations during the follicular and luteal cycle phases of SJW and placebo treatment.

![Figure 7.32 Mean IL-1ß concentration](image)

![Figure 7.33 Mean IL-6 concentration](image)

![Figure 7.34 Mean IL-8 concentration](image)

![Figure 7.35 Mean TNF-α concentration](image)

![Figure 7.36 Mean IFN-γ concentration](image)
The univariate tests of the main effects of treatment and cycle phase and of the treatment by cycle phase interactions of the cytokine analyses are shown in Table 7.24.

Table 7.24 Univariate statistics for the main effects of treatment and cycle phase, and treatment by cycle phase interactions for the cytokine analyses (n=12 to n=14)

<table>
<thead>
<tr>
<th>Treatment by cycle phase</th>
<th>Treatment</th>
<th>Cycle phase</th>
<th>Treatment</th>
<th>Cycle phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistic</td>
<td>Partial η²</td>
<td>Statistic</td>
<td>Partial η²</td>
<td>Statistic</td>
</tr>
<tr>
<td>- IL-1β</td>
<td>F(1, 12)=0.93</td>
<td>0.072</td>
<td>F(1, 12)=0.88</td>
<td>0.068</td>
</tr>
<tr>
<td>- IL-6</td>
<td>F(1, 12)=2.41</td>
<td>0.12</td>
<td>F(1, 12)=2.41</td>
<td>0.17</td>
</tr>
<tr>
<td>- IL-8</td>
<td>F(1, 11)=0.69</td>
<td>0.059</td>
<td>F(1, 12)=0.54</td>
<td>0.12</td>
</tr>
<tr>
<td>- IFN-γ</td>
<td>F(1, 13)=0.76</td>
<td>0.055</td>
<td>F(1, 13)=0.80</td>
<td>0.058</td>
</tr>
<tr>
<td>- TNF-α</td>
<td>F(1, 13)=1.53</td>
<td>0.11</td>
<td>F(1, 13)=0.017</td>
<td>0.001</td>
</tr>
</tbody>
</table>

* p<0.05  ~ p<0.06

Table 7.24 shows that the main effects for treatment and cycle phase, and the treatment by cycle phase interactions, were non-significant for each cytokine (p>0.05).

7.5 Discussion

The primary objective of this study was to determine whether SJW was more beneficial than placebo treatment in relieving symptoms of PMS in a 10 cycle double-blind randomised controlled crossover trial. In addition, although the intensity of the placebo response has been well documented in populations of PMS sufferers (e.g. Magos et al., 1986; Rapkin, 2003; Steiner et al., 1995), little research has addressed which PMS symptoms contribute most to this effect. Therefore, a subsidiary aim of this study was to explore which PMS symptoms are involved in the placebo effect over four menstrual cycles of placebo treatment, facilitated by the placebo run-in design of the study.

7.5.1 The effectiveness of SJW for PMS

The primary (complete) analysis conducted on the total DSR (Freeman et al., 1996) and subscale scores compared the symptoms that were reported during SJW and placebo treatment across four cycle phases of the two treatment cycles. The analysis of the total DSR scores revealed that SJW was superior to placebo treatment. Moreover, when the univariate effects from the multivariate analysis were examined, SJW was shown to benefit physical (food craving, swelling) and behavioural (poor coordination, insomnia, confusion, headaches, crying, fatigue) premenstrual symptoms. The non-significant
treatment by cycle interactions indicated that these beneficial effects of SJW on physical and behavioural symptoms were apparent within the first treatment cycle.

7.5.2 Secondary analyses

7.5.2.1 Comparison of commonly used analytical approaches in PMS trials

Researchers who conduct placebo controlled PMS treatment trials determine the effectiveness of the treatment under analysis through the comparison of the symptoms reported during active and placebo treatment. Various methods are used to do this. Some of the most common are the assessment of premenstrual scores (e.g. Halbreich et al., 2002; Walker et al., 1998), follicular to luteal change scores (e.g. Doll et al., 1989) and change from baseline scores (e.g. Atmaca et al., 2003; Hagen et al., 1985; Hicks et al., 2004; Schellenberg, 2001). The impact of these different approaches to the analysis of clinical trial data, and the conclusions that they lead to, was assessed through the comparison of the results arising from each of these widely used forms of analysis, with each other and with the primary (complete) analysis discussed in Section 7.5.1.

The complete analysis, and the analysis of the premenstrual and change from baseline scores, that were conducted on the total DSR scores, all suggested SJW to be superior to placebo treatment. However, a benefit of SJW was not indicated through analysis of the follicular to luteal change scores. When the DSR subscales were examined through the separate multivariate analyses, a significant multivariate treatment effect was only apparent from the analysis focusing on the premenstrual scores. This highlights that univariate assessment of total scale scores is not analogous to the multivariate effects produced from multivariate assessment of subscale scores. Whereas the univariate analyses simply examined the total scale score, the multivariate analyses of the subscales gave a global assessment of the treatment effect, taking into account the linear correlation among the subscales.

When the univariate effects from the multivariate analyses were examined, the complete analysis indicated SJW to benefit physical and behavioural symptoms. Although the analysis of the change from baseline scores also revealed a beneficial effect of SJW on behavioural symptoms, physical symptoms appeared unaffected. A benefit was also indicated for mood symptoms, while this trend was apparent for pain symptoms. The analysis of the premenstrual scores revealed all of these effects. Behavioural and physical symptoms were shown to be improved irrespective of treatment cycle, mood...
symptoms were shown to be improved during the second cycle of treatment, whilst this trend was shown for pain symptoms. The analysis of the follicular to luteal change scores revealed no benefit of SJW.

As was demonstrated in the primary (complete) analysis (see section 7.4.7), this amalgamation of results suggests that SJW benefits behavioural and physical PMS symptoms during the first treatment cycle in which it is taken. Additionally, these supplementary analyses suggest that mood symptoms are benefited during the second menstrual cycle of treatment with SJW. Furthermore, the comparison of results from these secondary analyses demonstrates that the outcome from a PMS treatment trial can be completely dependent upon the way in which the data is analysed. If only the follicular to luteal change scores had been assessed in this research, then it would have been concluded that SJW provides no benefit for PMS symptoms. However, the sole use of any other analysis type would have led to the conclusion that SJW is promising as a PMS treatment. However, the recommendation for particular PMS symptoms would have differed. This demonstrates the need for the different types of analysis that are commonly used in PMS treatment trials to be evaluated and a consensus on an appropriate analytical strategy to be reached.

7.5.2.2 DSR-20

The analyses that were conducted on the DSR-20 scores also suggested SJW to be beneficial for PMS symptoms. Although a non-significant multivariate effect of treatment was produced from the complete analysis and from the analysis of the follicular to luteal change scores, a significant multivariate treatment effect was produced through the analysis of the premenstrual and change from baseline scores. This trend was also apparent when the follicular to luteal scores from the second treatment cycle were assessed. SJW was consistently found to benefit physical PMS symptoms, as the inspection of the univariate effects produced from all multivariate analyses revealed a significant effect of treatment for the physical DSR-20 subscale (See Table 4.9), such that women reported a lower symptom severity on SJW than on placebo treatment. The non-significant treatment by cycle interactions that were produced suggest that these beneficial effects of SJW on physical PMS symptoms appear within the first cycle during which it is taken.
Although all analysis methods produced a non-significant effect of treatment for psychological symptoms, women began to report lower levels on all psychological symptoms that were assessed on SJW than placebo treatment towards the end of the treatment period (see Table 4.9). The treatment by cycle interaction produced from the analysis of premenstrual scores suggested that SJW may benefit psychological symptoms by the luteal phase of the second treatment cycle. Moreover, the analysis of the change from baseline scores, which solely assessed symptom levels reported at the end of treatment in comparison to baseline values, revealed a trend towards an effect of treatment. These findings suggest that SJW may benefit psychological PMS symptoms, but only during the second cycle during which it is taken. Further research assessing the effectiveness of SJW on psychological PMS symptoms over longer treatment duration is warranted.

7.5.2.3 The DSR (Freeman et al., 1996) versus the DSR-20

The findings from the DSR analyses which suggested SJW to benefit behavioural and physical PMS symptoms during the first treatment cycle, are compatible with those of the DSR-20 analyses which suggested SJW to benefit physical PMS symptoms during this period, as both symptoms comprising Freeman et al.’s (1996) physical factor, and all but one item comprising their behavioural factor, load onto the physical DSR-20 subscale. Moreover, the findings from the DSR analyses which suggested SJW may benefit mood symptoms during the second treatment cycle, are compatible with those of the DSR-20 analyses which suggested SJW may benefit psychological PMS symptoms during this treatment cycle, as all symptoms comprising the DSR mood factor load onto the DSR-20 psychological subscale.

The findings were more consistent across the different methods of analysis for the DSR-20 than for the DSR (Freeman et al., 1996). However, the amalgamation of results from the different analyses that were conducted on the DSR suggested SJW to benefit physical and behavioural symptoms during the first treatment cycle. As all analyses that were conducted on the DSR-20 consistently identified these findings, by finding SJW to benefit physical PMS symptoms, the factor on which all but one of Freeman et al’s (1996) physical and behavioural symptoms load, this would suggest that the DSR-20 is a more sensitive tool for assessing treatment effects than the DSR. This may be because the Principal Components Analysis conducted to identify the factor structure of the DSR-20 (chapter 4) used factor retention methods considered more sound than those
that were used to produce the factor structure of the DSR (Zwick and Velicer, 1986; Hayton et al., 2004; Humphreys and Montanelli, 1975). Moreover, it was conducted on data reported by the same sample upon which the treatment analyses were conducted. However, the DSR-20 may be more sensitive to treatment effects, as the DSR-20 subscales take into account the importance of each item to the factor by weighting items by their factor score coefficients.

7.5.2.4 Aggression, impulsivity, anxiety and depression in response to SJW
The non-significant findings produced from all of the treatment analyses that were conducted on the STAIS, BDI and BIS-11 suggest that SJW does not benefit the premenstrual symptoms of anxiety, depression and impulsivity. However, these are all psychological PMS symptoms, and as discussed above, SJW may provide benefit for psychological PMS symptoms after a longer treatment duration. The analyses that were conducted on the BPAQ all suggested SJW to provide no benefit for aggression, with the exception of the analysis of follicular to luteal change scores, which produced a significant treatment effect of physical aggression. The levels of physical aggression that were reported across the two cycles of SJW and placebo treatment appeared sporadic. Therefore, this effect may not have resulted from a beneficial effect of SJW, but from a Type I error, which highlights the problem inherent in analysing univariate effects in the absence of a significant multivariate effect. This was the only BPAQ subscale that did not differentiate PMS sufferers from other groups of women in the analyses that were conducted in Chapter 5. Therefore, if SJW does benefit premenstrual physical aggression levels, these effects are probably not limited to sufferers of PMS, but may generalise to other groups of women.

7.5.3 Previous PMS treatment trials of SJW
Few trials have assessed the effectiveness of SJW for PMS. At the time the protocol was written for this study, only one small pilot study had been published (Stevinson and Ernst, 2000). Although these researchers found SJW to significantly reduce PMS symptoms in a sample of 19 women, the study did not employ a placebo group or crossover design. Therefore, their treatment effects cannot be separated from those arising from the placebo response. Since the commencement of this study, three randomised, double-blind, placebo-controlled trials employing a parallel design have been published (Hicks et al., 2004; Pakgohar et al., 2004; 2005). Hicks et al (2004) found that women who were administered SJW reported fewer symptoms than those
administered placebo treatment on all of the symptom subgroups that were studied. However, these differences did not reach statistical significance and the authors concluded that this may have resulted from insufficient power, rather than from a lack of a beneficial effect of SJW. These conclusions may be supported by the beneficial effects of SJW that were found in this study, which increased power by reducing between subject variability through the employment of a crossover design. Hicks et al. used a low dose of 600mg/day. A higher dose of 900mg/day may be needed for beneficial effects to become apparent. The findings from the current study appear compatible with those of Pakgohar et al. (2004; 2005), who also found SJW to reduce physical PMS symptoms. However, only the abstracts of these Iranian articles have been translated into English. The omission of some details from the abstracts, including the dose of SJW that was used and the symptoms that were benefited, mean that more detailed comparisons could not be made.

7.5.4 Active ingredients
Although SJW tablets were produced according to GMP and were standardised to 0.3% hypericin, batch analysis (Appendix 8.5) provided by the manufacturers, Lichtwer Pharma, revealed that the tablets that were supplied actually contained 0.18% hypericin and 3.38% hyperforin. This is within the acceptable range for commercially available SJW, but highlights the great variability permitted for herbal products by the licensing authorisation. Although hypericin was originally considered to be the central active component of SJW (Bennett et al., 1998; Gaster and Holroyd, 2000; Mennini and Gobbi, 2004), many now believe that hypericin does not play a significant role (Cott, 1997; Mennini and Gobbi, 2004; Wheatley, 1998). Currently, hyperforin is considered more central to the beneficial properties of SJW (Cervo et al., 2002; Chatterjee et al., 1998; Laakmann et al., 1998; Muller et al., 1998; Zanoli, 2004). Hicks et al. (2004) found that 600mg of SJW/day, reportedly standardised to 0.3% hypericin, provided no benefit to PMS symptoms. However, the research presented in this thesis found that 900mg SJW/day, standardised to 0.18% hypericin, benefited physical PMS symptoms. As the quantity of hypericin administered to participants in this research was actually lower than that provided by Hicks et al. (2004), this may support the role of hyperforin in the beneficial effects of SJW.
7.5.5 Mechanism of action

7.5.5.1 Steroid hormones

Hormonal profiles do not differ between women with and without PMS (Connolly, 2001; Rubinow, 1992, see also section 5.5.6.1). However, women with PMS appear more sensitive to normal cyclical fluctuations in steroid hormones, which influence neurotransmitter function in the central nervous system (Eriksson et al., 2002; Rapkin, 1992; Reid and Yen, 1981; Steiner et al., 1997a). GnRH agonists, which suppress ovulation, reduce physical and behavioural PMS symptoms (Eriksson et al., 2002; Wyatt et al., 2004; Yonkers et al., 2008). The inhibition of ovarian cyclicity through ovariectomy has also been shown to ameliorate PMS symptoms (Yonkers et al., 2008). Therefore, it is reasonable to consider that stabilisation of steroid hormone levels may reduce PMS symptoms (Yonkers et al., 2008). The results from this study showed that although women demonstrated cyclicity of progesterone and FSH, SJW did not alter the concentration of any steroid hormone that was measured across the cycle, in comparison to placebo treatment. Therefore, the beneficial effects of SJW for PMS that were found in this study were not associated with an alteration of the fluctuation of normal circulating steroid hormone levels across the menstrual cycle, but this does not exclude the possibility that effects may have resulted from an altered sensitivity to them.

7.5.5.2 Serotonin

Although a clear mechanism has not been proposed to explain how SJW may reduce PMS symptoms, it can be postulated that the biological actions of SJW provide potential for symptom reduction. SJW has been shown to lead to increased brain serotonin levels (Calapai et al., 2001), and to an up-regulation of 5-HT1A and 2A receptors (Teufel-Mayer and Gleitz, 1997) in the frontal cortex (Muller et al., 1997). Women with PMS have lower serotonin levels premenstrually (Ashby et al., 1988; Rapkin et al., 1987; Rasgon et al., 2000; Taylor et al., 1984). Selective serotonin reuptake inhibitors (SSRIs), which facilitate serotonergic transmission, have been shown to be an effective PMS treatment (Dimmock et al., 2000; Steiner et al., 2003; Szegedi et al., 2005), for both physical and behavioural symptoms (Wyatt et al., 2002; Johnson, 2004). SSRIs improve PMS symptoms within a few days (Eriksson et al., 2002; Landen and Eriksson, 2003; Yonkers et al., 2008). As SJW appeared to benefit physical PMS symptoms within the first treatment cycle, it is possible that these beneficial effects may have arisen from an increase in serotonin production. However,
serotonin production was probably not solely involved in the psychological PMS symptom effects, as these only became apparent towards the end of treatment. Unfortunately, although this research aimed to assess 5-HT and 5-HIAA concentrations, it was not possible to perform this analysis (see section 5.3.2.2.3). Therefore, further research assessing the efficacy of SJW for PMS should also assess its impact upon serotonergic transmission.

7.5.5.3 Cytokines
SJW has also been shown to suppress pro-inflammatory cytokine production (Fiebich et al., 2001; Thiele et al., 1994). Researchers have shown that production of the pro-inflammatory cytokines IL-1 (Cannon and Dinarello, 1985; Polan et al., 1990) and IL-6 (Konecna et al., 2000) increases premenstrually. Increased production of these cytokines has been associated with symptoms characteristic of PMS, including depression (Maes et al., 1995; Schlatter et al., 2001), anxiety (Connor et al., 1998), fatigue (Kronfel and Remick, 2000), headaches (Martelletti et al., 1993), decreased memory performance (Reichenberg et al., 2001) and sleep disturbances (Kapas and Krueger, 1992; Vgontzas et al., 2005). Moreover, the analysis that was conducted in Chapter 5 revealed that PMS sufferers had greater proinflammatory cytokine concentrations than women without PMS, and women who met PMS criteria but who did not consider themselves to suffer from the condition, with differences between groups reaching statistical significance for IL-8 and TNF-α. In theory, it is plausible to suggest that SJW may benefit PMS symptoms by reducing pro-inflammatory cytokine levels during the luteal phase but this proposed explanation requires further corroboration.

SJW treatment did not affect the concentration of any pro-inflammatory cytokine that was measured, in comparison to placebo treatment. However, most of the premenstrual symptoms that have been associated with increased pro-inflammatory cytokine production have been psychological in nature (Connor et al., 1998; Maes et al., 1995; Schlatter et al., 2001; Kapas and Krueger, 1992; Vgontzas et al., 2005). As discussed previously, psychological symptoms began to abate towards the end of treatment, although these effects were not strong enough to produce a significant effect of treatment in most of the treatment analyses. Therefore, although pro-inflammatory cytokine levels were not significantly altered by the end of treatment, it is still possible that they play a role in the reduction of psychological PMS symptoms.
The antidepressant properties of SJW may arise from a combination of mechanisms, each being too weak alone to result in a therapeutic effect (Bennett et al., 1998; Nangia et al., 2000; Raffa, 1998). The antidepressant properties of SJW may arise through the interaction of cytokines with the serotonergic system (Calapai et al., 2001). Theoretically, this could also be the case for PMS symptoms. Therefore, the sole assessment of the influence of SJW on pro-inflammatory cytokines may not be enlightening. Effects may only become apparent when these levels are assessed in conjunction with effects on other systems, such as the serotonergic system. Further research is needed to assess whether SJW does benefit psychological PMS symptoms over a longer treatment duration and to assess the role of pro-inflammatory cytokines, in combination with the serotonergic system, in this effect.

7.5.6 Placebo effect

When the placebo effect was explored through the analysis of the symptoms reported on the DSR (Freeman et al., 1996), a placebo effect was found to operate on mood symptoms throughout the cycle, on pain symptoms premenstrually and on physical symptoms during the second half of the cycle. Approximately half of the sample received a further two cycles of placebo treatment after the placebo run-in phase (treatment group two). The analysis of their symptom profiles revealed that the placebo effect acting upon pain symptoms had abated by the time women entered the third cycle of placebo treatment, while mood symptoms remained reduced during the menstrual and luteal phases after four placebo treatment cycles. There was also a trend for physical symptoms to remain reduced during the second half of the cycle, although these effects did not reach statistical significance when familywise error was controlled for. However, as the sample size was effectively halved in this exploratory analysis because only 16 women received four cycles of continuous placebo treatment, this lack of statistical significance possibly resulted from a loss of power, rather than from a lack of a continued effect.

When the same symptoms were clustered together through the use of the DSR-20 factor loadings, with items being weighted by their factor score coefficients, the placebo effect was found to operate on psychological symptoms during the menstrual and luteal phases. These effects were still apparent after four cycles of placebo treatment. Symptom reporting appeared particularly reduced during the luteal phase of the third cycle of placebo treatment. Women were given a new set of tablets at this stage, and
may have assumed that they were different from the ones they were taking previously. No placebo effect was found to operate on physical PMS symptoms during any cycle of placebo treatment, which may be explained by these symptoms being less influenced by subjective expectation.

The assessment of the weekly measures revealed that a placebo effect operated on depression, anxiety, verbal aggression, anger and non-planning impulsivity during menses, while this trend was shown for attentional impulsivity. Hostility was also shown to be reduced during the majority of the cycle. The assessment of the symptom profiles reported by treatment group two revealed that all of these effects had abated after four placebo treatment cycles. Symptom reports of depression, anxiety and non-planning impulsivity had returned to screening levels by the time women entered the third placebo cycle, while verbal aggression, anger, attentional impulsivity and hostility levels had returned to screening values by the fourth.

All placebo effects noted from the administration of the daily measures were found to operate premenstrually. In addition, DSR mood symptoms, and DSR-20 psychological symptoms were reduced during menses. Symptom reduction during these phases was expected, as this is when premenstrual symptoms are experienced. Premenstrual symptoms are experienced for up to two weeks before the onset of menstruation, and only disappear a few days into menses (Altshuler et al., 2001; Bäckström et al., 1983; Endicott et al., 1986; Milewicz and Jedrzejuk, 2006; Reid, 1991; Yonkers et al., 2008). Interestingly, the placebo effects noted from the administration of the weekly measures only operated during menses, and were not apparent premenstrually. The lack of a placebo effect during the premenstrual phase on these measures may have resulted from the way in which they were administered. Women were asked to complete the weekly diaries on the day their period began, by recalling how they had felt over the past week. The onset of bleeding gave women a physical cue of where they were in their cycle, and could have resulted in them reminiscing about how they had felt over the past week, attributing their negative feelings/behaviours to being premenstrual (Connolly, 2001; Roy-Bryne et al., 1985). Therefore, although women may have felt less anxious, depressed, aggressive and impulsive during placebo treatment during the premenstrual phase, these effects may have been masked by selective symptom recall cued by menses. Retrospective symptom recall is notoriously inaccurate (Ainscough, 1990; Connolly, 2001; Gallant et al., 1992a), and has been previously shown to lead to
overestimations in symptom severity (Christensen et al., 1989; De Souza et al., 2000; Gallant et al., 1992b; Roy-Bryne et al., 1985).

Therefore, many of the PMS symptoms involved in the placebo effect appear to return to baseline values within four cycles of placebo treatment. However the reduced symptom severity reported on the psychological subscale of the DSR-20, and on the mood and physical subscales of the DSR after four cycles of placebo treatment, suggest that some symptoms may be subject to a placebo effect of a longer duration. Magos et al. (1986) found women who were administered a placebo implant showed a very strong initial placebo response, with symptom ratings only returning to pre-treatment levels after six cycles. Therefore, four cycles of placebo treatment may not be long enough for all placebo effects to disappear, which may suggest that PMS treatment studies would benefit from incorporating placebo run-in phases of greater than four menstrual cycles before active treatment is administered. Very few studies to date have included a placebo run-in phase at all.

Nevertheless, it must be acknowledged that the reduction in symptom severity that was apparent in this study after four placebo treatment cycles may not be due to a continued placebo effect, but rather could result from the study design. Women were informed at the beginning of the trial that the screening cycles would be used to assess their eligibility to enter the treatment phase. Motivated women, who have tried many treatments for their symptoms, tend to volunteer for studies of this type, and they are keen to be included in the trial. This may have resulted in women exaggerating their symptom severity during these screening cycles to ensure that they would be eligible to take part in the clinical trial. Therefore, the reduction in the mood and physical symptoms noted on the DSR and of the psychological symptoms noted on the DSR-20 during placebo treatment may not reflect a placebo effect, but result from women beginning to report accurate symptom levels as the study progressed. Most researchers who have investigated the placebo effect have done so through the comparison of the symptom profiles that have been reported during placebo treatment with those reported during baseline cycles which were also used to determine study eligibility (e.g. Freeman et al., 1995; Freeman et al., 1999). Therefore, the way in which the placebo response has been estimated may have resulted in an overestimation of true levels.
Women were found to report premenstrual symptoms at a lower severity during the third screening cycle than during the first two. This effect may have resulted from the continued exposure to the symptom lists reducing their sensitivity. However, women attended a study visit during the luteal phase of the third screening cycle, during which they were informed that they were eligible to enter the treatment phase. This reassurance that the diagnosis of PMS was confirmed could have led to the lower premenstrual symptom severity reported during this cycle, due to the women no longer feeling the need to exaggerate their symptom severity. Alternatively, the first two cycles could reflect accurate symptom levels. The relief of being entered onto the clinical trial and being given the opportunity to try a treatment for PMS, or having had the opportunity to discuss their condition with the PI, may in itself have had a therapeutic effect and led to the reduction in symptoms observed. Previously, researchers have found cognitive behavioural therapy, CBT, to be as effective as fluoxetine in reducing PMS symptoms, indicating the psychological benefits women find in having a medic listen to them and take their symptoms seriously (Hunter et al., 2002). Therefore, although this research suggests that placebo effects have a long duration for PMS sufferers, there is a need for further research to untangle these placebo effects from the effect of taking estimates of baseline symptomatology from cycles also used to assess eligibility to participate in a treatment study. This could be achieved by only offering women the opportunity to participate in a treatment trial after completing symptom reports for three menstrual cycles. Moreover, research evaluating the potential therapeutic effect of investigator contact is warranted.

7.7 Summary

SJW provided benefit for physical and behavioural PMS symptoms during the first menstrual cycle in which it was taken. The hormone and cytokine analyses suggested that biochemical changes within these systems were probably not involved in these beneficial effects, although their involvement in interaction with the serotonergic system could not be explored further. Secondary analyses revealed the DSR-20 to be a more sensitive tool than the original DSR (Freeman et al., 1996) to detect treatment effects in the sample recruited for this research, and demonstrated that SJW benefited the symptoms comprising the physical DSR-20 subscale. The comparison of the results that arose from the different commonly used analytical approaches that were explored, demonstrate that the outcome from a PMS treatment trial can be completely dependent upon the way in which the data is analysed, and highlight the importance for a
consensus to be reached on what is an appropriate analytical strategy. Moreover, these comparisons revealed that although most forms of analysis did not produce a significant main effect of treatment for mood (DSR) or psychological (DSR-20) PMS symptoms, these symptoms appeared to abate towards the end of the luteal phase of the second treatment cycle. Further research is needed to assess the effectiveness of SJW for mood and/or psychological PMS symptoms over a longer duration in order to be able to recommend its use for sufferers of PMS who find these symptoms problematic. The placebo response was shown to influence various symptoms of PMS. Although many of these effects had abated within four cycles of placebo treatment, some symptoms remained reduced. Although the placebo effect may have a long duration in PMS sufferers, it is possible that these continued effects could have resulted from participants reporting exaggerated symptom severity during the screening cycles in order to guarantee study eligibility. Further research investigating the placebo effect over a longer duration is needed. This research should not take baseline values from cycles that are also used to assess eligibility for a clinical trial.
Chapter 8: General Discussion

The primary aim of the research presented in this thesis was to evaluate the efficacy of SJW for the treatment of PMS in a robustly designed randomised controlled trial. A subsidiary outcome of this research was to improve understanding regarding the characterisation and treatment of PMS, and to gain a greater insight into the diagnosis and aetiology of the condition. The research that was conducted revealed a number of clear findings which have been discussed in depth in their respective chapters. Therefore, this chapter considers the main findings and their implications for the diagnosis and treatment of PMS.

8.1 Treatment

Many PMS sufferers find current PMS treatment strategies unsatisfactory and are often reluctant to take the medication that is offered to them (Bendich, 2000). Lifestyle changes often do not benefit symptoms sufficiently. The treatments of choice for most medical practitioners are SSRIs (Connolly, 2001; Domoney et al., 2003; Johnson, 2004). However, SSRIs for PMS can produce significant side effects, such as gastrointestinal disturbances, insomnia, fatigue, headache, dry mouth, dizziness, tremor, sweating, decreased libido and delayed orgasm (Eriksson et al., 2002; Johnson, 2004; Yonkers et al., 2008). Many women with PMS find their physicians unhelpful (Domoney et al., 2003), while women’s magazines and other media advertise alternative therapies as ‘safer’ than prescribed drugs. Many PMS sufferers turn to dietary supplements and herbal remedies to treat their symptoms (Bendich, 2000; Domoney et al., 2003; Eriksson et al., 2002; Fugh-Berman and Kronenberg, 2002; Girman et al., 2003), although rigorous scientific studies to test their efficacy are lacking (Bendich, 2000; Domoney et al., 2003; Eriksson et al., 2002). Therefore a systematic research review of commonly used dietary supplements and herbal remedies for PMS (chapter 6), and a ten-cycle double-blind randomised controlled trial to test the efficacy of SJW for the condition (chapter 7) were conducted.

The systematic research review concluded that although the evidence supported the use of calcium and continuous B6 treatment, the evidence for most of the dietary supplements and herbal remedies that were assessed was either conflicting or insufficient. Many of the studies included methodological weaknesses, including inadequate sample sizes, unrepresentative samples, retrospective diagnosis and
assessment of treatment efficacy and study designs that were not fully blind or counterbalanced (see sections 6.31 to 6.37). Therefore, more randomised, double-blind, placebo controlled trials, using larger, representative samples, strict, prospectively confirmed, diagnostic criteria and assessment of treatment efficacy are needed in order to clarify the role of the many dietary supplements and herbal remedies that are currently consumed by PMS sufferers. The clinical trial presented in Chapter 7 aimed to test the efficacy of SJW for PMS using a rigorous methodology, and demonstrated this design to be a sound scientific basis on which to base future research. Whilst many dietary supplements or herbal remedies could have been explored, SJW was chosen, due to its biological actions on the serotonergic and cytokine systems (see section 7.1.4), and its current application in depression (see section 7.1.2).

SJW was found to benefit physical (food cravings and swelling) and behavioural (poor coordination, insomnia, confusion, headaches, crying and fatigue) PMS symptoms, during the first menstrual cycle in which it was taken (see section 7.4.7). However, the results presented in Chapter 7 suggest that there is currently insufficient evidence to be able to recommend the use of SJW for PMS sufferers who find mood and pain related symptoms problematic, as SJW appeared to have little impact on these types of PMS symptom during two cycles of treatment. As mood related PMS symptoms did appear to improve towards the end of the treatment period in comparison to placebo treatment, further research is needed to determine whether these symptoms would respond with longer treatment duration. Taken together the results of the current study provide evidence which suggests that SJW confers some benefit for PMS sufferers who find physical and behavioural premenstrual symptoms problematic.

8.2 Variations in treatment assessment methods

Although researchers who conduct placebo-controlled PMS trials determine the effectiveness of the treatment under analysis through the comparison of symptoms reported during active and placebo treatment, various methods are used to do this (see section 7.4.8.1). Each of these has benefits and limitations.

Running a complete analysis (as was performed as the primary analysis in this trial) allows a detailed examination of the data by assessing symptoms across several phases during each cycle of treatment. However, this also reduces the available degrees of freedom, and therefore, reduces the power to detect an effect. Consequently, researchers
do not tend to use this form of analysis, and instead analyse premenstrual scores (e.g. Halbreich et al., 2002; Walker et al., 1998), follicular to luteal change scores (e.g. Doll et al., 1989), or change from baseline scores (Atmaca et al., 2003; Hagen et al., 1985; Hicks et al., 2004; Schellenberg et al., 2001). Although the analysis of only premenstrual scores increases the power to detect an effect by solely examining symptom scores during the time period when a treatment effect is desired, the variability in symptom reporting across the cycle is not taken into account. The use of change from baseline scores also suffers from the same problem, but in addition, compares symptom scores on active and placebo treatment with baseline values taken from cycles used to assess study eligibility. Symptom reports from screening cycles may not reflect accurate symptom levels, as women may exaggerate their symptom severity in these cycles through demand characteristics, increased introspection or a desire for treatment (see section 7.5.6). Moreover, as this form of analysis only considers symptom levels from the last treatment cycle, no indication can be given as to the length of time that is needed for a treatment effect to become apparent. The analysis of follicular to luteal change scores determines the effectiveness of a treatment by assessing the variability of symptoms between the two cycle phases that are central to the diagnosis of PMS (Connolly, 2001; see also chapter 6). This approach can obscure a treatment effect if women experience a general improvement in symptoms, because the follicular to luteal change score would not appear to be reduced if symptom severity is decreased in both the follicular and luteal phases.

The clinical trial data presented in Chapter 7 was analysed using each of these commonly used forms of analysis subsequent to performance of the primary analysis. The comparison of the results that were produced from these different analytical strategies suggests that the analysis of premenstrual scores may be the treatment assessment method most likely to detect a significant effect. The complete (primary) analysis and the analysis of change from baseline scores that were conducted on the DSR data (Freeman et al., 1996) each revealed some, albeit different, beneficial effects of SJW (see section 7.4.7.8.1). However, the analysis of the premenstrual scores revealed all of these effects. Moreover, when the DSR-20 data was examined, all psychological symptoms appeared reduced by the luteal phase of the second treatment cycle. When these different analytical approaches were applied to this data (see section 7.4.8.2), the analysis of the premenstrual scores was the only form of analysis to show this effect to be statistically significant. The analysis of the follicular to luteal change
scores appeared to be the least sensitive treatment assessment method. This was the only form of analysis to reveal no beneficial effect of SJW when the DSR scores were examined. Furthermore, this form of analysis, along with the complete (primary) analysis were the only treatment assessment methods to reveal no multivariate effect of treatment when the DSR-20 scores were analysed.

The comparison of the results that were produced from these different commonly used analytical approaches demonstrate that the outcome of a PMS treatment trial can be wholly dependent upon the way in which the data are analysed. As various forms of analysis are commonly used (see section 7.4.8.1), this may partly explain why PMS treatment studies often produce conflicting results (see chapter 6). Moreover, this highlights a need for the different assessment methods to be examined. Although this research suggests that the analysis of premenstrual scores may be the most sensitive method by which to assess a treatment effect, more research in this area is needed, and a consensus reached on what is the most appropriate analytical strategy.

8.3 The DSR-20 as a tool to assess treatment effects
Although the DSR was considered the most appropriate tool by which to capture and track women’s symptoms across the menstrual cycle (see section 1.3.1.3), its use raised certain content related, methodological and statistical concerns (see section 4.3). This measure omits the items anger, aggression and impulsivity, which are distressing symptoms that propel PMS sufferers to seek treatment (see section 1.4.2). Moreover, during the construction of the DSR, the items were reduced into overarching components through the use of Principal Components Analysis (PCA) with the number of factors retained being determined through the use of the K1 rule. However, PCA only produces components that are applicable to the sample that has been studied (Field, 2005), and the K1 rule has been heavily criticised as a factor retention method (Fabrigar et al., 1999; Gorsuch, 1997; Zwick and Velicer, 1986). Hence Chapter 4 explored the factor analysis of this measure to determine whether Freeman et al.’s factor structure could be reproduced in the sample recruited for this research. Moreover, the factor structure was examined during different cycle phases, in other diagnostic groups, and when the items anger, aggression and impulsiveness were added to the scale, to determine whether the DSR is a valid tool for assessing PMS symptoms in phases other than the luteal phase, in samples other than women meeting criteria for PMS, and when
additional symptoms were added to the scale. A new two-factor structure was produced, the DSR-20.

The results presented in Chapter 7 suggest that the DSR-20 was a more sensitive tool to detect treatment effects than the original DSR (Freeman et al., 1996) in the sample recruited for this research (see section 7.5.2.3). The findings were more consistent across the different methods of analysis for the DSR-20 than they were for the DSR (Freeman et al., 1996). In fact, all forms of analyses that were conducted on the DSR-20 data consistently found SJW to benefit physical PMS symptoms. The DSR-20 may have been a more sensitive treatment assessment tool than the DSR (Freeman et al., 1996) in this research because it was produced using a Principal Components Analysis (PCA) that was conducted on data reported by the same sample upon which the treatment analyses were conducted (see section 7.5.2.3). Moreover, the factor retention method (PA) that was used to construct the DSR-20 was sounder than that used to produce the factor structure of the DSR (Zwick and Velicer, 1986; Hayton et al., 2004; Humphreys and Montanelli, 1975; see also section 4.2.1.1). Furthermore, the subscales of the DSR-20 take into account the importance of each item to that subscale by weighting items by their factor score coefficients (see section 4.8.1). Although the DSR-20 appeared to be a more sensitive tool to detect treatment effects in the sample recruited for this research, further studies are needed to validate this factor structure. These studies should use a sound factor retention method (see section 4.2.1.1), and calculate factor scores using an 'exact method,' which takes into account the relevant importance of the items to the factor (Bogner and Wiseman, 2002; Field, 2005; Tabachnick and Fidell, 2007a; see also section 4.8.1).

8.4 The placebo effect in PMS

Women participating in PMS treatment studies show a strong response to placebo treatment (see section 7.1.6). Although the strength of the placebo effect is fairly well documented (Eriksson et al., 1995; Freeman and Rickels, 1999; Magos et al., 1986; Steiner et al., 1995), little research has examined which PMS symptoms are involved, and the length of time needed for symptom reports to return to baseline values i.e. for placebo responding to abate. The results presented in Chapter 8 revealed that although the symptoms of depression, anxiety, verbal aggression, anger and non-planning impulsivity were subject to a placebo effect during menses, these effects had abated by the fourth cycle of placebo treatment. However, mood (DSR) and psychological (DSR-
20) PMS symptoms remained reduced during the menstrual and luteal phases even after this length of time. Therefore, some placebo effects appear to persist for a long duration, suggesting that PMS treatment studies would benefit from having a placebo run-in phase of greater than four cycles when the outcome variables include these psychological symptoms. However, this would be extremely expensive to perform in practice, and may reduce compliance and sample size.

8.5 The characterisation and diagnosis of PMS

Aggression and impulsivity profiles reported by PMS sufferers across the menstrual cycle were compared with those reported by other groups of women, to determine the symptom profiles characteristic of PMS. Focus was on aggressive and impulsive symptoms, as despite PMS sufferers often reporting these symptoms to be amongst their most distressing, little previous research has assessed the variability of aggression and impulsivity across the menstrual cycle (see section 1.4.2).

PMS sufferers, who self-diagnosed with PMS and who met PMS criteria (30% increase criterion) (see section 3.5.2.4), were found to report variable profiles of aggression and impulsivity across the menstrual cycle (see section 5.6.1), and demonstrated a follicular to luteal symptom increase in these symptoms. Debate exists as to whether these cyclical mood changes represent an abnormal symptom profile, or are variations of a normal cyclical pattern (Sanders et al., 1983). Some researchers address this issue by comparing the symptom profiles reported by PMS sufferers, with those reported by women who do not meet PMS criteria (e.g. Bond et al., 2003; De Ronchi et al., 2005; Howard et al., 1994). When this comparison was applied to these symptoms, the symptom profiles reported by PMS sufferers were compared with those reported by control women, who did not self-report problematic PMS symptoms and who did not meet the 30% increase criterion. Comparing the PMS sufferers and controls suggested that the cyclical symptom profiles of the PMS sufferers was indicative of an abnormal process, as the controls were found to report stable levels of aggression and impulsivity across the cycle.

However, this comparison is overly simplistic when liberal PMS criteria (e.g. 30% increase criterion) have been used to form diagnostic groups (see section 5.5.3.1). In line with previous research (e.g. Gallant et al., 1992a; Morse et al., 1988), approximately one third of women who did not self-diagnose with PMS in this research
did not meet the 30% increase criterion (controls). However, approximately two thirds of women who did not self-diagnose with PMS did meet this criterion (PMS non-reporters) (see section 3.5.2.1). As discussed in Section 3.5.2.3, the PMS non-reporters reported premenstrual symptoms at an almost identical severity to the controls during the luteal phase. The main difference between these groups was that the PMS non-reporters reported symptoms at a slightly lower severity than the controls during the follicular phase. The 30% increase criterion bases its diagnosis on the change in symptoms that are reported between the follicular and luteal phases (see section 1.3.1.1). Hence its use resulted in the division of normally cycling women into two groups on the basis of women’s follicular phase symptom severity (see section 3.5.2.3). As such, the use of the 30% increase criterion resulted in the confirmation of a PMS diagnosis in a group of women who should not have been considered to be PMS sufferers, but who should instead have been considered to experience ‘normal cyclicity,’ and included in the control group. Essentially these women were “false positives” (see section 3.5.2.3). The finding that the PMS non-reporters displayed similar physiological (hormone and pro-inflammatory cytokine) profiles to the control group, which were different from those of the PMS sufferers, supports this contention (see section 5.5.5).

The PMS non-reporters and controls were combined into one group of normally cycling women (see section 3.5.3). The aggression and impulsivity profiles reported by this group were compared with those reported by PMS sufferers, to establish what separates ‘normal’ from ‘abnormal’ symptom profiles (e.g. Hallman et al., 1987). Here, the ‘normal symptom profile’ was shown to be cyclical, with the follicular to luteal change in symptoms and premenstrual symptom severity separating ‘normal’ from ‘abnormal’ symptom profiles (see section 5.5.2). The difference in results that were produced when only control women comprised the ‘control’ group compared to when controls and PMS non-reporters were included, demonstrates the importance of taking into account the diagnostic criteria that have been used to form diagnostic groups when selecting a ‘control’ sample in PMS studies. If only control women had been included here, then it would have been concluded that normally cycling women report stable levels of aggression and impulsivity across the cycle. However, if a more representative ‘control’ sample had been selected (controls and PMS non-reporters), the ‘normal symptom profile’ would have been shown to be cyclical, with the follicular to luteal change in symptoms and the premenstrual symptom severity separating ‘normal’ from ‘abnormal’ symptom profiles.
Research should be conducted on PMS samples who self-report problematic PMS symptoms, who are seeking help for the condition, as these are the women whom the research should ultimately identify and benefit. Some researchers retrospectively allocate women to groups after they have completed symptom reports to reduce the impact of expectation effects and increased introspection (e.g. Christensen and Oei, 1988; Dougherty et al., 1998). This can result in PMS non-reporters, who should be included in control groups, actually being considered PMS sufferers following the use of liberal PMS criteria (see section 5.5.3.2). The comparison of the aggression and impulsivity profiles that were reported by the PMS sufferers and the PMS non-reporters demonstrated that the inclusion of PMS non-reporters in PMS groups could result in an underestimation of the premenstrual severity and the cyclicity of symptoms that PMS sufferers experience (see section 5.5.3.2). As such, findings from studies that have retrospectively allocated women to groups following the use of liberal PMS criteria, cannot be generalised to women self-diagnosing with PMS and seeking help for the condition.

Many women with co-morbidities, especially anxiety and depression, present with PMS (Freeman, 2003; Johnson, 2004; Landen and Eriksson, 2003; Pearlstein, 1995). These women may believe they suffer from PMS due to their symptoms becoming magnified premenstrually (see section 1.3.1.4), or due to a tendency to attribute negative premenstrual emotions to internal, hormonal factors, and negative emotions that they experience during other cycle phases to external factors, such as work or other life stresses (Bains and Slade, 1988; Koeske and Koeske, 1975). Just over a quarter of women who self-diagnosed with PMS in this research met criteria for anxiety and/or depression (see section 3.5.1). However, these women do not meet a diagnosis for PMS (APA, 2000), and require different treatment from women who do suffer from the syndrome. Therefore, researchers should identify and exclude these women from participating in PMS studies, so that the findings from these studies can be generalised to PMS sufferers needing treatment in clinical practice. However, as this process is time-consuming, this is often not performed in practice (e.g. Van der Ploeg, 1987; Watts et al., 1987). The comparison of the aggression and impulsivity profiles that were reported by the PMS sufferer and co-morbid groups (see section 3.5.1) demonstrated that the inclusion of women with anxiety and/or depression in PMS groups in research settings could obscure the cyclicity of symptoms that PMS sufferers experience (see section 5.5.3.3).
This variation in the diagnosis of PMS has resulted in heterogeneous samples of PMS sufferers and control groups being studied. This may partly explain the inconsistencies apparent in the literature on PMS (see sections 1.4.1.3 and 7.5).

8.6 Aetiology
The aetiology of PMS is unclear (Halbreich, 2003; see also section 1.5). Various physiological mechanisms have been proposed to be involved. The research presented in this thesis supports a role for some of these mechanistic processes, but not others. Some researchers have proposed that PMS symptoms arise from an oestrogen excess and progesterone deficit (Bäckström and Carstenses, 1984; Mortola, 1998), or through an elevation of prolactin levels (Carroll and Steiner, 1978; Halbreich, 1976). However, in line with other research, the comparison of the hormonal profiles of PMS sufferers and normally cycling women (see section 5.4.4.2.1), did not support the involvement of oestrogen, progesterone (Connolly, 2001; Dennerstein et al., 1998; Rubinow, 1992) or prolactin (Puolakka et al., 1985) in the production of PMS symptoms.

Women with PMS have been shown to have lower calcium levels (Shamberger, 2003; Thys-Jacobs et al., 2007) and a lower bone-mass (Thys-Jacobs et al., 1995) than symptom-free women. Moreover, similarities exist between PMS symptoms and symptoms of hypocalcaemia, such as irritability, fatigue, depression, anxiety and muscle cramps (Thys-Jacobs, 2000; Thys-Jacobs et al., 2007). This has led researchers to propose that PMS symptoms may arise from calcium deficiency. The finding that calcium administration provided benefit to PMS sufferers in the systematic review presented in Chapter 6 supports this suggestion (see section 6.3.1).

Many researchers believe that serotonin plays an important role in the pathophysiology of PMS (Eriksson et al., 2002; Halbreich, 1997, Rapkin, 2002; Steiner et al., 1997a; Yonkers et al., 2008; see also section 1.5.3). The finding that continuous vitamin B6 administration benefited PMS symptoms in the systematic research review (see section 6.3.2) supports the involvement of the serotonergic system in PMS symptom production, as the active form of B6, pyridoxal phosphate, is an essential cofactor in the synthesis of serotonin.

Other researchers have proposed that increased testosterone levels (Rapkin, 2003) and the up-regulation of pro-inflammatory cytokines (Konecna et al., 2003) may be
involved in the production of PMS symptoms. The comparison of the testosterone and cytokine profiles exhibited by PMS sufferers and normally cycling women (see section 5.4.4.1), who did not self-report problematic premenstrual symptoms, supports these suggestions. PMS sufferers were found to have greater testosterone and pro-inflammatory cytokine levels than normally cycling women (controls and PMS non-reporters) during both the follicular and luteal phases (see section 5.4.4.2).

Increased testosterone production may produce PMS symptoms through interaction with the serotonergic system (Steiner et al., 2003b). Cytokines may also produce PMS symptoms through this mechanism of action, as they have been shown to interact with this system (Gemma et al., 1997; Linthorst et al., 1994; Pousset et al., 1996; Silverman et al., 1989). Blood samples were taken from the women who participated in the research presented in this thesis with the intention of measuring whole blood serotonin levels, and levels of its metabolite, 5-HIAA. Unfortunately this was not possible (see section 5.2.2.2.3). Therefore, future research is needed to assess testosterone and pro-inflammatory cytokine levels in women with and without PMS with concurrent measurement of serotonin activity.

8.7 PMS diagnostic criteria

Although various diagnostic criteria are currently used to diagnose PMS, the majority of researchers follow the NIMH (1983) guidelines. However, these criteria have been defined in two ways, and both have been widely adopted (30% increase criterion, modified 30% increase criterion) (see section 1.3.1.1). As discussed in Section 3.5.2.4, the NIMH (1983) criteria were operationalised through the use of the 30% increase criterion for the research presented in this thesis for various reasons. The primary aim of the research presented in this thesis was to determine whether an OTC herbal remedy, SJW, was more beneficial than placebo supplementation for the treatment of premenstrual symptoms (see Chapter 2). The 30% increase criterion identified women who were experiencing PMS symptoms at a mild severity (see section 3.5.2.1), a severity representative of the women who buy OTC preparations to treat their symptoms (Bendich, 2000; Domoney et al., 2003; Eriksson et al., 2002; Freeman et al., 2002; Fugh-Berman and Kronenberg, 2003; Girman et al., 2003), and who are advised to try this form of treatment by their clinicians, based on their judgment of PMS severity (Connolly, 2001; Johnson, 2004; see also section 1.3.1.1). In contrast, the modified 30% increase criterion only identified PMS sufferers who reported extremely
severe PMS symptoms (see section 3.5.2.2), a group of women who would be less likely to buy OTC preparations to treat their symptoms, and more likely to receive prescribed medication (Dimmock et al., 2000; Eriksson et al., 2002; Johnson, 2004; Wyatt et al., 2002). Moreover, the 30% increase criterion identified the majority of women who self-diagnosed with PMS, and as such identified women representative of the women likely to buy OTC preparations. The modified 30% increase criterion only identified a minority of these women (see section 3.5.2.4). As such, its use would have failed to identify the majority of the women likely to try this form of treatment to relieve symptoms they attribute to the premenstrual phase. Furthermore, the 30% increase criterion is often used to diagnose PMS in RCTs assessing the effectiveness of OTC preparations for PMS symptoms (e.g. Bryant et al., 2005; De Souza et al., 2000; Facchinetti et al., 1991; Freeman et al., 2002; Hicks et al., 2004; Pearlstein et al., 1997; Sayegh et al., 1995; Walker et al., 1998). For example, the researchers who developed the DSR (Freeman et al., 1996) used the 30% increase criterion to diagnose PMS in their treatment trial of the carbohydrate-rich beverage Escape. In contrast, the modified 30% increase criterion is usually used when researchers desire PMS sufferers who are experiencing severe symptoms, for example when the effectiveness of treatments such as alprazolam (e.g. Evans et al., 1998) and SSRIs (e.g. Tung-Ping et al., 1997) are being assessed. Although it was considered appropriate to use the 30% increase criterion as the method by which to diagnose PMS in this research, its use renders the findings that were produced, applicable only to PMS sufferers experiencing PMS symptoms at a mild to moderate severity. Future research is needed if researchers wish to generalise the findings that have arisen from this research to PMS sufferers experiencing PMS symptoms at a greater severity. In particular, it would be interesting for future studies to assess the potential impact that SJW may have for PMS sufferers experiencing severe premenstrual symptoms.

8.8 Limitations of the research

8.8.1 Standardisation of herbal products

It is unclear which bioactive compound is/are responsible for the beneficial effect of SJW on depression (Muller et al., 1997; Nangia et al., 2000; Teufel-Mayer and Gleitz, 1997; Wheatley, 1998; Zanoli, 2004). Both hypericin (Bennett et al., 1998; Gaster and Holroyd, 2000; Mennini and Gobbi, 2004) and hyperforin (Cervo et al., 2002; Chatterjee et al., 1998; Laakmann et al., 1998; Muller et al., 1998) have been proposed to be efficacious. Although the commercially available SJW products that have been
tested in research settings have varied greatly in their hyperforin content (Mennini and Gobbi, 2004), they have usually been standardised according to the percentage of hypericin that they contain (Gaster and Holroyd, 2000; Linde et al., 2005). Most studies that have assessed the efficacy of SJW for depression have reported that the tablets used were standardised to 0.3% hypericin (e.g. Montgomery et al., 2000; Shelton et al., 2001; Szegedi et al., 2005). In particular, the pilot study (Stevinson and Ernst, 2000) and the RCT (Hicks et al., 2004) which assessed the efficacy of SJW for PMS also reported that their tablets were standardised to this level (0.3%). Therefore, the manufacturers of the SJW tablets that were administered in the RCT presented in Chapter 8 were requested to standardise the product to 0.3% hypericin. However, the subsequent batch analysis provided by the manufacturers, Lichtwer Pharma, indicated that the tablets contained approximately half the required amount (0.18%). Whilst this falls within the acceptable range for commercially available SJW, it highlights the great variability permitted by the licensing authorities for herbal products. Moreover, this has implications for the findings reported by studies which have relied solely on pack based standardisation information without performing a batch analysis of the product administered in the study. There is a need for tighter controls to be implemented in the production of herbal products, for these controls to be policed by regulatory bodies and for researchers to report the actual percentage of hypericin and hyperforin that is contained in the products of SJW that they administer to participants.

8.8.2 Demand characteristics
All groups of women who were recruited for the research presented in this thesis were informed that the study focus involved the menstrual cycle. Previous research has demonstrated that women who do not consider themselves to have problematic PMS symptoms (Gallant et al., 1992b) report significantly more negative psychological and somatic symptoms during the premenstrual and menstrual phases if they know that menstrual cycle symptomatology is being assessed (AuBuchon and Calhoun, 1985). Knowledge of the study focus in this research may have resulted in demand characteristics and expectations increasing cyclicity of symptom reporting in the group of normally cycling women (PMS non-reporters and controls), since this group did not consider themselves to suffer from PMS. However, it would not have been practical, and possibly ethically questionable, to disguise the focus of the study. The normally cycling women were identified through their participation in the menstrual cycle study (see section 3.4.1.2). To be eligible to participate in this study, women were required
not to self-diagnose with PMS. If the study focus had been disguised, it would have been necessary for the women to complete symptom reports before asking them whether they considered themselves to suffer from PMS. This would have resulted in both women who did and who did not self-diagnose with PMS taking part. As the study was solely conducted by the PI, it would not have been pragmatic to recruit a larger sample size. Therefore, disguising the study focus would have resulted in fewer women who did not self-diagnose with PMS participating, which would have resulted in less women being classified as normally cycling women (PMS non-reporters and controls). This would have reduced the power in the analyses that were conducted on these samples. Moreover, in order to disguise the study focus, additional measures would have needed to be administered to reduce the likelihood of women deducing the study focus, which may have reduced compliance. Nevertheless, demand characteristics in the group of normally cycling women were probably minimized by informing these women that the study required women who experienced ‘normal’ menstrual cycles. Therefore, these women may have actually reported less cyclicity of symptoms in order to please the experimenter.

Women volunteering for the clinical trial who met the initial entry criteria (see section 7.2.4) were informed that the symptom reports that they completed during the screening cycles would be used to assess their eligibility to enter the treatment phase of the clinical trial (see section 3.4). Therefore, these women may have been motivated to report higher symptom levels during these cycles in order to ensure that they would be eligible to enter the clinical trial (Gallant et al., 1992b). This may have resulted in an inflated estimation of their symptom severity. If this were the case, then this has implications for the way in which the placebo effect was examined in this research and in previous studies (e.g. Freeman et al., 1995; Freeman et al., 1999). Symptom profiles reported during placebo treatment were compared with those reported during screening cycles which were also used to assess study eligibility. If women had exaggerated their symptom severity during the screening cycles, then this would result in an overestimation of the placebo effect. Although any such effects could have been reduced by only offering women the opportunity to take part in the clinical trial after completing the screening cycles, the resultant sample would not have been representative of women seeking treatment for PMS, in whom the efficacy of the treatment was to be assessed.
8.8.3 Co-morbidities

The diagnosis of PMS requires that symptoms are limited to the luteal phase (Freeman, 2003; Gladys and Silverman, 2000). Many women who present with severe PMS are found not to be suffering from the syndrome, but are found to be suffering from another psychiatric disorder, most commonly anxiety and depression (Freeman, 2003; Johnson, 2004; Landen and Eriksson, 2003; Pearlstein, 1995; see also section 1.3.1.4). Therefore, for the research presented in this thesis, women who were found to report clinical levels of anxiety and/or depression during the follicular phase were not considered to be PMS sufferers, and instead were categorised as co-morbid. However, other psychiatric conditions that were not directly assessed, such as bulimia nervosa, substance abuse and seasonal affective disorder, can become exacerbated premenstrually (Freeman, 2003; Pearlstein, 1995). Therefore, it is possible that some of the women who were categorised as PMS sufferers in the research presented in this thesis were not suffering from PMS but were suffering from these co-morbid conditions. Although it is unlikely that this was the case for the PMS sufferers who participated in the RCT, as the study doctor (Dr Julie Ayres) had ensured that these women were in good psychological health before they were entered onto the trial, it is possible that some of the PMS sufferers who were screened by post were suffering from these co-morbid conditions. Although it may have been possible to detect the presence of these conditions through the administration of additional questionnaires, it was thought that the protocol was already sufficiently demanding, and that this would have reduced compliance and sample size. Moreover, these women were already asked to complete questionnaires relating to anxiety and depression. Therefore, it was considered ethically questionable to ask these women about additional topics of a sensitive nature, as these women were participating by post, and did not have the opportunity to meet with the study doctor.

8.8.4 Weekly versus daily diary collection

All of the women who participated in the research presented in this thesis were asked to complete daily diary booklets throughout their participation. Each daily diary booklet contained seven copies of the DSR (Freeman et al., 1996), and one copy of the BDI (Beck et al., 1961), STAI (Spielberger, 1983b), BPAQ (Buss and Perry, 1992) and BIS-11 (Patton et al., 1995) (see section 3.1.1). Women were asked to return each daily diary booklet as soon after its completion as possible, to minimise the possibility of retrospective reporting (Keenan et al., 1992; see also section 3.4.5). Participants were explicitly informed that they were to complete one copy of the DSR each evening, and
to provide the date in the space provided (see Appendix 3.1). However, as each booklet contained seven copies of the DSR, it must be acknowledged that some women may not have completed their daily ratings in this prospective manner. Furthermore, although each copy of the DSR was presented on a separate page to minimise the likelihood of women spotting patterns in their symptom reporting (see section 3.1.1.1), it is possible that some ratings may have been biased by participants looking at their previous diary completions from that week. Although some researchers ask women to return their ratings daily (e.g. Evans et al., 1999), it was thought that applying this method of data collection to a study ten menstrual cycles in length, would have led to greater study burden, a higher attrition rate, and a larger amount of missing data. In order to reduce the likelihood of ratings being biased in this way, future studies could ask women to complete their ratings on the internet, and to submit these ratings to the study’s website at the end of each day.

8.8.5 Absence of data on appetite, alcohol and caffeine intake

Previous research has shown that PMS sufferers demonstrate variability in certain behaviours across cycle phases that were not directly measured in this research. For example, PMS sufferers have been shown to demonstrate an increase in food cravings (for both sweet and savoury foods) and energy intake (particularly from foods with a high fat content) in their premenstrual phase compared to their postmenstrual phase, and to experience a greater follicular to luteal increase in these behaviours than women without PMS (Dye and Blundell, 1997; Reed et al., 2008). Moreover, female drinkers who experience premenstrual symptoms have been shown to consume more alcohol during the premenstrual phase than relatively asymptomatic women (Mello et al., 1990). Furthermore, women with moderate to severe premenstrual symptoms have been shown to demonstrate a premenstrual increase in caffeine intake, in comparison to women with minimal premenstrual symptoms, who showed stable caffeine consumption across the cycle (Rossignol et al., 1991). Given that PMS sufferers have been shown to demonstrate a premenstrual increase in alcohol, caffeine and energy consumption during the premenstrual phase, one limitation of the research presented in this thesis, is that these behaviours were not directly measured and so their impact on the symptoms reported could not be assessed.
8.8.6 Multiple statistical testing

If one hundred statistical tests are conducted, on five occasions, the test statistic will by chance, falsely be large enough to infer the presence of a genuine effect in the population (Type I error) (Field, 2005). A conservative approach was taken throughout this thesis to minimise the chances of rejecting the null hypothesis in this way. For example, the self-report data presented in Chapter 5 was analysed using doubly multivariate analysis of variance models, for the simultaneous assessment of the BPAQ (Buss and Perry, 1992) and BIS-11 (Patton et al., 1995) subscales. This was done to avoid the inflation of the Type I error rate that occurs when separate analyses are conducted on each subscale (Bray, 1995; Field, 2005; Tabachnick and Fidell, 2007a; see also section 5.3.1). Moreover, throughout this thesis, when simple effects and interaction contrasts analyses were performed to explore the results from RM ANOVAs and doubly multivariate MANOVAs, Bonferroni corrections were applied in such a way that the \( \alpha \) levels were divided by the number of tests that could have been performed, rather than by the number of tests that were actually conducted (see section 5.3.1). However, it must be conceded that a large number of statistical tests were conducted on a relatively small sample size (n=34), in the exploratory analyses that were conducted in Chapter 7 following performance of the primary analysis (see section 7.4.8). This raises the possibility that some of the findings from these analyses may have arisen from an increased familywise error rate, rather than from a genuine effect (Rosenthal and Rosnow, 1999). However, in order to minimise the over interpretation of chance findings, the data was interpreted through identifying patterns in the occurrence of significant results. Certain key consistencies were found between the results that were produced from the various commonly used analytical strategies that were applied to the DSR (see section 7.5.2.1), and the DSR-20 (see section 7.5.2.2). Furthermore, the results that arose from the analyses that were conducted on the DSR were compatible with those conducted on the DSR-20 (see section 7.5.2.3).

8.9 Future Research

Whilst this thesis has extended current knowledge regarding the treatment, diagnosis and aetiology of PMS, future studies are required to verify and substantiate some of these findings.

The research presented in this thesis suggests that SJW confers some benefit for PMS sufferers who find physical and behavioural premenstrual symptoms problematic.
However, this widely consumed herbal remedy appeared to have little impact upon mood related PMS symptoms, the symptoms which cause PMS sufferers the greatest distress (Corney and Stanton, 1991; Freeman, 2003; Halbreich et al., 2007). Nevertheless, the results presented in Chapter 7 suggest that SJW may have potential to reduce this type of PMS symptom subject to longer study, since mood related symptoms appeared to improve towards the end of the second cycle of SJW treatment. Therefore, further studies are needed to clarify whether SJW could be an effective alternative therapy to treat PMS sufferers who find mood related symptoms problematic, by determining whether this type of PMS symptom shows an enduring benefit following longer treatment duration.

The role of increased testosterone (Rapkin, 2003) and pro-inflammatory cytokine levels (Konecna et al., 2003) in PMS symptom production was supported by the research presented in this thesis (see section 5.6.3). However, no previous research has prospectively examined cytokine levels across the cycle of PMS sufferers, and the results from previous studies that have assessed testosterone levels have been contradictory (Bäckström and Aakvaag, 1981; Dougherty et al., 1997a; Eriksson et al., 1992; 1994; Watts et al., 1985). Therefore, more studies are needed to clarify whether PMS sufferers do have elevated testosterone and pro-inflammatory cytokine levels. Increased testosterone and pro-inflammatory cytokine levels appeared correlated in this research i.e. both occurred together. However, future research is needed to investigate the direction and causality of such a relationship. Although it is possible that increased testosterone and pro-inflammatory cytokine levels produce PMS symptoms through interaction with the serotonergic system (see section 1.5.3), 5-HT and 5-HIAA levels could not be determined in this research (see section 5.3.2.2.3). Therefore, future research assessing testosterone and pro-inflammatory cytokine levels in women with and without PMS concurrently with serotonin activity is warranted.

8.10 Conclusion

PMS is a distressing condition that affects many women of reproductive age. The findings presented in this thesis add significantly to the understanding of the syndrome, in terms of its characteristics, diagnosis and treatment. This research highlights the need to carefully screen women using appropriate and parsimonious symptom scales, which differentiate women with PMS from women with clinical anxiety and/ or depression. Moreover, this thesis demonstrates the importance of utilizing appropriate experimental
design and analysis. The rigorous methodological approach adopted in the research
presented in this thesis differs from past research, especially with regard to the
assessment of herbal treatment effects in PMS studies. The results presented in this
thesis merit further investigation and corroboration to explicate further the questions
and issues examined. This is particularly the case with regards to the aetiology of the
syndrome and the methods for the assessment of efficacy in PMS treatment studies.
REFERENCES


related to interferon-alpha-induced changes in the serotonergic system. *Journal of Clinical Psychopharmacology*, 22, 86–90


alpha administration on some behavioural, neurochemical, endocrine and immune parameters in the rat. *Neuroscience*, 84, 923-933


Horn, J.L. (1965). A rationale and test for the number of factors in factor analysis. Psychometrika, 30, 179-185


Horrobin, D.F. (1983). The role of essential fatty acids and prostaglandins in the premenstrual syndrome. Journal of Reproductive Medicine, 28(7), 465-8


Hsiao, C-C., Liu, C-Y., & Hsiao, M-C. (2004). No correlation of depression and anxiety to plasma estrogen and progesterone levels in patients with premenstrual dysphoric disorder. *Psychiatry and Neurosciences*, 58, 593-599


McIntosh, J.E., Matthews, C.D., Crocker, J.M., Broom, T.J., & Cox, L.W. (1980). Predicting the luteinizing hormone surge: Relationship between the duration of the follicular and luteal phases and the length of the human menstrual cycle. Fertility and Sterility, 34 (2), 125-130


Moos, R. (1968). The development of a menstrual distress questionnaire. *Psychosomatic Medicine, 30*(6), 853-867


NIMH. (1983). Workshop on Pre-menstrual Syndrome. Rockville, MD: Co-sponsored by the Centre for Studies of Affective Disorders and the Psychobiological Processes and Behavioral Medicine Section, 14-15


Quah-Smith, J.I., Tang, W.M., & Russell, J. (2005). Laser acupuncture for mild to moderate depression in a primary care setting—a randomised controlled trial. *Acupuncture in Medicine, 3*, 103-11


Appendix 3.1 DSR (Freeman et al., 1996) administered on a daily basis

Day: M T W T F S S Date: ___/___/___

Time of Day: ______________

Please indicate your feelings and experiences today by checking the box that best describes your experience by using the following scale:

0 = not present at all
1 = minimal, only slightly apparent to you
2 = moderate, aware of symptom but does not affect daily routine
3 = a lot, continuously bothered by symptom and/or symptom interferes with daily activity
4 = severe, symptom is overwhelming and/or unable to carry out daily activity

Even if none of the categories is exactly correct, please choose the one that is most similar to your experience.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Do you have your period today? No _____

When did your period start? Day M T W T F S S Date: ___/___/___
Appendix 3.1 continued: Questionnaires administered on a weekly basis

BDI (Beck et al., 1961)

For each question, please put a tick by the statement that best describes how you have felt this week.

1.)
- I do not feel sad
- I feel blue or sad
- I am blue or sad all the time and I can’t snap out of it
- I am so sad or unhappy that it is very painful
- I am so sad or unhappy that I can’t stand it

2.)
- I am not particularly pessimistic or discouraged about the future
- I feel discouraged about the future
- I feel I have nothing to look forward to
- I feel that I won’t ever get over my troubles
- I feel that the future is hopeless and that things cannot improve

3.)
- I do not feel like a failure
- I feel I have failed more than the average person
- I feel I have accomplished very little that is worthwhile or that means anything
- As I look back on my life all I can see is a lot of failures
- I feel I am a complete failure as a person (parent, husband, wife)

4.)
- I am not particularly dissatisfied
- I feel bored most of the time
- I don’t enjoy things the way I used to
- I don’t get satisfaction out of anything anymore
- I am dissatisfied with everything

5.)
- I don’t feel particularly guilty
- I feel bad or unworthy a good part of the time
- I feel quite guilty
- I feel bad or unworthy practically all the time now
- I feel as though I am very bad or worthless

6.)
- I don’t feel I am being punished
- I have a feeling that something bad may happen to me
- I feel I am being punished or will be punished
- I feel I deserve to be punished
- I want to be punished

7.)
- I don’t feel disappointed in myself
- I am disappointed in myself
- I don’t like myself
- I am disgusted with myself
- I hate myself

8.)
- I don’t feel I am any worse than anybody else
- I am very critical of myself for my weaknesses or mistakes
- I blame myself for everything that goes wrong
- I feel I have many bad faults
9.)
- I don’t have any thoughts of harming myself
- I have thoughts of harming myself but I would not carry them out
- I feel I would be better off dead
- I have definite plans about committing suicide
- I feel my family would be better off if I were dead
- I would kill myself if I could

10.)
- I don’t cry any more than usual
- I cry more now than I used to
- I cry all the time now. I can’t stop it
- I used to be able to cry but now I can’t cry at all even though I want to

11.)
- I am no more irritated now than I ever am
- I get annoyed or irritated more easily than I used to
- I feel irritated all the time
- I don’t get irritated at all at the things that used to irritate me

12.)
- I have not lost interest in other people
- I am less interested in other people now than I used to be
- I have lost most of my interest in other people and have little feeling for them
- I have lost all my interest in other people and don’t care about them at all

13.)
- I make decisions about as well as ever
- I am less sure of myself now and try to put off making decisions
- I can’t make decisions any more without help
- I can’t make any decisions at any more

14.)
- I don’t feel that I look any worse than I used to
- I am worried that I am looking old or unattractive
- I feel that there are permanent changes in my appearance and they make me look unattractive
- I feel that I am ugly or repulsive looking

15.)
- I can work about as well as before
- It takes extra effort to get started at doing something
- I don’t work as well as I used to
- I have to push myself very hard to do anything
- I can’t do any work at all

16.)
- I can sleep as well as usual
- I wake up more tired in the morning than I used to
- I wake up 1-2 hours earlier than usual and find it hard to get back to sleep
- I wake up early every day and can’t get more than 5 hours sleep

17.)
- I don’t get any more tired than usual
- I get tired more easily than I used to
- I get tired from doing anything
- I get too tired to do anything

18.)
- My appetite is no worse than usual
- My appetite is not as good as it used to be
- My appetite is much worse now
- I have no appetite at all any more

19.)
- I haven’t lost much weight, if any, lately
- I have lost more than 5 pounds
- I have lost more than 10 pounds
- I have lost more than 15 pounds
20.
- I am no more concerned about my health than usual
- I am concerned about aches and pains or upset stomach or constipation or other unpleasant feelings in my body
- I am so concerned with how I feel or what I feel that it's hard to think of much else
- I am completely absorbed in what I feel

21.
- I have not noticed any recent change in my interest in sex
- I am less interested in sex than I used to be
- I am much less interested in sex now
- I have lost interest in sex completely
Appendix 3.1 continued. STAI (Spielberger et al., 1983b)

Directions: A number of statements which people might use to describe themselves are given below. Read each statement and then circle the number to the right of the statement to indicate how you have felt over the past week. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which best describes how you have felt this week.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Not at all</th>
<th>Somewhat</th>
<th>Moderately So</th>
<th>Very Much So</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I felt calm</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. I felt secure</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. I felt tense</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4. I felt strained</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5. I felt at ease</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6. I felt upset</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7. I worried over possible misfortunes</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8. I felt satisfied</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>9. I felt frightened</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>10. I felt comfortable</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>11. I felt self-confident</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>12. I felt nervous</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>13. I was jittery</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>14. I felt indecisive</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>15. I was relaxed</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>16. I felt content</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>17. I was worried</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>18. I felt confused</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>19. I felt steady</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>20. I felt pleasant</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Directions: A number of statements which people might use to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you generally feel. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe how you generally feel.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Almost Never</th>
<th>Sometimes</th>
<th>Often</th>
<th>Almost Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>21. I feel pleasant</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>22. I feel nervous and restless</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>23. I feel satisfied with myself</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>24. I wish I could be as happy as others seem to be</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>25. I feel like a failure</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>26. I feel rested</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>27. I am “calm, cool, and collected”</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>28. I feel that difficulties are piling up so that I cannot overcome them</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>29. I worry too much over something that really doesn’t matter</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>30. I am happy</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>31. I have disturbing thoughts</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>32. I lack self-confidence</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>33. I feel secure</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>34. I make decisions easily</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>35. I feel inadequate</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>36. I am content</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>37. Some unimportant thought runs through my mind and bothers me</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>38. I take disappointments so keenly that I can’t put them out of my mind</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>39. I am a steady person</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>40. I get in a state of turmoil as I think over my recent concerns and interests</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Appendix 3.1 continued. BIS-11 (Patton et al., 1995)

Please answer the following questions as accurately as possible, indicating how much each statement refers to you and your thoughts and behaviour in the last week. Please circle only one number for each statement.

Circle 1 if the behaviour in question has been extremely uncharacteristic of yourself in the last week (i.e. you rarely or never have done this or felt this way in the last week) and circle 4 if the behaviour has been extremely characteristic of yourself in the last week (i.e. you almost always have done this or always felt this way in the last week).

<table>
<thead>
<tr>
<th>In the last week I have:</th>
<th>Rarely/Never</th>
<th>Occasionally</th>
<th>Often</th>
<th>Almost Always/Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Planned tasks carefully</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. Done things without thinking</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. Made up my mind quickly</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4. Been happy-go-lucky</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5. Not “paid attention”</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6. Had “racing” thoughts</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7. Planned trips well ahead of time</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8. Been self-controlled</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>9. Concentrated easily</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>10. Saved regularly</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>11. “Squirmed” at plays or lectures</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>12. Been a careful thinker</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>13. Planned for job security</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>14. Said things without thinking</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>15. Liked to think about complex problems</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>16. Changed jobs</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>17. Acted “on impulse”</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>18. Got easily bored when solving thought problems</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>19. Acted on the spur of the moment</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>20. Been a steady thinker</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>21. Changed residences</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>22. Bought things on impulse</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>23. Only been able to think about one problem at a time</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>24. Started a new hobby</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>25. Spent more than I earn</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>26. Often had outside thoughts when thinking</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>27. Been more interested in the present than the future</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>28. Been restless in the theatre or in lectures</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>29. Enjoyed puzzles</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>30. Been future oriented</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Appendix 3.1 continued. BPAQ (Buss and Perry, 1992)

Please answer the following questions as accurately as possible, using the scales provided, by indicating how much each statement applies to your thoughts and behaviour in the last week. Please circle one number for each statement. Circle 1 if the statement is extremely characteristic of you during the last week and the number 5 if the behaviour is extremely uncharacteristic of you during the last week.

<table>
<thead>
<tr>
<th>In the last week:</th>
<th>Extremely Characteristic</th>
<th>Extremely Uncharacteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. My friends have said that I’m somewhat argumentative</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>2. Sometimes I have flown off the handle for no good reason</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>3. I have told my friends openly when I have disagreed with them</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>4. Some of my friends have thought that I was a hothead</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>5. Once in a while I couldn’t control the urge to strike another person</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>6. Given enough provocation, I could have hit another person</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>7. At times I felt I have got a raw deal out of life</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>8. I have sometimes been eaten up with jealousy</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>9. I have sometimes felt like a power keg ready to blow</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>10. I have felt that other people always seem to get the breaks</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>11. I have wondered why sometimes I feel so bitter about things</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>12. If somebody had hit me, I would have hit back</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>13. I have sometimes felt that people are laughing at me behind my back.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>14. When people annoyed me, I told them what I think of them</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>15. When frustrated, I have let my irritation show</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>16. If I had to resort to violence to protect my rights, I have</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>17. I have been an even-tempered person</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>18. There are people who pushed me so far, we came to blows</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>19. I have flared up quickly but got over it quickly</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>20. I couldn’t help getting into arguments when people disagreed with me</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>21. I have been suspicious of overly friendly strangers</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>22. I have had trouble controlling my temper</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>23. I could not think of a good reason for hitting a person</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>24. I have threatened people I know</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>25. I have become so mad that I have broken things</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>26. I have often found myself disagreeing with people</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>27. When people were especially nice, I have wondered what they want.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>28. I have got into fights a little more than the average person</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>29. I know that ‘friends’ have talked about me behind my back</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 3.2: Preparations serving as exclusion criteria

- HIV protease inhibitors (indinavir, nelfinavir, ritonavir, saquinavir)
- HIV non-nucleoside reverse transcriptase inhibitors (efavirenz, nevirapine)
- Warfarin
- Cyclosporin
- Photoreactive agents
- Anticonvulsants (e.g. carbamazepine, phenobarbitone, phenytoin)
- Digoxin
- Theophylline
- Triptans (sumatriptan, naratriptan, rizatriptan, zolmitriptan, frovatriptan)
- SSRIs (citalopram, fluoxetine, fluvoxamine, paroxetine, sertaline etc.)
- Psoralen (oxsoralen, trisoralen, ciprofloxacin, tetracycline)
- Lithium
- Antipsychotic drugs
- Neuroleptic drugs
- Cimetidine
- Oral corticosteroids
- Thyroid hormones
- Immunosuppressive agents
Appendix 3.3: Menstrual cycle study handbook

The Study
Research suggests that most women who menstruate experience some changes in their feelings and behaviours across the menstrual cycle. This research aims to gain a better understanding about these mood and behavioural changes. This research will also look at how women’s serotonin (a chemical in the brain that influences behaviour) and cytokine (a protein involved in the body’s immune responses) levels change across the menstrual cycle.

What are the possible benefits of taking part?
- You will be making an important contribution to new scientific research
- You will have the opportunity to experience taking part in research
- You will gain feedback about your feelings and behaviour across three menstrual cycles
- You will be compensated for your time and travel

What is involved in taking part in the study?
Your involvement in the study will last for three consecutive menstrual cycles. During this time you will be asked to complete daily symptom diaries and to visit the researcher three times at the Institute of Psychological Sciences, University of Leeds, to perform a simple computer task. On two of these occasions, you will also be asked to have a small blood sample taken. The following pages will discuss what is involved in the study in more detail.

Weekly Symptom Diaries
Throughout the study you will be given weekly diaries which will ask you about various feelings and behaviours. You will be asked to complete a checklist about your feelings and behaviour at the end of each day, which will involve rating 20 feelings/experiences on a scale from 0 (not present at all) to 4 (severe). At the end of the week, you will be asked to complete a longer questionnaire about your feelings and behaviour in more detail, and to return each daily diary booklet in a Freepost envelope.

Computer Task (Emotional Stroop Task)
On the three occasions that you visit the researcher, you will be asked to perform a simple computer task. During this task, various words will appear on the screen, one by one, in blue, red, green or yellow. You will be asked to simply say what colour the word is written in as quickly as possible, whilst ignoring its meaning. This task takes approximately 5 minutes to complete. You will be guided through this task by the researcher (Sarah Canning) on your first visit.
Blood Tests
You will be asked to have 2 blood samples taken during course of the study. These will be taken by the researcher (Sarah Canning) when you come in for your second and third visits. We are interested in how your serotonin and cytokine levels change across your menstrual cycle. Cytokines are a diverse group of proteins involved in immune responses to injury and infection. Serotonin is a brain neurotransmitter. Neurotransmitters are chemical messengers that deliver information from one part of the brain to another.

Visits
You will be asked to visit the researcher 3 times throughout the course of the study, at approximately the same time in the day. On your first visit, you will be asked to sign a consent form to say that you have had the opportunity to ask any questions that you may have and are happy to take part in the study. You will be introduced to the computer task that you will be asked to perform on the later visits. On the second and third visits, you will also be asked to have a small blood sample taken.

Your first visit can occur at any time during your menstrual cycle. Your second visit will need to take place during your postmenstrual phase of your third menstrual cycle (5-10 days after your third period begins). Your third visit will need to take place during your premenstrual phase of this cycle (within 5 days of your fourth period starting). We will be able to estimate when you are in your postmenstrual and premenstrual phases from the length of your normal menstrual cycle.

We will ask you to let us know when you begin your third period, so that we can check that your second visit has been booked during your postmenstrual phase. We will also telephone you 5 days after your premenstrual visit to make sure that you have begun your period. If you have not, you will be asked to return to the Institute of Psychological Sciences, University of Leeds, to have an additional visit, where you will be asked to perform the computer task again and have a small blood sample taken.

Am I suitable to take part in the research?

Yes
If you are between the ages of 18-45
If you have reasonably regular menstrual cycles

NO
If you have been diagnosed with, or feel that you may be suffering from depression, anxiety or the premenstrual syndrome
If you are currently taking antidepressants
If you are using hormonal preparations including the oral contraceptive pill, mirena coil or HRT
If you are pregnant or breast feeding.
General Information

- This study has been set up by the University of Leeds. Your visits would be to the Institute of Psychological Sciences, University of Leeds.
- You will be allocated a participant number, which will be used instead of your name to identify your data. This ensures that all data remains anonymous. Only the principal investigator (Sarah Canning) and her supervisors (Dr Louise Dye and Dr Mitch Waterman) will know what number corresponds to your data. This is in line with the ethical guidelines set out by the British Psychological Society.
- The results of this study may be published in an academic journal and be presented at academic conferences. However, no one will be able to tell that you participated in the study. If a paper is published, you will be notified of this, so that you can obtain a copy.
- It is your right to withdraw from this study at any stage without giving a reason as to why. You are also not obliged to answer any question you do not wish to. If you do choose to leave the study, please inform Sarah Canning immediately.
- Once the study has been completed and the data analysed, you will be given detailed feedback, showing you how your feelings and behaviour change over the course of your menstrual cycle.
- In the unlikely event that you feel distressed after completing one of the weekly diaries, please contact one of the following:
  Samaritans: Tel: (0113) 245 6789, Email: jo@samaritans.org
  Nightline: Tel: (0113) 380 1381, Email: nightline@leeds.ac.uk
- If there is anything you would like advice on throughout the study, please contact Sarah Canning at Leeds University (Tel (0113) 3433198; Email s.e.canning@leeds.ac.uk). This will be your point of contact throughout the study:
Appendix 3.4: RCT Information Sheet
Testing the effectiveness of St John’s Wort (SJW) as a treatment for PMS

You are being invited to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Please take time to decide whether or not you wish to take part.

What is the purpose of the study?
This study is interested in testing whether St John’s Wort, a herbal remedy that is available over the counter, is an effective treatment for Premenstrual Syndrome (PMS). It will look at the effect taking SJW has on the major symptoms of PMS, including depression, anxiety, aggression and impulsivity levels. This study will last for the duration of ten of your menstrual cycles.

Why have I been chosen?
You have been chosen for this study as you have either shown an interest in the research or attend the premenstrual clinic at Leeds General Infirmary, and seem to suffer from premenstrual symptoms. This study will test approximately 25 women in total.

Do I have to take part?
It is up to you to decide whether or not to take part. If you do decide to take part, you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part, you are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive.

What will happen to me if I take part?
As a participant in the research, you would be required to take part for 10 consecutive menstrual cycles. As women differ in their menstrual cycle length, the study will take approximately one year to complete. During the time you take part in the study, you would have to fill in a two minute daily symptom checklist, which simply names 20 symptoms (e.g. aches, anger) and asks you to rate the strength that you have suffered from these each day. You would also be required to complete a questionnaire bundle at the end of each week, which would ask you about feelings of depression, anxiety, impulsivity and aggression in more detail. This would take around 20 minutes to complete. You would be required to visit the researcher a total of 11 times during the study. On all of these occasions you would be required to perform a simple cognitive task, which would involve you looking at a series of words written in different colours, where you would be asked to simply say what colour the word is written in. On these occasions you would see the principal investigator and be able to ask any questions you might have. On 10 of these visits you would be asked to have a small blood sample taken (30mls). Your visits to the clinic would last approximately 30 minutes.

The study will involve comparing the effectiveness of SJW with placebo tablets for the treatment of PMS. Placebo tablets are dummy tablets that look like the real thing, but are not. They contain no active ingredients. You will not know when you are receiving SJW tablets and when you are receiving placebo tablets, although you will be given both during the course of the study. As the trial is double-blind, your doctor will not know this either, but would be able to find out if this was needed.

By taking part in this study, you would be expected to complete the daily symptom checklist, the weekly questionnaire bundle and to visit the researcher on these 11 occasions.

What would I have to do?
This study will not effect what you normally eat or drink. However, you should avoid being exposed to excessive sunlight or artificial UVA light. Also, SJW cannot be taken with a range of medications, including HIV protease inhibitors, HIV non-nucleoside reverse transcriptase inhibitors, warfarin, cyclosporine, oral contraceptives, photoreactive agents, anticonvulsants, digoxin, theophylline, triptans, SSRIs, psoralens, lithium, antipsychotic drugs, neuroleptic drugs, cimetidine, oral corticosteroids, thyroid
hormones or immunosuppressive agents. However, you would not have been given this information sheet if you had said that you were taking any of these medications in the recruitment questionnaire. If you become pregnant during the study, then you would be asked to withdraw from the study, as your menstrual cycles would be affected.

**What is the drug or procedure that is being tested?**
St John’s Wort (SJW) has been used since ancient times but has become increasingly popular in the last ten years in Europe and the US. It is a herbal remedy and is available over the counter. When taking SJW, you would be given a dose of 900mg/day, in tablet form. You will be given a card with the details of this study written on it, and you would be asked to carry this card with you at all times.

**What are the alternatives for diagnosis or treatment?**
PMS is a poorly understood condition which has few proven treatments, including many unlicensed over-the-counter preparations – this is the reason why we are carrying out this study. If you were still uncertain whether to take part in the study you could be seen by your family doctor and/or be referred to our PMS clinic instead.

**What are the side effects of any treatment received when taking part?**
No serious side effects have been reported in previous studies with SJW, but minor side effects include tummy discomfort, tiredness, dry mouth, dizziness, skin rash, and hypersensitivity to sunlight.

**What are the possible disadvantages and risks of taking part?**
The effects of taking SJW in pregnancy are still not clear. Therefore, you must use an effective contraceptive during the course of this study (though not the oral contraceptive pill). If you find out that you have become pregnant while taking part in the study, you should immediately tell the principal investigator.

**What are the possible benefits of taking part?**
You would discover if you do actually suffer from PMS. If you do, you may find that SJW will help you. However, this cannot be guaranteed. You would also be helping research to move forward, which may show that SJW, a herbal medication, is effective for treating PMS. This would be of huge benefit as many women prefer to turn to herbal medications rather than take prescribed medication, especially if their symptoms are milder. Also, many products are targeted at women even though there is little evidence to show that they are effective. This study would clarify whether SJW should be used for PMS.

**What if new information becomes available?**
Sometimes during the course of a research project, new information becomes available about the drug being studied. If this happens, your research doctor will tell you about it and discuss with you whether you want to continue in the study. If you decide to withdraw, your research doctor will make arrangements for your care to continue. If you decide to continue in the study, you will be asked to sign an updated consent form. Also, on receiving new information, your research doctor might consider it to be in your best interest to withdraw you from the study. They will explain the reasons and arrange for your care to continue.

**What happens when the research study stops?**
If you would like to take SJW after the study has finished, it is available over the counter and is easily accessible.

**What if something goes wrong?**
If you are harmed by taking part in this research project, there are no special compensation arrangements. If you are harmed due to someone’s negligence, then you may have grounds for legal action, but you may have to pay for it. Regardless of this, if you wish to complain, or have any concerns about any aspect of the way you have been approached or treated during the course of this study, the normal National Health Service complaints mechanisms should be available to you.

**Will my taking part in the study be kept confidential?**
All information which is collected about you during the course of the research will be kept strictly confidential. Any information about you which leaves the hospital will have your name and address removed so that you cannot be recognized from it.

Your GP will have to be notified of your participation in the trial.
What will happen to the results of the research study?
After the study has finished, you will be told when you were given SJW tablets, when you were given placebo tablets and you will be given a summary of your results. There will also be an evening where a presentation will be given to inform you more about the study, and also to give you the opportunity to ask any questions you may have about the study, and to give you the chance to meet other women who took part.

The results of this study may be published in an academic journal and be presented at academic conferences. However, no one will be able to tell that you participated in the trial. If the journal is published, you will be notified of this, so that you can obtain a copy.

Who is organising the funding for the research?
This research is able to be funded by an extremely generous donation from a former PMS clinic patient, Rosalind Bolton. The Rosalind Bolton Bequest have funded the services of a PhD student, Sarah Canning, to carry out this research. She will not be paid for including you in the study.

Who has reviewed the study?
Leeds (West) NHS research ethics committee have reviewed the study and agreed that it can go ahead.

If you have any questions, or would like to discuss any aspect of the study, then please don’t hesitate to contact the principle researcher, Sarah Canning, using the contact details below.

Telephone: 0113 3433198
Email: s.e.canning@leeds.ac.uk
Address: Institute of Psychological Sciences
University of Leeds
Leeds
LS2 9JT

You will be given a copy of this information sheet to keep.

Thank you for your time!
Appendix 3.5: Information sheet for women who were screened by post

You are being invited to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Please take time to decide whether or not you wish to take part.

What is the purpose of the study?
This study is interested in gaining a better understanding about how PMS should be characterised. Many symptoms have been associated with PMS and this research aims to discover which are the central symptoms involved.

Why have I been chosen?
You have been chosen for this study as you have either shown an interest in the research or attend the premenstrual clinic at Leeds General Infirmary, and seem to suffer from premenstrual symptoms.

Do I have to take part?
It is up to you to decide whether or not to take part. If you do decide to take part, you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part, you are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive.

What will happen to me if I take part?
As a participant in the research, you would be required to take part for three consecutive menstrual cycles. During the time you take part in the study, you would have to fill in a two minute daily symptom checklist, which simply names 20 symptoms (e.g. aches, anger) and asks you to rate the strength that you have suffered from these each day. You would also be required to complete a questionnaire bundle at the end of each week, which would ask you about feelings of depression, anxiety, impulsivity and aggression in more detail. This would take around 20 minutes to complete.

What are the alternatives for diagnosis or treatment?
If you were still uncertain whether to take part in the study you could be seen by your family doctor and/or be referred to our PMS clinic instead.

What are the possible benefits of taking part?
You would discover if you do actually suffer from PMS. You would also be helping research to move forward, helping us to gain knowledge about how PMS should be characterised. This would be of huge benefit as PMS is a poorly understood condition.

What if something goes wrong?
If you are harmed by taking part in this research project, there are no special compensation arrangements. If you are harmed due to someone’s negligence, then you may have grounds for legal action, but you may have to pay for it. Regardless of this, if you wish to complain, or have any concerns about any aspect of the way you have been approached or treated during the course of this study, the normal National Health Service complaints mechanisms should be available to you.

Will my taking part in the study be kept confidential?
All information which is collected about you during the course of the research will be kept strictly confidential. Any information about you which leaves the University of Leeds will have your name and address removed so that you cannot be recognized from it.
What will happen to the results of the research study?
After the study has finished, you will be given a summary of your results. There will also be an evening where a presentation will be given, to inform you more about the study, and also to give you the opportunity to ask any questions you may have about the study, and to give you the chance to meet other women who took part.

The results of this study may be published in an academic journal and be presented at academic conferences. However, no one will be able to tell that you participated in the trial. If the journal is published, you will be notified of this, so that you can obtain a copy.

Who is organising the funding for the research?
This research is able to be funded by an extremely generous donation from a former PMS clinic patient, Rosalind Bolton. The Rosalind Bolton Bequest have funded the services of a PhD student, Sarah Canning, to carry out this research. She will not be paid for including you in the study.

Who has reviewed the study?
Leeds (West) NHS research ethics committee have reviewed the study and agreed that it can go ahead.

If you have any questions, or would like to discuss any aspect of the study, then please don’t hesitate to contact the principle researcher, Sarah Canning, using the contact details below.

Telephone: 0113 3433198
Email: s.e.canning@leeds.ac.uk
Address: Institute of Psychological Sciences
University of Leeds
Leeds
LS2 9JT

You will be given a copy of this information sheet to keep.

Thank you for your time!
Appendix 3.6: Consent form for women being screened by post

Name of Researcher: Sarah Canning

1. I confirm that I have read and understand the information sheet (Version 5, 20/12/2005) for the above study and have had the opportunity to ask questions.

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.

3. I understand that I do not have to answer any questions that I would prefer not to answer.

4. I consent to information collected in the study being used for future research studies on PMS, so long as those studies have received research ethics committee approval.

5. I agree to take part in the above study.

Name of Participant: __________________________ Date: ______________ Signature: __________________________
Thank you for agreeing to take part in this study. The purpose of this form is to make sure that you are happy to take part and that you know what is involved in doing so. Please read the statements below and tick each box if you agree with the statement.

1. I have read and understood the study information booklet

2. I have had the opportunity to ask questions about the study

3. I have had satisfactory answers to any questions I have asked

4. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason as to why

5. I understand that I do not have to answer any question that I would prefer not to

6. I agree to take part in the above study

Signed: ________________________________

Full name in BLOCK LETTERS: ________________________________

Date: ________________________________
Appendix 3.8: RCT Admission Form

SJW for PMS - 04/Q1205/173

Investigator Name: ____________________________

Participant Number: __________

Participant Initials: __________

Inclusion Criteria:
If any item is ticked ‘NO,’ DO NOT ADMIT patient to study.

1. Has provided written informed consent
2. Female between 18 and 45 years
3. Of Good health
4. Have regular menstrual cycles (25-35 days)

Exclusion Criteria:
If any item is ticked ‘YES,’ DO NOT ADMIT patient to study.

1. Taking prescribed medication for PMS.
2. Taking OTC medication for PMS and unwilling to abstain for the duration of the trial.
3. Using hormonal preparations, including oral contraception or Hormone Replacement Therapy (HRT)
4. Is pregnant, planning pregnancy or lactating.
5. Has known photosensitivity.
6. Taking HIV protease inhibitors (indinavir, nelfinavir, ritonavir, saquinavir); HIV non-nucleoside reverse transcriptase inhibitors (efavirenz, nevirapine); warfarin; cyclosporin; oral contraceptives; photoreactive agents; anticonvulsants (carbamazepine, phenobarbitone, phenytoin); digoxin; theophylline; triptans (sumatriptan, naratriptan, rizatriptan, zolmitriptan, frovatriptan); SSRIs (citalopram, fluoxetine, fluvoxamine, paroxetine, sertaline), psoralens, (oxsoralen, trisoralen, ciprofloxacin and tetracycline) lithium, antipsychotic drugs, neuroleptic drugs, cimetidine, oral corticosteroids, thyroid hormones, or immunosuppressive agents.
Appendix 3.9: RCT Consent Form

St John’s Wort and Premenstrual Syndrome: A research study

Name of Researcher: Sarah Canning
Participant Identification Number:

1. I confirm that I have read and understand the information sheet (Version 3, 23/6/2005) for the above study and have had the opportunity to ask questions

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.

3. I understand that I do not have to answer any questions that I would prefer not to answer.

4. I understand that sections of any of my medical notes may be looked at by responsible individuals, including the sponsors and regulatory authorities where it is relevant to my taking part in research. I give permission for these individuals to have access to my records.

5. I consent to information collected in the study being used for future research studies on PMS (for example, details of response to medication) so long as those studies have received research ethics committee approval.

6. I agree to take part in the above study.

Please tick box

Name of Participant ___________________________ Date _______________ Signature ___________________________

Name of Person taking consent ___________________________ Date _______________ Signature ___________________________

(if different from researcher)

Researcher ___________________________ Date _______________ Signature ___________________________
Appendix 3.10: Cut-offs for clinically significant anxiety

1. The 90% probability limit criteria for defining clinical significance involve the use of the following formula:

Cut-off = 1.645 sd above the mean

Therefore, the cut-off value produced when this formula was applied to Spielberger et al.’s (1983b) norms for working adults aged 19-49 years (mean 35.59, sd 9.42) was:

$$((1.645 \times 9.42) + 35.59) = 51.09$$

2. Jacobson’s criteria for defining clinical significance (Fisher and Durham, 1999; Jacobson and Truax, 1991) involve the use of the following formula:

$$\text{Cut-off} = \frac{S1M2 + S2M1}{S1 + S2}$$

$S1 = \text{Standard deviation for GAD sample at pre-treatment}$
$S2 = \text{Standard deviation for the well functioning sample}$
$M1 = \text{Pre-treatment mean of the STAIT for the GAD sample}$
$M2 = \text{Mean of the well functioning sample on the STAIT}$

Therefore the cut-off value produced when Spielberger et al.’s (1983) norms for the functional population and Fisher and Durham’s (1999) norms for the GAD sample were used was:

$$\frac{(9.45 \times 35.59) + (9.42 \times 57.00)}{(9.45 + 9.42)} = 46.28$$

1. NAME OF THE MEDICINAL PRODUCT
Jarsin 450mg

2. QUALITATIVE AND QUANTITATIVE COMPOSITION
One coated tablet contains 450 mg of Hypericum dry extract\(^1\), extractive agent methanol 80 %v/v. Drug-Extract-Ratio (DER) 3-6:1.

3. PHARMACEUTICAL FORM
Oval, yellow, coated tablet. No markings. Odour of vanilla.

4. CLINICAL PARTICULARS

4.1 Therapeutic indications
Herbal medicine for treatment of mild temporary depressive episodes.

4.2 Posology and method of administration
Unless otherwise prescribed, adults and children over 12 years of age take one tablet two times daily.

The tablets should be swallowed whole with some liquid at mealtimes.

In general, it is necessary to take the product for 4-6 weeks until the symptoms show marked improvement. If symptoms continue unchanged after 4 weeks or become stronger a medical practitioner should be consulted.

4.3 Contraindications
Not to be used at the same time as:
- cyclosporin,
- tacrolimus,
- indinavir and other protease inhibitors in anti-HIV treatment,
- irinotecan and other cytotoxic agents,
- other antidepressants.

Not to be used:
- in known hypersensitivity (allergy) to the active substance or to any of the excipients.
- in severe depressive episodes.
- in known hypersensitivity to light.

Use during pregnancy and lactation only if strictly indicated and after careful consideration of the benefits and possible risks. To date insufficient data are available with pregnant women and breastfeeding mothers.

No adequate information is available about the use of this product by children, so that it should not be used by children below 12 years old.

4.4 Special warnings and special precautions for use
The therapeutic activity of cumarin anticoagulants (e.g. phenprocoumon, warfarin), theophylline, cyclosporin, digoxin and some antidepressants may be significantly reduced.

Treatment with Jarsin at the same time as these products should thus only take place after taking medical advice.

Particularly when starting, changing the dose or stopping treatment with Jarsin the thromboplastin time or digoxin plasma level should be carefully monitored.

Avoid exposure to intense UV-light (long sunbathing, high-altitude sunshine, solaria) when taking the product.

Women using oral contraceptives should be warned of the possibility of breakthrough bleeding and be counselled regarding possible additional contraceptive measures.

Studies in patients with restricted liver function and patients with renal insufficiency requiring dialysis show that hypericin and pseudohypericin are almost completely metabolised and excreted via the liver. Decreased kidney function thus has no effect on hypericin levels. In mild liver insufficiency (Child-Pugh A) there is also no effect but there is in major liver damage (Child-Pugh B). Such patients should thus be dosed with caution to reduce any possible risk of liver toxicity.

4.5 Interaction with other medicinal products
In rare cases an interaction which could lead to reduced therapeutic activity has been observed with the following products:
- cumarin-type anticoagulants (e.g. phenprocoumon, warfarin),
- cyclosporin.
Summary of Product Characteristics

Lichtwer Pharma

- tacrolimus,
- digoxin,
- indinavir and other protease inhibitors,
- irinotecan and other cytostatics,
- amitriptyline, nortriptyline,
- midazolam,
- theophylline,
- hormonal contraceptives.

The effects of serotonergic antidepressants (e.g. Paroxetine, Sertralin, Nefazodon) may be increased. Undesired effects such as nausea, vomiting, anxiety, restlessness and confusion may be seen.

Interaction with other photosensitising medicinal products to increase the undesired effects is theoretically possible.

Medical advice should be sought whenever other medicines are taken.

4.6 Use during pregnancy and lactation
Only use during pregnancy and lactation if strictly indicated and after careful consideration of the benefits and possible risks. To date insufficient data are available with pregnant women and breast-feeding mothers.

Animal studies showed no signs of teratogenic effects. However, in isolated cases, human plasma levels have been obtained which were associated with embryo- or fetotoxic effects in animal studies.

4.7 Effects on ability to drive and use machines
No known effects.

4.8 Undesirable effects
Photosensitivity may lead to sunburn-like skin reactions, especially in fair-skinned persons exposed to strong sunlight.

Gastrointestinal complaints, allergic skin reactions (redness, itching), drowsiness or agitation may occur in rare cases.

4.9 Overdose
To date there have been no reports of acute poisoning in humans. If massive overdoses have been taken, the patient should be protected from sunlight and other sources of ultraviolet radiation for a period of 1-2 weeks (stay indoors, full clothing, use sun-protecting agents).

The side effects described above may present more strongly.

5. PHARMACOLOGICAL PROPERTIES

5.1 Pharmacodynamic properties
In preparations of mouse and rat brain the hypericum extract LI 160 inhibits synaptosomal absorption of the neurotransmitters norepinephrine and serotonin. Down-regulation of serotonin and norepinephrine beta receptors has also been demonstrated. The extract has also been tested in a number of animal behavioural models with rats and mice. These showed typical effects such as reserpine antagonism, shortening of anaesthesia duration and immobility time. On the basis of the present pharmacological data the extract is classified as an "atypical antidepressant".

5.2 Pharmacokinetic properties

Because of the complex composition of the plant extract, pharmacokinetic studies can only be conducted with mark substances in the extract; one of which is hypericin (a dianthrene derivative).

Oral administration of 300 or 900 mg of the extract LI 160 to healthy male volunteers resulted in peak mean plasma hypericin levels of 1.3 or 7.2 ng hypericin/ml plasma respectively. Absorption began after approx. 2 hours, the mean elimination half-life was 24 hours.

During 14-day long-term medication (3 x 300 mg/day) a steady-state concentration of 8.8 ng/ml was reached after 6-7 days. The volume of distribution was 162 litres with a total clearance of 0.55 ml/min.

5.3 Preclinical safety data
Preclinical data reveal no special hazard for humans based on conventional studies of single- and repeated-dose toxicity, reproduction toxicity, genotoxicity and carcinogenic potential.

Single-dose toxicity
Extract LI160 has low toxicity after single doses

<table>
<thead>
<tr>
<th>Species</th>
<th>Route</th>
<th>LD 50 mg/kg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouse</td>
<td>Oral</td>
<td>&gt; 5000</td>
</tr>
<tr>
<td>Rat</td>
<td>Oral</td>
<td>&gt; 5000</td>
</tr>
<tr>
<td>Mouse</td>
<td>i.p.</td>
<td>1780</td>
</tr>
<tr>
<td>Rat</td>
<td>i.p.</td>
<td>1000</td>
</tr>
</tbody>
</table>

Repeated-dose toxicity
No substance-related changes seen in a 26-week study in dogs.

Reproductive toxicity
No evidence of teratogenicity in rats or rabbits in studies with doses which were toxic to the dams. Embryonic findings (retarded growth, morphological changes) were associ-
ated with maternal plasma levels of hypericin of more than 10 ng/ml.

No effect on fertility in a study in rats.

In the rat hypericin accumulates in the milk and can exceed plasma concentrations by several times.

**Mutagenicity**

No evidence of mutagenicity in in vivo or in vitro studies.

Long-term animal studies for tumorigenic potential gave no indication of substance-related changes.

**Phototoxicity**

It is known that hypericum herb can have a marked phototoxic effect when eaten in large amounts by grazing animals; the phototoxic dose in cattle is about thirty times higher than the therapeutic dose in man.

Single administration of 900, 1800 or 3600 mg extract to healthy male volunteers (n=13) produced no significant change in sensitivity to UV light. In contrast, long-term administration of 1800 mg/day/15 days (equivalent to about 5.4 mg hypericin/day) (n=50, both sexes) led to 20% decreased pigmentation dose and increased sensitivity to UV-A light. Light sensitive subjects showed significantly more erythema after 15 days treatment than after the first treatment.

6. **PHARMACEUTICAL PARTICULARS**

6.1 **List of excipients**

Cellulose, soy polysaccharides, long-chain partial glycerides, lactose monohydrate, colloidal anhydrous silica, hypromellose, stearic acid, titanium dioxide (E 171), yellow ferric oxide (E 172)

6.2 **Major incompatibilities**

None known.

6.3 **Shelf life**

18 months.

6.4 **Special precautions for storage**

Do not store above 25°C. Store in the original package. Protect from moisture.

6.5 **Nature and contents of container**

PVC/PVDC/aluminium foil blisters in cardboard cartons. Packs of 25, 60 and 100 tablets. Hospital packs of 1000 tablets. Not all pack sizes may be marketed.

6.6 **Special precautions for disposal**

No special requirements.

7. **MARKETING AUTHORITY HOLDER**

(in Germany)

Lichtwer Pharma GmbH
Wallenroder Strasse 8-10
13435 Berlin, Germany

Tel.: +49 (0) 30 - 40370 - 0
Fax: +49 (0) 30 - 40370 - 103
Internet: http://www.lichtwer.de

8. **MARKETING AUTHORITY NUMBER**

39112.00.01 (in Germany)

9. **DATE OF FIRST AUTHORISATION**

28 May 2003 (in Germany)

10. **DATE OF REVISION OF TEXT**

May 2004
Appendix 8.2: RCT Recruitment Questionnaire

If you would be interested in taking part in the study, then please complete the following questionnaire, which will ask you about various aspects of your life, including your general and gynaecological health. Some of the questions may be of a sensitive nature, but remember, the answers you give will be completely confidential.

**Personal Details**

1. Height: _______ cm
2. Weight: _______ Kg
3. Are you in a relationship: No _____ Yes _____
4. If so, how long have you been in this relationship? _____ Years _____ Months
5. Do you have any children? No _____
   Yes _____ Number _____
6. What is your occupation?
   - Employed _____
   - Unemployed _____
   - Retired _____
   - Housewife _____
   - Student _____
   - Other, please specify ___________________________
7. What is/ was your most recent full-time job? ___________________________

**Your General Health**

8. Do you suffer from any medical condition (allergies, heart condition, diabetes etc)?

_____________________________________________________

9. Are you currently taking any medication? No _____ Yes _____
10. If so, what are you taking? ___________________________

_____________________________________________________

11. What is this for? ___________________________

_____________________________________________________

**Your Gynaecological Health**

12. How long does your period (days of bleeding) usually last? _____ days
13. What is the length of your menstrual cycle? (i.e. From the 1st day of one period to the first day of the next period e.g. 26 days) _____ days
14. What was the date of your last period? (The 1st day of bleeding) dd/mm/yyyy

____________________

15. How regular are your periods (Please circle one of the following):
   - Very regular (always starting within 1-3 days of expected date)
   - Quite regular (starting within 4-7 days of expected date)
   - Somewhat irregular (some variation in cycle length)
   - Very irregular (lots of variation in cycle length)
16. What method of contraception do you use? (Please circle one of the following):
• None  
• Pill  
• Condom  
• IUD/ coil  
• Sterilised  
• Past menopause  
• Hysterectomy  
• Other, please specify ________________________________

17. How long have you been using this method of contraception? ______________

18. Have you ever used the pill?
   • No _____
   • Yes, _____ months/ years ago
   • Yes, I am taking it currently _____

19. Are you, or have you ever taken hormone replacement therapy (HRT)?
   • No _____
   • Yes _____ months/ years ago
   • Yes, currently _____

20. Have you ever been pregnant? No _____ If no, go to question 25
   Yes _____

21. Are you currently pregnant? Yes _____ No_____

22. Are you currently breast feeding? Yes _____ No_____

23. If you have been pregnant: Number of births ______
   Number of miscarriages ______
   Number of terminations ______

24. Did you suffer from depression after the birth of any of your children?
   • No _____
   • Yes, but didn’t receive treatment _____
   • Yes, and received treatment _____
   • Don’t know _____

25. Do you, or have you ever suffered from depression (apart from at childbirth)?
   • Don’t know _____
   • No _____
• Yes, currently _____
  • Yes ______ months/ years ago

26a. What, if any, treatment did you receive/ are you receiving? (E.g. none, prescribed medication, psychotherapy etc) ________________________________

26b. When did you receive this treatment? (E.g. Currently, 2 years ago etc)

______________________________

27. Do you suffer from any other psychiatric conditions? (Anxiety, schizophrenia etc)
  • No _____ (If no, please go to question 29)
  • Yes, please specify ________________________________

If yes,
28a.) Are you / have you previously seen a psychiatrist for this?
  • No _____
  • Yes, _____ months/ years ago
  • Yes, I am currently _____

28b.) Are you / have you received medication for this?
  • No _____
  • Yes, _____ months/ years ago. Please specify ________________________________
  • Yes, I currently am _______. Please specify ________________________________

29. Do any of your family members suffer from any psychiatric conditions?
  • No _____
  • Yes, please specify ________________________________

30. Do you believe you suffer from premenstrual syndrome (PMS)?
  • No ______ (If no, please go to question 34a)
  • Yes, definitely ______
  • Yes, possibly ______
  • Don't know ______

If yes:
31a) How long do you think you have suffered from PMS for? ____ months/ years

31b) Do your PMS symptoms tend to ease after the start of your period? Yes _____  No _____

31c) Is there at least one week after the end of your period (after bleeding has stopped) where you feel totally free from these symptoms?  Yes _____  No _____

31d) Do you take any medication for PMS?  Yes _____  No _____
32. If yes, what do you take? __________________________

33. Would you be willing to discontinue this medication for the duration of a study?
   Yes _____ No _____

34. Would you be happy to be contacted regarding other studies that we are conducting?
   Yes _____ No _____

If you would like to leave your contact details below then you will be contacted soon with more
information about the clinical trial.

Name: __________________________

Address: ____________________________________________________________________________

Telephone Number: ____________

Date of birth: ____________

Email: __________________________

If you have any questions or would like to discuss anything in more detail, please contact Sarah
Canning (Telephone: 0113 3433198   Email: s.e.canning@leeds.ac.uk)

If you are not suitable to take part in this study, would you like us to keep your details and
inform you of future studies that you may take part in? Yes _____ No _____
Appendix 8.3 Interview Guide for Safety Assessment

Questions:

1. How are you finding doing the study – any problems?
2. Have you been able to consume the tablets ok?
3. Are there any points you feel may need further clarification?
4. Have you felt as well as usual? (query any side-effects)

Issues to be addressed:

1. Future dates for testing
2. Reiteration of what participant is required to do until next meeting
3. Administration of supplement products if necessary
Appendix 8.4 Example symptom profile for women screened by post

What is Premenstrual Syndrome (PMS)?

PMS is now recognised as a medical condition, which affects women of childbearing age. It is a group of symptoms that occurs any time from a couple of days to a couple of weeks before a period. The symptoms disappear abruptly within the first few days of the bleed. PMS is sometimes called premenstrual tension (PMT), but increased tension may not be the only symptom.

What are the Symptoms?

Although a large range of symptoms have been associated with PMS, the more common symptoms are listed below:

Feelings:
- Tension
- Low mood
- Aggression

Irritability
Weepiness
Tiredness

Feeling of anger
Lack of concentration
Anxiety / panic

Physical symptoms:
- Breast tenderness
- Headaches

Bloating
Clumsiness

Weight gain
Back ache

People suffering from PMS do not always experience all of these symptoms. Also, people may experience some of these symptoms and not suffer from PMS. It is the timing of these symptoms that distinguishes PMS from other conditions.

How is PMS Caused?

Many theories of what causes PMS have been proposed. People used to think that hormone imbalances were responsible for the symptoms of PMS. However, it now seems clear that this is not the case and that women with PMS are simply more ‘sensitive’ to the normal hormonal changes that occur each month. PMS can become worse after childbirth, a miscarriage or termination of pregnancy. Poor diet and stress can affect PMS and it tends to get worse as you get older.

How do I know if I have PMS?

Diagnosis involves 2-3 months of daily monitoring of your symptoms using a PMS symptom checklist. The scores from each day are totalled and divided into separate types of symptoms to show which symptoms a person suffers from most. Postmenstrual (after the period) and
premenstrual (before the period) averages are taken and compared. There needs to be a symptom increase of at least 30% premenstrually for two out of three menstrual cycles to confirm PMS.

You completed the Daily Symptom Report (DSR) every day for three menstrual cycles. We summed up your ratings from cycle days 5-10 to determine your postmenstrual score (day 1 being the first day of each bleed) and the ratings of the 6 days before the next period for the premenstrual score. We then calculated the percentage increase in your symptoms premenstrually.

We found that in all three menstrual cycles you had a premenstrual increase in symptoms of at least 30%. Your symptoms that increased the most consistently premenstrually were swelling, cramps, aches, poor coordination, irritability, feeling anxious and out of control.

Some other conditions can worsen premenstrually, making you think you might have PMS when you do not. Two common symptoms that worsen premenstrually are feeling low and feeling anxious. Therefore, we also asked you to complete a questionnaire at the end of each week, which we used to assess these feelings.

**Feeling Low**

You were asked to complete the Beck Depression Inventory at the end of each week. The score for each item was added together to produce a total score. Scores of 0-12 indicate no depression. If the score is above 12, this may mean that a person feels low more than other people.

As feeling low is a major symptom of PMS, we looked at this score in the week after the period to make sure that your mood was not being affected by being premenstrual.

We took your answers from this questionnaire on the three occasions that you completed it when you were postmenstrual. We then averaged these scores. Your average postmenstrual score was in the normal range.

**Anxiety Levels**

It is also important to look at levels of anxiety to determine whether you do have PMS. You were asked to complete the State-Trait Anxiety Inventory. Trait anxiety refers to how you react generally, while state anxiety refers to how you react to particular situations e.g. when you are premenstrual.

You were asked to rate 40 statements on a scale of 0-4 to assess your state and trait anxiety levels. The questions that asked you how you generally feel were used to assess your trait anxiety levels, while those asking you about how you feel at that particular moment were used to look at your state anxiety levels.
We looked at your postmenstrual trait anxiety scores to gain the best estimate of your trait anxiety levels, not being influenced by being premenstrual. A score of around 35-40 is considered normal. While scores of above 51 may mean that you have higher anxiety levels than normal. Your average postmenstrual score was in the normal range.

**Summary**

Your diaries indicated that you may be suffering from PMS, as you showed a premenstrual increase of at least 30% in the feelings and behaviours that were measured in all three menstrual cycles and your scores on the anxiety and depression measures were in the normal range. The following graphs show you how your feelings and behaviours changed across each of the three menstrual cycles that you took part in the study. The mood symptoms that we measured included irritability, mood swings, nervous tension, anxiety, depression and feeling out of control. The behavioural symptoms included poor coordination, confusion, insomnia, crying and fatigue. The pain symptoms included breast tenderness, cramps and aches. Finally the physical symptoms included food cravings and swelling.
Your symptoms during the third menstrual cycle

You may not find these symptoms to be a problem. However, this may mean that you feel monthly changes in your feelings and behaviour a little more than other people. If you feel that this is something that troubles you, it may be worth seeing your GP. If you would like to do this, we could send him/her a copy of your profile of scores to help them decide how best to help you.

If you live in the Yorkshire area, then you can ask your GP to refer you to the Rosalind Bolton PMS clinic, Leeds General Infirmary. This clinic provides personalised help, advice and treatment to women suffering from PMS and runs on Tuesday mornings.

More Information
More information about PMS is available from the National Association of Premenstrual Syndrome:
- Tel: 0870 7772178
- Website: www.pms.org.uk
- Address: 41 Old Road, East Peckham, Kent, TN12 5AP

Thank you for your interest in our research.

If you have any questions about the study, please contact Sarah Canning at Leeds University.

Telephone: (0113) 3433198
E-mail: s.e.canning@leeds.ac.uk
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Given value</th>
<th>Actual value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average weight</td>
<td>717,8 - 762,2 mg</td>
<td>735,9 mg</td>
</tr>
<tr>
<td>Weight min.</td>
<td>703,0 - 777,0 mg</td>
<td>727,0 mg</td>
</tr>
<tr>
<td>Weight max.</td>
<td>703,0 - 777,0 mg</td>
<td>746,0 mg</td>
</tr>
<tr>
<td>Height</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height min.</td>
<td>6,90 - 7,30 mm</td>
<td>6,42 mm *</td>
</tr>
<tr>
<td>Height max.</td>
<td>6,90 - 7,30 mm</td>
<td>6,50 mm *</td>
</tr>
<tr>
<td>Length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length min.</td>
<td>9,00 - 9,40 mm</td>
<td>9,25 mm</td>
</tr>
<tr>
<td>Length max.</td>
<td>9,00 - 9,40 mm</td>
<td>9,30 mm</td>
</tr>
<tr>
<td>Diameter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diameter min.</td>
<td>17,90 - 18,30 mm</td>
<td>18,31 mm</td>
</tr>
<tr>
<td>Diameter max.</td>
<td>17,90 - 18,30 mm</td>
<td>18,39 mm</td>
</tr>
<tr>
<td>Fracture resistance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min.</td>
<td>120 - 300 N</td>
<td>201 N</td>
</tr>
<tr>
<td>Max.</td>
<td>120 - 300 N</td>
<td>237 N</td>
</tr>
<tr>
<td>Disintegration time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max.</td>
<td>60,0 min.</td>
<td>19 min.</td>
</tr>
<tr>
<td>Loss on drying</td>
<td>0,00 - 5,00 %</td>
<td>4,29 %</td>
</tr>
<tr>
<td>Identity (DC)</td>
<td>St. John's wort dry extract</td>
<td>passed</td>
</tr>
<tr>
<td>Total Hypericin</td>
<td>720 - 1260,00 µg/DF</td>
<td>799,39 µg/DF</td>
</tr>
<tr>
<td>Hyperforin</td>
<td>0,00 - 27,00 mg/DF</td>
<td>15,2 mg/DF</td>
</tr>
<tr>
<td>Recovery extract (Hyperosid)</td>
<td>427,5 - 472,5 mg</td>
<td>440,7 mg</td>
</tr>
</tbody>
</table>

* Abnahme- und akzeptiert.

Tested: Date 13.12.05  
Signature: [Signature]

Head of Quality Control: yes ☐  no ☐  
(Qualified Person) Date 13.12.05  
Signature: [Signature]
Dietary supplements and herbal remedies for premenstrual syndrome (PMS): a systematic research review of the evidence for their efficacy

SARAH CANNING, MITCH WATERMAN & LOUISE DYE
Institute of Psychological Sciences, University of Leeds, UK

Abstract Many women with PMS use alternative therapies, although there has been little research to demonstrate their efficacy. This systematic review provides a comprehensive discussion of dietary supplements and herbal remedies commonly used for premenstrual syndrome (PMS), including calcium, magnesium, vitamin B6, evening primrose oil, Vitex agnus castus, ginkgo biloba and St John's Wort. Randomized controlled trials of magnesium and evening primrose oil have produced conflicting results, in contrast to the substantial evidence for the efficacy of calcium and vitamin B6. There are insufficient data to advocate the use of ginkgo biloba, Vitex agnus castus and St John's Wort, although preliminary data seem supportive. Greater standardization of PMS diagnosis and assessment, with randomized, double-blind, placebo-controlled trials using larger, representative samples, strict, prospectively confirmed diagnostic criteria and assessment of treatment efficacy, would help to clarify the role of these alternative PMS treatments. Although much of the clinical research is preliminary and/or inadequately controlled, this review will be relevant to the practicing clinician looking for greater understanding of the alternative therapies available to their patients with PMS.

Introduction
This paper presents a systematic research review of various dietary and herbal interventions used by women suffering from premenstrual syndrome (PMS). PMS is a cyclical condition occurring 7–10 days before the onset of menstruation, and is relieved shortly after menstrual flow begins (Reid, 1991). The most commonly reported symptoms include depression, anxiety, irritability and mood swings, as well as physical symptoms such as breast tenderness and bloating. PMS involves milder symptoms than Premenstrual Dysphoric Disorder (PMDD) (Ismail & O'Brien, 2001), although PMS symptoms are severe enough to disrupt women's lives (Shaw et al., 2003). PMDD...
characterizes a subgroup of women whose symptoms are particularly severe and is confirmed using DSM-IV-TR criteria.

**PMS assessment**

The diagnosis of PMS is based on the pattern of symptoms over the menstrual cycle. Both retrospective and prospective measures are used for this purpose. Retrospective questionnaires promote incorrect symptom timing, bias due to cultural expectations and heavy reliance upon memory of symptoms (Connolly, 2001), resulting in an inflated estimate of symptom severity (De Souza et al., 2000). Prospective daily self-report instruments are the only accepted means of confirming PMS diagnosis (Steiner & Wilkins, 1996). These are easy to administer, less reliant on memory (Haywood et al., 2002) and highlight intercycle variability in symptom type and severity (Connolly, 2001). However, they are demanding to complete and may bias symptom patterns though nonadherence (Connolly, 2001).

Various measures have been used to prospectively confirm PMS diagnosis. Budeiri et al. (1994) identified over 65 instruments, but concluded that there is not yet agreement about which is the most appropriate. At least five different daily diaries are widely used but there is no consensus as to which one is best (Steiner et al., 2003).

The diagnosis of PMS is also based on symptom severity through the application of severity criteria to symptoms, such that an increase of at least 30% from the follicular to luteal phase scores is required to confirm PMS (NIMH, 1983).

Many women with PMS use alternative therapies (Girman et al., 2003), even though the efficacies of these are not established (Domoney et al., 2003). This review will evaluate evidence for the effectiveness of dietary and herbal interventions for PMS assessed by randomized, double-blind, placebo-controlled trials, in women whose PMS was diagnosed using validated measures. The studies considered in this review did not distinguish between PMS and PMDD.

**Methods**

**Selection criteria**

For review inclusion, studies were required to include a placebo/comparison group, as the placebo effect has been shown to be large in women with PMS (Freeman & Rickels, 1999). They were required to test the efficacy of one treatment, taken for at least one cycle, taken throughout or premenstrually. Participants had to be randomized to treatments in the case of parallel designs, or have treatment orders counterbalanced in the case of crossover designs.

Studies required participants of reproductive age, with PMS or PMDD, diagnosed prospectively or retrospectively. Few trials employed prospective diagnosis or assessment of efficacy hence, in order to provide a comprehensive review, retrospective measures were included. Women had to have no other pre-existing psychiatric conditions (e.g. depression, anxiety), although studies including women presenting with depression or anxiety only premenstrually were included.

Studies employing outcome measures which examined combined PMS symptoms, global scores or specific symptoms, e.g. cyclical breast pain, were included. Trials including women taking oral contraceptives were also included.
Search strategy

Electronic databases

Trials were identified by searching Embase (1980-2006), Medline (1966-2006), AMED (1985-2006), Cinahl (1982-2006), PsycINFO (1967-2006), and the Cochrane Controlled Trials Register database.

A general search on these databases revealed dietary supplements and herbal remedies used for PMS, including vitamin B6, magnesium, calcium, *Vitex agnus castus*, evening primrose oil, St John’s Wort and ginkgo biloba supplements. Databases were searched using all Latin and English names for these supplements. Hence, the following keywords were used:

- Pms, pmt, pmdd, llpd, llpdd, premenstrual syndrome, pre menstrual syndrome, pre-menstrual syndrome, premenstrual syndrome, premenstrual tension, pre menstrual tension, pre-menstrual tension, premenstrual dysphoria, pre-menstrual dysphoria, pre menstrual dysphoria, premenstrual dysphoric disorder, pre-menstrual dysphoric disorder, pre menstrual dysphoric disorder, late luteal
- Vitex, agnus castus, vitex agnus castus, vitex agnus-castus, chaste tree, chasteberry, chaste-berry, monk’s pepper, monks pepper, hemp tree, agnolyt, agnufemil, castufemin, cefanorm, femicur, gynocastus, hewekliman, kytaienti, strotan, agnomens
- Evening primrose oil, oenothera biennis, evening primrose, primrose oil, oenothera, biennis, fever plant, oep, sundrop, essential fatty acids, efamol
- Calcium, calcium supplements, calcium therapy
- Magnesium, magnesium therapy
- Vitamin B6, vitamin B6, vitamin b-6, vitamin B-6, vitamin therapy, vitamins, pyridoxine, B-vitamins, B vitamins, pyridoxine hydrochloride
- St john’s wort, st johns wort, hypericum perforatum, hypericum, perforatum, hypericin, hypericins, kira
- Gingko, gingko biloba, ginko, ginko biloba, biloba, living fossil, Japanese Silver Apricot, Kew Tree, Maidenhair Tree, Yinsing
- Alternative therapy, alternative therapies, herbal therapy, nutritional supplements
- Clinical trial, trial, randomized controlled trial, controlled trial, randomized, randomized controlled trial, blind, double blind, double-blind, doubleblind, crossover, cross-over, parallel, prospective, retrospective

There were 113 articles remaining when duplicates were removed, with 32 articles kept as trials relevant to the research area. Seven articles did not meet the selection criteria, five of which had no placebo or comparison group (Berger et al., 2000; Brush et al., 1988; Larsson et al., 1989; Loch et al., 2000; Stevinson & Ernst, 2000). Two articles were excluded as they tested the efficacy of more than one treatment. Krutan Berman et al. (1990) tested the efficacy of pyridoxine for PMS, in combination with a dietary intervention, while Callender et al. (1988) tested evening primrose oil using efamol tablets which also included efavit (containing vitamin C, pyridoxine, niacin, and zinc sulfate). Hence, 25 studies meeting the selection criteria were retained (see Table 1). One additional article (Ockerman et al., 1986) was identified from the reference lists from review articles in this area.
Data synthesis for the dietary supplements and herbal remedies for PMS

Two studies of calcium, four of magnesium, 10 of B6, four of evening primrose oil, four of *Agnus castus*, one of St. John’s Wort and one of gingko biloba met the inclusion criteria. Fifteen trials suggested some benefit of the treatment under investigation, while 11 trials found no such benefit.

Table 1 describes and evaluates these studies. The trials are ordered alphabetically and sub-divided into studies finding positive and negative effects for each treatment. Aspects of the methodological quality (e.g. sample size, design, dose and duration of intervention, screening and assessment tools employed) are considered in order to provide a context for discussion of the reliability of the results. The table assumes that studies excluded women taking the oral contraceptive pill, unless otherwise stated.

Two well-designed trials rigorously assessed the efficacy of calcium supplementation for PMS (Thys-Jacobs et al., 1989, 1998). The similarities between their findings suggest that calcium supplementation of at least three cycles may be of benefit to women suffering from PMS.

Positive evidence was also found for vitamin B6. A systematic review to evaluate the efficacy of vitamin B6 for the treatment of PMS identified 25 published trials and included nine trials, representing 940 patients (Wyatt et al., 1999). The overall odds ratio for efficacy of B6 was 2.32 (1.95-2.54). Wyatt *et al.* suggest that doses of vitamin B6 up to 100 mg/day may benefit premenstrual depression and other symptoms. However, conclusions from the meta-analysis were limited by the methodological weaknesses of some of the trials included. Findings from studies assessing vitamin B6 in this review are also mixed, although studies which demonstrated benefits appeared to be methodologically stronger. The evidence suggests that continuous vitamin B6 treatment at doses of 50 to 150 mg/day may be beneficial for some PMS symptoms, since the intermittent treatment (at 50 to 300 mg/day) did not prove effective (Diegoli *et al.*, 1998; Mal mgren *et al.*, 1987; Stokes & Mendels., 1972). However, more trials, using stricter diagnostic criteria, are required to confirm its benefit.

Trials looking at magnesium supplementation produced mixed findings and many methodological limitations were apparent. Although most trials evaluating *Agnus castus* treatment reported positive effects, many studies also suffered methodological problems. Though some studies suggest evening primrose oil may benefit PMS symptoms (Ockerman *et al.*, 1986; Puolakka *et al.*, 1985), currently the evidence is not convincing, especially as the most methodologically sound study (Collins *et al.*, 1993) found no benefit for mood or physical symptoms. Therefore, trials with longer treatment durations, tighter controls and larger samples are required to evaluate *Agnus castus*, evening primrose oil and mg supplementation in PMS.

Only single studies have been performed to test the efficacy of St John’s Wort and gingko biloba, so further investigation of these therapies is needed. Hicks *et al.* (2004) suggested that their study had insufficient statistical power to demonstrate the efficacy of the dose of St John’s Wort that they used (600 mg, standardized to 0.3% of hypericin/day). Most studies have used a dose of 900 mg/day standardized to 0.3% hypericin for depression (Szegedi *et al.*, 2005), and this is the dose recommended for depression by most manufacturers of St John’s Wort. There is currently no recommended dose for PMS. Tamborini and Taurelle (1993) found gingko biloba to benefit congestive PMS symptoms, particularly breast tenderness. However, the sample of PMS sufferers studied was atypical in that women were required to report...
<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample size</th>
<th>Study design</th>
<th>Dose and duration</th>
<th>Diagnosis</th>
<th>Assessment measures</th>
<th>Outcome</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Calcium (Ca)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thys-Jacobs <em>et al.</em> (1989)</td>
<td>N=60, 33 completed; 45% dropout</td>
<td>R, DB</td>
<td>2 tablets providing 1000 mg elemental Ca/d; 3 cycles</td>
<td>14 symptoms rated daily on a 4-pt scale; 1 cycle</td>
<td>Daily symptom change; global assessment measured retrospectively</td>
<td>✓ negative affect, water retention, pain; retrospective assessment</td>
<td>Well designed trial. No washout. Prospective diagnosis but for only 1 cycle.</td>
</tr>
<tr>
<td>Thys-Jacobs <em>et al.</em> (1998)</td>
<td>N=479, 441 completed; 8% dropout</td>
<td>R, DB PGs, PC, MC</td>
<td>2 x 2, 750 mg CaCO3 tablets/d (1200 mg elemental Ca/d); 3 cycles</td>
<td>17 symptoms (4 factors: negative affect, water retention, food and pain) rated daily on a 4-pt scale; 2 cycles</td>
<td>Symptom complex score during luteal and menstrual phases; factor scores; rescue medication used</td>
<td>✓ symptom complex scores; all symptom factors</td>
<td>Well designed trial. Placebo treatment not specified. Strict diagnostic criteria. OC users included. Although no difference between users non-users found.</td>
</tr>
<tr>
<td><strong>Magnesium (mg)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facchinetti <em>et al.</em> (1991)</td>
<td>N=32, 28 completed; 13% dropout</td>
<td>R, DB PGs</td>
<td>3 tablets providing 360 mg mg ion/d from cycle day 15; 2 cycles</td>
<td>DSM III-R, MDQ daily during the 2nd and 4th cycles</td>
<td>✓ total MDQ score, especially negative affect and arousal</td>
<td></td>
<td>Small sample.</td>
</tr>
<tr>
<td>Walker <em>et al.</em> (1998)</td>
<td>N=54, 24 completed; 56% dropout</td>
<td>R, DB</td>
<td>200 mg of mg as mgO; 2 cycles</td>
<td>22 symptoms daily on a 4-pt scale, covering 6 categories (anxiety, craving, depression, hydration, other and total)</td>
<td>✓ hydration symptoms only (weight gain, swelling, breast tenderness and abdominal bloating)</td>
<td></td>
<td>No washout. Retrospective diagnosis. Low dose. OC users included. No baseline diary measurements taken.</td>
</tr>
<tr>
<td>Reference</td>
<td>Sample size</td>
<td>Study design</td>
<td>Dose and duration</td>
<td>Diagnosis</td>
<td>Assessment measures</td>
<td>Outcome</td>
<td>Comments</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>------------------------------------------</td>
<td>--------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>----------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>De Souza et al. (2000)</td>
<td>N=44, 37 completed; 16% dropout</td>
<td>R, DB</td>
<td>Each daily for 1 cycle: 1) 200 mg mg, as mgO 2) 50 mg B6 3) 200 mg mg+50 mg B6 4) placebo</td>
<td>MHQ; 1 mth menstrual diary</td>
<td>30 symptoms daily on a 5-pt scale, covering 6 categories (anxiety, craving, depression, hydration, other and total)</td>
<td>X except combined treatment, which reduced mild anxiety related symptoms</td>
<td>OC users included (18%). Diagnosis methods not specified. Treatment for only 1mth. No washout.</td>
</tr>
<tr>
<td>Walker et al. (2002)</td>
<td>N=85</td>
<td>R, DB, CO, PC, sorbitol</td>
<td>2 of the following for 2 cycles each, with 1 mth washout: placebo, mg (200, 350 or 500 mg/d, containing 1050, 813 and 717 mg of sorbitol)</td>
<td>MHQ</td>
<td>20 symptoms daily on a 5-pt scale, covering 5 categories (anxiety, craving, depression, hydration, total)</td>
<td>X Sorbitol significantly better than mg for anxiety and total PMS scores</td>
<td>OC users included (28%).</td>
</tr>
<tr>
<td>Vitamin B6 (B6) Abraham &amp; Hargrove (1980)</td>
<td>N=25</td>
<td>R, DB</td>
<td>1 × B6 tablet/d containing 500 mg of pyridoxine HCl; 3 cycles</td>
<td>MSQ (Abraham, 1980)</td>
<td>19 symptoms daily on a 4-pt scale</td>
<td>✓ total symptom score; premenstrual weight gain</td>
<td>Only 3 women with premenstrual weight gain included. Completer rate inconsistent. Unclear when symptoms were rated for diagnosis.</td>
</tr>
<tr>
<td>Barr (1984)</td>
<td>N=48, 36 analysed; 25% dropout assumed</td>
<td>R, DB</td>
<td>1 × 100 mg pyridoxine HCl/d from cycle days 10–3; 2 cycles</td>
<td>Not specified</td>
<td>8 symptoms daily for 2 wks prior to menstruation</td>
<td>✓ overall scores</td>
<td>OC users included. Entry criteria not specified. Symptoms only recorded with positive/negative responses.</td>
</tr>
<tr>
<td>Reference</td>
<td>Sample size</td>
<td>Study design</td>
<td>Dose and duration</td>
<td>Diagnosis</td>
<td>Assessment measures</td>
<td>Outcome</td>
<td>Comments</td>
</tr>
<tr>
<td>----------------------</td>
<td>----------------------</td>
<td>--------------</td>
<td>-------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Doll et al. (1989)</td>
<td>N=68, 32 completed; 53% dropout; moderate to severe PMS</td>
<td>R, DB CO, PC</td>
<td>50 mg of pyridoxine/d; 3 cycles</td>
<td>9 symptoms (emotional, somatic and menstrual) rated on a 4-pt scale; 1 cycle</td>
<td>9 symptoms on a 4-pt scale throughout treatment</td>
<td>✔ emotional type symptoms only—depression, irritability, tiredness</td>
<td>OC users included. No washout. It was not specified how the scale was used for diagnosis or treatment assessment.</td>
</tr>
<tr>
<td>Kendall &amp; Schnurr (1987)</td>
<td>N=74, 55 completed; (B6, 29; Pl, 26); 26% dropout</td>
<td>R, DB PGs, PC</td>
<td>3 x 50 mg B6 tablets/d; 2 cycles</td>
<td>PAF (Halbreich et al., 1982)</td>
<td>MDQ, every other day</td>
<td>✔ autonomic reactions (dizziness, vomiting); behaviour changes (poor performance, decreased social activities) only</td>
<td>Baseline MDQ not used for exclusion. Half the women in each condition received capsules and half tablets.</td>
</tr>
<tr>
<td>Williams et al. (1985)</td>
<td>N=617, 434 completed; 30% dropout</td>
<td>R, B PGs, PC</td>
<td>100 mg pyridoxine tablet/d; 3 cycles. The dose could be doubled/halved by the patients.</td>
<td>Through the general practitioner, if menstruation relieved 1 or more symptom</td>
<td>Symptom rating during the last week of each cycle</td>
<td>✔ for final assessment compared to entry but not for individual symptoms</td>
<td>Retrospective diagnosis and treatment assessment. ‘Other medication’ could be taken. Many women changed their dose.</td>
</tr>
<tr>
<td>Diegoli et al. (1998)</td>
<td>N=120 (30 in each group)</td>
<td>R, DB CO, PC</td>
<td>One for 3 cycles: 1) 300 mg pyridoxine/d from day 15 2) $3 \times 0.25$ mg of alprazolam/d from day 15 3) 10 mg fluoxetine/d 4) 20 mg propranolol/d, 40 mg when menstruating</td>
<td>24 symptoms rated retrospectively on a 4-pt scale</td>
<td>24 symptoms on a 4-pt scale at the end of each cycle</td>
<td>✔ only for the pyridoxine group</td>
<td>Doses and frequency of treatment differed between groups. Retrospective diagnosis and treatment assessment. It was unclear when treatment effect was assessed.</td>
</tr>
<tr>
<td>Reference</td>
<td>Sample size</td>
<td>Study design</td>
<td>Dose and duration</td>
<td>Diagnosis</td>
<td>Assessment measures</td>
<td>Outcome</td>
<td>Comments</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-------------------</td>
<td>------------------------------</td>
<td>----------------------------------------------------------</td>
<td>---------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Hagen <em>et al.</em></td>
<td>N=42, 34 completed; 19% dropout</td>
<td>R, DB CO, PC</td>
<td>2 × 50 mg pyridoxine tablets/d; 2 cycles</td>
<td>Interview conducted by the research team</td>
<td>VAS at baseline and after each treatment; ranking of 6 symptoms, adding others</td>
<td>x</td>
<td>Treatment order not fully counter-balanced. When rating, women were allowed to see their previous ratings, and it was unclear when this was performed.</td>
</tr>
<tr>
<td>Malmgren <em>et al.</em></td>
<td>N=19</td>
<td>DB, CO PC</td>
<td>300 mg pyridoxine/d from cycle day 15; 1 cycle</td>
<td>MDQ; STAI (Spielberger <em>et al.</em>, 1970) on cycle days 5-7 and 25-27</td>
<td>Not specified</td>
<td>x</td>
<td>No washout. B6 and placebo were only given for one cycle each.</td>
</tr>
<tr>
<td>Smallwood <em>et al.</em></td>
<td>N=42, severe cyclical mastalgia</td>
<td>R, DB CO, PC</td>
<td>200 mg B6 (benadon)/d; 2 cycles</td>
<td>Not specified</td>
<td>Monthly, by a clinician within the last 5 days of the cycle; daily VAS for breast pain and tenderness; paracetamol requirements</td>
<td>x for all measures</td>
<td>Only women having severe pain premenstrually for at least 6 consecutive mths included. Exclusion of OC users not stated. Method of diagnosis not specified. Only a small proportion of symptoms assessed.</td>
</tr>
<tr>
<td>Reference</td>
<td>Sample size</td>
<td>Study design</td>
<td>Dose and duration</td>
<td>Diagnosis</td>
<td>Assessment measures</td>
<td>Outcome</td>
<td>Comments</td>
</tr>
<tr>
<td>--------------------</td>
<td>----------------------------------</td>
<td>--------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>-------------------------</td>
<td>---------------------</td>
<td>---------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Stokes &amp; Mendels (1972)</td>
<td>N=13, premenstrual tension, depression</td>
<td>R, DB CO, PC</td>
<td>50 mg B6 or placebo/d for 18 days/mth, for 8–12 mths, with the order of placebo and B6 months being random</td>
<td>Not specified</td>
<td>MDQ</td>
<td>✓</td>
<td>Method of diagnosis not mentioned. Tiny sample size. It was unclear for how long each treatment was taken. Treatment assessment method not explicit.</td>
</tr>
<tr>
<td><em>Evening Primrose Oil (EPO)</em> Ockerman <em>et al.</em> (1986)</td>
<td>N=36, severe PMS</td>
<td>DB, PGs PC, olive oil</td>
<td>8 x 0.5g Efamol capsules/d; 3 cycles</td>
<td>Not specified</td>
<td>Not specified</td>
<td>✓</td>
<td>Treatment resistant sample. Extremely brief report failing to specify diagnosis methods, symptoms measured and measures used.</td>
</tr>
<tr>
<td>Puolakka <em>et al.</em> (1985)</td>
<td>N=30, incapacitating PMS</td>
<td>R, CO PC</td>
<td>2 x 3 Efamol capsules from cycle day 15; 4 cycles</td>
<td>19 symptoms rated on a 3-pt scale on the last treatment day</td>
<td>✓ depression; total score; responder rate: Efamol (62%), placebo (40%)</td>
<td>✓</td>
<td>Placebo treatment not defined. Double-blinded? Retrospective measurement? Exclusion of OC users not specified. Data for only 22 women analysed? Analysis unclear.</td>
</tr>
</tbody>
</table>
Table 1. (Continued.)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample size</th>
<th>Study design</th>
<th>Dose and duration</th>
<th>Diagnosis</th>
<th>Assessment measures</th>
<th>Outcome</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collins et al. (1993)</td>
<td>N=68, 38 completed; 44% dropout</td>
<td>R, DB CO, PC, paraffin oil</td>
<td>3 × 4 Efamol capsules/d for 4 cycles</td>
<td>DSM-III-R criteria; VAS prospective 16 symptom VAS (Hammarbeck et al., 1989)</td>
<td>× all mood and physical symptoms</td>
<td>Well controlled study.</td>
<td></td>
</tr>
<tr>
<td>Khoo et al. (1990)</td>
<td>N=38; no dropouts; moderate PMS</td>
<td>R, DB CO, PC, liquid paraffin</td>
<td>8 EPO capsules/d; 3 cycles</td>
<td>Retrospective 10 symptom scale</td>
<td>10 symptoms (4 psychological, 6 physical) completed retrospectively each cycle</td>
<td>× total PMS score; psychological, fluid retention, breast, and menstrual symptoms</td>
<td>Dose not stated. Diagnosis process unclear and seemed to be retrospective. It was unclear when symptoms were recorded in each cycle.</td>
</tr>
<tr>
<td>Vitex agnus castus</td>
<td>N=41, 38 completed; 7% dropout (fluoetine, 21; AC, 20) PMDD</td>
<td>R, SB PGs, PC</td>
<td>Fluoxetine or AC (20–40 mg/d); 2 cycles</td>
<td>DSM-IV; Penn daily symptom reports; 2 cycles</td>
<td>DSM-IV criteria for PMDD; premenstrual score of the DSR, HAM-D (Hamilton, 1960); CGI-I; agreement of improvement by the authors</td>
<td>✓ No significant difference between groups on DSR, CGI-SI scores or responder rates</td>
<td>AC compared to fluoxetine – no placebo group. Participants and raters blinded but prescribing physician not.</td>
</tr>
<tr>
<td>Halaska et al. (1998)</td>
<td>N=100, completion: AC 48, placebo 49, 3% dropout</td>
<td>DB, PGs PC</td>
<td>2 × 30 AC drops (1.8ml)/d; 3 cycles</td>
<td>Not specified</td>
<td>VAS scale</td>
<td>✓ breast pain</td>
<td>Only women suffering from mastalgia - 5 days/cycle included.</td>
</tr>
<tr>
<td>Reference</td>
<td>Sample size</td>
<td>Study design</td>
<td>Dose and duration</td>
<td>Diagnosis</td>
<td>Assessment measures</td>
<td>Outcome</td>
<td>Comments</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>--------------</td>
<td>----------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Schellenberg</td>
<td>N=170 (active, 86; placebo 84)</td>
<td>R, DB PGs, PC</td>
<td>1 x 20 mg AC tablet/d; 3 cycles</td>
<td>DSM-III-R</td>
<td>6 symptoms (irritability, mood alteration, anger, headache, other menstrual symptoms including bloating, breast fullness) on a VAS compared to previous 3 cycles; responder rate</td>
<td>✔ combined and individual symptoms (not bloating); responder rates 52% AC v 24% placebo</td>
<td>Assessment scale and timing of symptom rating was unclear. OC users included.</td>
</tr>
<tr>
<td>Turner &amp; Mills</td>
<td>N=600, completion: AC, 105; placebo 112, 64% dropout</td>
<td>R, DB PGs, PC</td>
<td>Treatment; 3 mths MDQ</td>
<td></td>
<td>MDQ at end of treatment; a shortened version administered at the end of cycles 1 and 2</td>
<td>✗ except for the symptom 'feel jittery or restless' and responder rates 25% AC v 16% placebo</td>
<td>Dose and frequency of administration was unclear. Only women high in negative affect included. Huge drop out rate, but evenly distributed across groups.</td>
</tr>
<tr>
<td>St John's Wort</td>
<td>N=169, 125 completed; 26% dropout</td>
<td>R, DB PGs, PC, lactose/cellulose</td>
<td>2 x 300 mg tablets standardized to 900 μg hypericin/d; 2 cycles</td>
<td>Retrospective assessment; 25 symptoms rated daily; 1 cycle</td>
<td>VAS assessing 25 symptoms grouped into 6 categories (anxiety, craving, depression, hydration, other and total)</td>
<td>✗ all symptom subgroups</td>
<td>Although diagnosis was prospectively confirmed, this process was unclear.</td>
</tr>
<tr>
<td>Reference</td>
<td>Sample size</td>
<td>Study design</td>
<td>Dose and duration</td>
<td>Diagnosis</td>
<td>Assessment measures</td>
<td>Outcome</td>
<td>Comments</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------------------------------------</td>
<td>--------------</td>
<td>-------------------</td>
<td>-----------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>----------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Ginkgo Biloba</td>
<td>N=165, congestive PMS symptoms</td>
<td>R, DB</td>
<td>EGb 761 from cycle day 16; 2 cycles</td>
<td>Observation of 1 menstrual cycle</td>
<td>Daily scale measuring congestion, breast tenderness and mood; practitioner observation premenstrually before and after treatment</td>
<td>✓ congestive symptoms, particularly breast symptoms</td>
<td>Method of diagnosis not specified.</td>
</tr>
</tbody>
</table>

**Key:**

**Study design**
- R Randomized
- DB Double-blind
- SB Single-blind
- PC Placebo-controlled
- PGs Parallel groups
- CO Crossover
- MC Multi-centred
- OC Oral contraceptives

**Measures**
- MHQ Mental Health Questionnaire (Warner & Bancroft, 1990)
- MDQ Menstrual Distress Questionnaire (Moos, 1968)
- VAS Visual analogue scale
congestive premenstrual symptoms for 7 days per cycle for the three cycles prior to recruitment.

Discussion

Many women with PMS use alternative therapies, despite the lack of established efficacy (Domoney et al., 2003). This review included 26 trials that assessed the efficacy of seven different dietary supplements and herbal remedies for PMS. The most substantial positive evidence was found for calcium and continuous vitamin B6 treatment. Trials assessing magnesium and evening primrose oil produced conflicting findings, whilst insufficient data were found to advocate the use of vitex agnus castus, gingko biloba or St John’s Wort.

The studies considered in this review differed greatly in the diagnostic methods they used. It is generally accepted that prospective daily self-report measures are needed to confirm PMS (Steiner & Wilkins, 1996). Some studies did diagnose by use of the DSM criteria, and confirmed their diagnoses prospectively. However, others relied upon retrospective diagnosis, which has been criticized (Connolly, 2001), since these often result in inflated estimates of symptom severity (De Souza et al., 2000).

The methods used for the assessment of treatment efficacy also differed. Many studies used the total symptom score of a rating scale as their primary outcome measure, and simultaneously considered symptom clusters, and some also considered individual symptoms. This increases the chances of finding symptom effects. Assessment measures were used prospectively with daily symptom ratings or by averaging these ratings across phases in some trials. Others assessed treatments retrospectively, at the end of each cycle or at the end of treatment, using a variety of methods, including questionnaires, interviews (Kendall & Schnurr, 1987; Loch et al., 2000) and general practitioner assessments (Smallwood et al., 1986; Williams et al., 1985). Some authors did not specify their treatment assessment methods (Mal mgren et al., 1987; Ockerman et al., 1986).

Women taking the oral contraceptive, which has previously been used as a PMS treatment (Girman et al., 2003), were not excluded from some studies, while others focused on specific groups of women, including women with ‘premenstrual tension depression’ (Stokes & Mendels, 1972), severe cyclical mastalgia (Smallwood et al., 1986) and congestive PMS symptoms (Tamborini & Taurelle, 1993). It is difficult to compare such studies with those examining a more representative sample.

Conclusion

The variations apparent in diagnostic procedures, assessment methods, and outcome measures make it difficult to assess treatment efficacy (Connolly, 2001). More consensus about the diagnosis, measurement and assessment of PMS is required, as are randomized, double-blind, placebo-controlled trials. Such carefully controlled trials with strict diagnostic criteria, prospective confirmation of PMS and prospective assessment of symptoms in response to treatment would help to clarify the efficacy of the many alternative treatments used for PMS. At the moment, calcium and continuous vitamin B6 treatment seem likely to be beneficial.
References


DIETARY SUPPLEMENTS AND HERBAL REMEDIES FOR PMS

John's Wort (Hypericum perforatum L.) for premenstrual symptoms. *Journal of Alternative and Complementary Medicine, 10*, 925–932.


