

University of Nottingham
UK | CHINA | MALAYSIA

The Investigation of The Relationships between Pre-treatment Mood, Lifestyle, Cortisol, and Outcomes of Assisted Reproductive Treatment

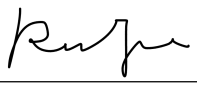
Thesis submitted to the University of Nottingham for the degree of **Doctor of Philosophy, March 2022.**

Ru Jia, MSc.

14303394

Supervised by

Professor Kavita Vedhara
Dr. Walid Maalouf

Signature 

Date 26 / August / 2022

Abstract

Patients seeking fertility treatment often experience considerable stress. Whether this stress influences treatment outcomes is one of the most frequently asked questions by patients undergoing *in vitro* fertilisation (IVF) treatment. However, the answer to this question is still the subject of much debate. One salient issue with current studies is that they have primarily focused on self-reported measures of stress while evidence on the physiological mechanisms that might give rise to a relationship between psychological stress and treatment outcomes is relatively absent. A candidate biomarker exists to explore precisely this relationship. The hypothalamic-pituitary-adrenal (HPA) axis has been identified as being central to the body's physiological response to stress, particularly in the context of chronic stress. The final hormonal product of the HPA axis, cortisol, is considered an appropriate 'biomarker of stress' and appropriate for use in studies investigating the effect of stress on pregnancy outcomes. However evidence is again inconclusive about the association between cortisol and IVF treatment outcomes. One outstanding methodological limitation of the previous evidence was the reliance on short-term measures of cortisol which, by their very nature, only provide information about cortisol levels over minutes (e.g., in saliva/blood) to hours (e.g., in urine/follicular fluid). The development of methods to measure systemic levels of cortisol over longer periods in hair has been an important innovation in the attempts to examine the relationship between stress and IVF treatment outcomes. Compared with other acute measures of cortisol, hair cortisol sampling is less invasive, less affected by compliance, and reflects the cumulative cortisol exposure over extended periods of up to six months. Recent studies have indicated an inverse relationship between higher hair cortisol levels and the likelihood of achieving pregnancy following IVF treatment. Such studies, however, are limited by small sample sizes and lack of inclusion of lifestyle factors (e.g., physical activity and sleep) which may influence both cortisol and IVF treatment outcomes.

This thesis presents findings from a prospective cohort study which aimed to build upon previous work by: (1) examining the relationship between hair cortisol levels before commencing IVF treatment, and clinical pregnancy; but also ex-

amining the relationship across two important treatment outcomes: biochemical pregnancy and discontinued treatment; (2) examining the relationship between lifestyle and mood factors and the three treatment outcomes both directly, and indirectly through the hair cortisol before commencing IVF treatment.

This study included 234 women undergoing IVF or intracytoplasmic sperm injection (ICSI) from a fertility clinic in the United Kingdom (January 2018 – August 2019), with an average age of 34.4 years ($SD = 4.6$). This study found no significant association between hair cortisol prior to commencing IVF/ICSI treatment and clinical pregnancy, biochemical pregnancy, or discontinued treatment. Lifestyle and mood factors were not significantly associated with any of the three treatment outcomes either directly or indirectly through pre-treatment hair cortisol. However, it was observed that women undergoing IVF/ICSI treatment were likely to report poor sleep quality and high levels of depression before treatment. Poor sleep quality was significantly correlated with greater anxiety, depression, negative mood, and lower positive mood. A positive association between hair cortisol levels and depression was also observed.

Implications of this study are three fold. Firstly, the null associations indicated that cortisol may not be a major pathway for chronic stress to influence treatment outcomes of IVF/ICSI treatment. Other pathways such as the sympathetic and parasympathetic systems could be focus of future studies. Secondly, although this study did not find significant associations between lifestyle and mood factors and treatment outcomes either directly or indirectly through hair cortisol, it suggested that women were likely to experience poor sleep quality before treatment, which might be relevant to their experience of stress. Thirdly, this study supported an association between hair cortisol and depression among women undergoing IVF/ICSI treatment. This study added to the evidence base that hair cortisol might be a useful tool to monitor long-term psychological distress. Under the current environment of a global pandemic which has been ongoing for more than two years at the time of writing, utilising hair cortisol as a tool to measure stress is valuable for understanding the long-term health consequences of chronic stress.

The unique contributions of this study lie in its methodological and theoretical

considerations: (1) use of hair cortisol to reflect chronic levels of stress; (2) inclusion of lifestyle and mood factors and use of mediation analysis to examine the possible indirect associations between these factors and treatment outcomes; (3) exploration of the association of lifestyle and mood factors and long-term cortisol levels to extend the literature that focused on acute measures of cortisol. Based on the findings of this thesis, future studies are encouraged to investigate how sleep quality and depression are associated and how they influence patients undergoing IVF treatment. Tools such as hair cortisol might be helpful to explore the underlying mechanism of these possible associations.

Keywords: stress, *in vitro* fertilisation, cortisol, lifestyle, mood

List of publications

PhD project:

Jia, R., Carlisle, S., & Vedhara, K. (2022). The Association of Lifestyle and Mood with Long-Term Levels of Cortisol: A Systematic Review. *Cogent Psychology*, 9(1), 2036487. (Winner of Cogent Psychology Research Spotlight Awards)

COVID-19 projects:

Jia, R., Ayling, K., Chalder, T., Massey, A., Broadbent, E., Coupland, C., & Vedhara, K. (2020). Mental health in the UK during the COVID-19 pandemic: cross-sectional analyses from a community cohort study. *BMJ Open*, *10*(9), e040620.

*Ayling, K., **Jia, R.**, Chalder, T., Massey, A., Broadbent, E., Coupland, C., & Vedhara, K. (2020). Mental Health of Keyworkers in the UK during the COVID-19 Pandemic: a Cross-sectional Analysis of a Community Cohort. *medRxiv*.

Jia, R., Ayling, K., Chalder, T., Massey, A., Broadbent, E., Morling, J., Coupland, C. & Vedhara, K. (2020). Young people, mental health and COVID-19 infection: The canaries we put in the coal-mine. *Public Health*, *189*, 158-161.

***Jia, R.**, ... Ayling, K. & Vedhara, K. (2020). Experiences of the COVID-19 pandemic: cross-sectional analysis of risk perceptions and mental health in a student cohort. *medRxiv*.

Knight, H., **Jia, R**., Ayling, K., Bradbury, K., Baker, K., Chalder, T., ..., & Vedhara, K. (2021). Understanding and Addressing Vaccine Hesitancy in the Context of COVID-19: Development of a Digital intervention. *Public Health*, *201*, 98-107.

Blake, H.,..., **Jia, R.**, & Vedhara K. (2021). Students' Views towards Sars-Cov-2 Mass Asymptomatic Testing, Social Distancing and Self-Isolation in a University Setting during the COVID-19 Pandemic: A Qualitative Study. *International Journal of Environmental Research and Public Health*, *18*(8):4182.

Gasteiger, N.,..., **Jia, R**.,..., & Broadbent, E. (2021). Depression, anxiety and stress during the COVID-19 pandemic: Results from a New Zealand cohort study on mental wellbeing. *BMJ Open*, *11*(5), e045325

Vedhara, K., Ayling, K., **Jia, R.**, Fairclough, L., Morling, J. R., Ball, J. K., ... & Tighe, P. (2022). Relationship Between Anxiety, Depression, and Susceptibility to Severe Acute Respiratory Syndrome Coronavirus 2 Infection: Proof of Concept. *The Journal of infectious diseases*.

Ayling, K., **Jia, R.**, Coupland, C., Chalder, T., Massey, A., Broadbent, E., & Vedhara, K. (2022). Psychological Predictors of Self-reported COVID-19 Outcomes: Results From a Prospective Cohort Study. *Annals of Behavioral Medicine*, *56*(5), 484-497.

Jia, R., Ayling, K., Chalder, T., Massey, A., Broadbent, E., Coupland, C., & Vedhara, K. (2022). The prevalence, incidence, prognosis and risk factors for symptoms of depression and anxiety in a UK cohort during the COVID-19 pandemic. *BJPsych open*, *8*(2)

Other project:

Carlisle, S., Ayling, K., **Jia, R.**, Buchanan, H & Vedhara, K. (2021). The effect of choice interventions on retention-related, behavioural and mood outcomes: A systematic review with meta-analysis. *Health Psychology Review*, 1-37.

*Pre-print

Acknowledgements

I would like to thank my supervisors Prof. Kavita Vedhara and Dr. Walid Maalouf for giving me this opportunity. Thank you for your guidance and support. I would also like to thank Dr. Adam Massey who laid foundations for this work and has guided me with his experience from the very beginning. I have learned how to be a researcher and how to think beyond from all of you.

I must give a special thank you to Prof. Kavita Vedhara. Thank you for believing in my potential and keeping inspiring me with your thoughts and actions. I am more than grateful to have the opportunity to work with you and learn from you. You have always taken the time to make sure that I am on track with my work. With everything that has been going on, especially the global pandemic, I feel very lucky that you have always been there to support me.

To my father Defeng Jia, and my mother Zhumei Jiang, words are simply not enough. I would not have been able to come this far without your support. Thank you for believing in me and encouraging me to pursue my dreams.

Finally, I would like to say thank you to all the staff at CARE Fertility Nottingham, the United Kingdom for helping me collect data for this thesis, and thank you to each patient for taking part in this study and sharing their experiences.

COVID-19 Impact Statement

Towards the end of this PhD, the outbreak of the SARS-CoV-2 virus was declared a global pandemic (COVID-19) by the World Health Organisation (WHO). Data collection for this thesis was fortunately completed before this and hence not affected. However, the progress of this thesis was influenced in several ways.

At the beginning of the COVID-19 pandemic, my ability to work was restricted. At the time when the first lockdown was administrated in the UK, I was living in student accommodation with no working equipment (e.g., personal computer) and unpleasant flatmates. Therefore, my work progress was delayed until these issues were resolved.

However, I was soon actively involved in several COVID-19 related research programmes led by Prof. Kavita Vedhara: COVID-19 Stress and Health Study and Track-COVID Study. In COVID-19 Stress and Health Study, I was able to extend my knowledge on stress and cortisol and was involved in the design and delivery of a longitudinal observational study investigating the levels of emotional distress, levels of hair cortisol and cortisone, and COVID-19 related outcomes (e.g., SARS-CoV-2 infection and symptoms) in the general population. I contributed to the rapid recruitment of more than 3,000 participants to the study within weeks during the first lockdown in April 2019. Building on this work, I was then involved in the Track-COVID study where emotional distress and COVID-19 related health behaviours (e.g., social distancing, self-isolation) among a university student cohort were investigated. I was also involved in the work to address COVID-19 vaccine hesitancy among these established cohorts, followed by the development of a digital intervention to tackle COVID-19 vaccine hesitancy among the general population. This work has all taken place whilst I was completing this thesis and has resulted in the eight papers and a further one which is currently under review on which I am the first author.

To many people, the COVID-19 pandemic has unfortunately taken away their opportunities. I was fortunate to have had these new opportunities and have embraced some of the challenges brought by the COVID-19 pandemic.

Contents

Abstract i

List of Publications iv

Acknowledgements vi

COVID-19 Impact Statement vii

List of Tables xi

List of Figures xii

Chapter 1 Infertility and *In Vitro* Fertilisation 1

1.1 The incidence of infertility 2

1.1.1 Types and causes of infertility 3

1.1.2 Treatment options for infertility 7

1.2 *In vitro* fertilisation 11

1.2.1 Development and facts 11

1.2.2 Cost and cost-effectiveness 12

1.2.3 A typical treatment cycle 13

1.3 The success of an IVF cycle 20

1.3.1 Defining success 20

1.3.2 Typical success rate of IVF 21

1.3.3 Risks of the treatment 22

1.3.4 Towards success: predictors of treatment outcomes 25

Chapter 2 Psychological Stress, Physiological Aspects of Stress, and Their Roles in IVF 28

2.1 Stress theories 29

2.1.1 The concept of stress and stressor 29

2.1.2 The stress and coping theory 30
2.1.3 The allostasis model of stress 31
2.2 Physiology of stress 32
2.2.1 The nervous system 32
2.2.2 The neuroendocrine system 33
2.2.3 The Sympathetic-Adrenal-Medullary (SAM) pathway . . . 34
2.2.4 The Hypothalamic-Pituitary-Adrenal (HPA) pathway . . . 35
2.3 Stress in the context of fertility and IVF treatment 38
2.3.1 Psychological stress among patients with fertility problems 39
2.3.2 Does stress influence IVF treatment? 40
2.4 Does cortisol influence pregnancy outcomes in IVF treatment? . . 47
2.4.1 Cortisol and health 47
2.4.2 Cortisol: fractions, indices and measures 50
2.4.3 The association of cortisol and IVF treatment: physiological pathway 55
2.4.4 The association of cortisol and IVF treatment: systematic review by Massey et al. (2014) 56
2.4.5 The role of hair cortisol in pregnancy in IVF treatment: Massey et al. (2016) 58
2.4.6 Recent evidence of the association of hair cortisol with IVF treatment outcomes 60
Chapter 3 The Role of Lifestyle 65
3.1 The role of lifestyle in fertility, infertility, and IVF treatment . . . 66
3.1.1 Behavioural determinants of reproductive health 66
3.2 Lifestyle behaviours among patients undergoing IVF treatment . . 67
3.2.1 Diet and body mass index 68
3.2.2 Physical activity and exercise 70
3.2.3 Smoking 72
3.2.4 Alcohol consumption 74

3.2.5 Caffeine consumption 77
3.2.6 Sleep 79
3.3 An individual pathway to fertility 84
Chapter 4 Lifestyle, Psychological Factors and Cortisol 87
4.1 Associations of lifestyle and psychological factors with long-term levels of cortisol 89
4.2 Systematic review: methods 92
4.2.1 Eligibility criteria 92
4.2.2 Exclusion criteria 94
4.2.3 Information source 95
4.2.4 Data extraction and meta-analysis 95
4.2.5 Assessment of risk of bias 96
4.3 Systematic review: results 97
4.3.1 Study characteristics 97
4.3.2 Population 99
4.3.3 Type of cortisol samples and time frame investigated . . . 101
4.3.4 Lifestyle correlates of cortisol 103
4.3.5 Mood correlates of cortisol 120
4.3.6 Effects of mood interventions on cortisol 130
4.4 Systematic review: discussion 139
4.4.1 Heterogeneity in how lifestyle and mood were measured . . 141
4.4.2 Heterogeneity in approaches to measuring cortisol 143
4.4.3 The temporal issue of the association between cortisol and lifestyle or mood 147
4.4.4 Issues from the interventions 151
4.4.5 Strengths and limitations 154
4.4.6 Recommendations for future research 155
Chapter 5 Cohort Study: Aims and Methods 157
5.1 Study aims and hypotheses 159

5.1.1 Aim 1 . . . 159
5.1.2 Aim 2 . . . 160
5.1.3 Aim 3 . . . 161
5.2 Methods . . . 162
5.2.1 Participants . . . 162
5.2.2 Ethical approval and considerations . . . 163
5.2.3 Procedure . . . 164
5.2.4 Data management . . . 166
5.2.5 Measures . . . 166
5.2.6 Clinical outcomes . . . 174
5.2.7 Hair cortisol assay . . . 175
5.2.8 Approaches to data analysis . . . 176
5.2.9 Hair cortisol data . . . 177
Chapter 6 Cohort Study: Results 185
6.1 Patient flow . . . 186
6.2 Cohort characteristics . . . 188
6.3 Part 1: cortisol levels and *in vitro* fertilisation treatment outcomes 191
6.3.1 Hair cortisol concentration among patients undergoing *in vitro* fertilisation treatment . . . 191
6.3.2 Predicting three treatment outcomes: examining demographic and clinical covariates . . . 192
6.3.3 Relationship between hair cortisol concentration in the three months before treatment and outcomes of IVF treatment: univariate analysis . . . 196
6.3.4 Relationship between hair cortisol concentrations in the three months before treatment and outcomes of IVF treatment: multivariate analysis with results from the IBL assay 197

6.3.5 Relationship between hair cortisol concentration in the three months before treatment and outcomes of IVF treatment: multivariate analysis with results from the DRG assay . . 199
6.3.6 Relationship between hair cortisol concentration in the three months before treatment and outcomes of IVF treatment: multivariate analysis with combined results from two assays 200
6.3.7 Relationship between hair cortisol concentrations in the month before treatment and outcomes of IVF treatment: univariate analysis 202
6.3.8 Relationship between hair cortisol concentrations in the month before treatment and outcomes of IVF treatment: multivariate analysis with results from the IBL assay . . . 203
6.3.9 Relationship between hair cortisol concentrations in the month before treatment and outcomes of IVF treatment: multivariate analysis with results from the DRG assay . . 205
6.3.10 Relationship between hair cortisol concentrations in the month before treatment and outcomes of IVF treatment: multivariate analysis with combined results from two assays 206
6.3.11 Sensitivity analysis using multiple imputation 208
6.4 Part 2: lifestyle of women undergoing IVF treatment 209
6.4.1 Lifestyle in patients undergoing IVF treatment 209
6.4.2 Lifestyle and treatment outcomes in patients undergoing IVF treatment: univariate analyses 212
6.4.3 The association between lifestyle factors and 3 cm hair cortisol concentrations: univariate analyses 213
6.4.4 The association between lifestyle and 1 cm hair cortisol concentrations: univariate analyses 213
6.5 Part 3: mood of women undergoing IVF treatment 216

6.5.1 Mood and treatment outcomes in patients undergoing IVF treatment . . . 216
6.5.2 The association between mood and 3 cm hair cortisol concentrations: univariate analyses . . . 219
6.5.3 The association between mood and 1 cm hair cortisol concentrations: univariate analyses . . . 220
6.5.4 Correlations between lifestyle and mood factors . . . 221
6.5.5 Depression predicting biochemical pregnancy through 3 cm hair cortisol concentrations: mediation analysis . . . 224

Chapter 7 Discussion 228

7.1 The association between hair cortisol concentrations and pregnancy outcomes . . . 229
7.1.1 Hair cortisol and clinical pregnancy . . . 229
7.1.2 Is cortisol the pathway through which stress influences *in vitro* fertilisation treatment? . . . 230
7.1.3 Is stress associated with pregnancy outcomes following *in vitro* fertilisation treatment? . . . 231
7.1.4 Hair cortisol and biochemical pregnancy . . . 236
7.1.5 The association between hair cortisol and discontinued treatment . . . 238
7.2 The direct and indirect association between lifestyle factors and *in vitro* fertilisation treatment outcomes . . . 241
7.2.1 Sleep . . . 241
7.2.2 Physical activity . . . 244
7.2.3 Caffeine consumption . . . 247
7.2.4 Alcohol consumption . . . 251
7.2.5 Smoking . . . 254
7.3 Mood among women undergoing assisted reproductive treatment and treatment outcomes . . . 255

7.3.1 Direct association between mood and treatment outcomes 256
7.3.2 Indirect association between mood and treatment outcomes 257
7.4 High levels of depression among women undergoing assisted reproductive treatment . . . 260
7.4.1 Differences in hair cortisol values from the two assays . . . 261
7.5 Study limitations . . . 263
7.6 Alternative designs . . . 264
7.7 Future recommendations . . . 266
7.7.1 Emotional impact of the COVID-19 pandemic on women undergoing fertility treatment . . . 267
7.8 Conclusion . . . 269
Bibliography **271**
Appendices **323**
Appendix A Systematic Review Supplementary Materials **324**
A.1 Summary of studies . . . 324
A.1.1 Table S1 Characteristics of studies . . . 325
A.1.2 Table S2 Summary of studies investigating long-term cortisol and lifestyle factors . . . 328
A.1.3 Table S3 Summary of studies investigating long-term cortisol and mood factors . . . 336
A.2 Risk of bias assessment . . . 346
A.2.1 Table S4 Summary of the risk of bias of studies assessed by the 27-item Downs and Black checklist (Downs and Black 1998) . . . 347
A.2.2 Table S5 Categorisation of the total risk of bias scores assessed by the 27-item Downs and Black checklist (Downs and Black 1998)[1] . . . 348
A.3 Search Terms . . . 349

A.3.1 Table S6 Terms of cortisol, lifestyle and mood factors for database Search 349

A.3.2 Search terms used for MEDLINE, Embase, PsycInfo 349

Appendix B Research Ethics 352

B.1 Patient Information Sheet 353

B.2 Participant Consent Form 359

B.3 Ethics approval letter: Research Ethics Committee 361

B.4 Ethics approval letter: Health Research Authority 362

Appendix C Questionnaires 363

C.1 Lifestyle and mood questionnaires 364

List of Tables

2.1 Summary of the characteristics and measures of cortisol indices . 54

5.1 HCC absolute values before transformation 178

6.1 Cohort characteristics 189

6.2 Univariate logistic regression models showing hair characteristic predictors of hair cortisol levels 192

6.3 Univariate logistic regression models showing individual predictors of treatment outcomes 194

6.4 Hair cortisol concentration (HCC) values in patients by treatment outcome and reference group 196

6.5 Univariate logistic regression models showing 3 cm hair cortisol concentrations (HCC) predicting treatment outcomes 197

6.6 Multivariate logistic regression models showing 3 cm hair cortisol concentrations (HCCs) from the IBL assay predicting treatment outcomes while controlling for age and miscarriage history 198

6.7 Multivariate logistic regression models showing combined 3 cm hair cortisol concentrations (HCCs) from the DRG assay predicting treatment outcomes while controlling for age and miscarriage history 200

6.8 Multivariate logistic regression models showing combined 3 cm hair cortisol concentrations (HCCs) from both assays predicting treatment outcomes while controlling for age and miscarriage history . 202

6.9 Univariate logistic regression models showing 1 cm hair cortisol concentrations (HCCs) predicting treatment outcomes 203

6.10 Multivariable logistic regression models showing 1 cm hair cortisol concentrations (HCCs) from the IBL assay predicting treatment outcomes while controlling for age and miscarriage history 204
6.11 Multivariate logistic regression models showing combined 1 cm hair cortisol concentrations (HCCs) from the DRG assay predicting treatment outcomes while controlling for age and miscarriage history 206
6.12 Multivariate logistic regression models showing combined 1 cm hair cortisol concentrations (HCCs) from both assays predicting treatment outcomes while controlling for age and miscarriage history . 207
6.13 Missing value table 208
6.14 Sensitivity analysis multivariate logistic regression models using multiple imputation datasets to show combined 3 cm hair cortisol concentrations (HCCs) from both assays predicting treatment outcomes while controlling for age and previous miscarriage 208
6.15 Sensitivity analysis multivariate logistic regression models using multiple imputation datasets to show combined 1 cm hair cortisol concentrations (HCCs) from both assays predicting treatment outcomes while controlling for age and previous miscarriage 209
6.16 Lifestyle factors of patients undergoing *in vitro* fertilisation/intracytoplasmic sperm injection treatment by treatment outcomes . . 211
6.17 Univariate logistic regressions showing the association between lifestyle factors and *in vitro* fertilisation/intracytoplasmic sperm injection treatment outcomes 212
6.18 Univariate linear regression models showing lifestyle predictors of 3 cm hair cortisol concentrations (HCCs) from the independent assays and combined assay results 214

6.19 Univariate linear regression models showing lifestyle predictors of 1 cm hair cortisol concentrations (HCCs) from independent assays and the combined assay results 215
6.20 Mood among IVF patients by treatment outcomes 217
6.21 Univariate linear regression models showing mood predictors of 3 cm hair cortisol concentrations (HCCs) from the independent assays and the combined assay results 218
6.22 Univariate linear regression models showing mood predictors of 3 cm hair cortisol concentrations (HCCs) from the independent assays and the combined assay results 220
6.23 Univariate linear regression models showing mood predictors of 1 cm hair cortisol concentrations (HCCs) from the independent assays and the combined assay results 221
6.24 Correlations between lifestyle and mood factors 223

List of Figures

1.1 Flow diagram of a typical cycle of IVF treatment 19

2.1 Examples of the circadian rhythm of the adrenocorticotropic hormone (ACTH) and cortisol 37

4.1 Thesis conceptual model 88

4.2 The PRISMA flow diagram of systematic review search and study selection 98

4.3 Random effect meta-analysis for the correlation between long-term levels of cortisol and BMI 105

4.4 Random effect meta-analysis for the correlation between long-term levels of cortisol and physical activity 109

4.5 Random effect meta-analysis for the correlation between long-term levels of cortisol and perceived stress 122

4.6 Random effect meta-analysis for the correlation between long-term levels of cortisol and mood impact of stressors 123

4.7 Random effect meta-analysis for the correlation between long-term levels of cortisol and other indices of stress 125

4.8 Random effect meta-analysis for the correlation between long-term levels of cortisol and anxiety 126

4.9 Random effect meta-analysis for the correlation between long-term levels of cortisol and depression 129

5.1 Thesis conceptual model 158

5.2 Flow of the treatment and recruitment procedure 165

5.3 HCC absolute values (pg/mg) by ELISA 179
5.4 Transformed HCC values (pg/mg) by ELISA 179

6.1 Flow diagram of patient recruitment and data collection 187
6.2 Path analysis with depression predicting biochemical pregnancy through 3cm HCC from the IBL assay 224
6.3 Path model 1 showing the association between Depression and biochemical pregnancy through 3cm HCCs from the IBL assay 226
6.4 Path model 2 showing the association between Depression and biochemical pregnancy through 3cm HCCs from the IBL assay 227

7.1 Modified thesis model based on findings from the cohort study and systematic review 270

Chapter 1

Infertility and *In Vitro* Fertilisation

Infertility, as a worldwide public health issue (World Health Organization [WHO], 2002), is the starting point of this thesis. *In vitro* fertilisation (IVF), as a commonly performed medical technique to help people experiencing infertility, has gained popularity over the past 30 years (Human Fertilisation and Embryology Authority, 2018). However, the success rate of IVF treatment appears not to have advanced as quickly as its popularity, undoubtedly compounded by the fact that many of the factors known to affect treatment success are largely unmodifiable, for instance: age, ovarian reserve, and sperm quality (Human Fertilisation and Embryology Authority, 2018; van Loendersloot et al., 2010). Therefore, identifying any modifiable factors that are associated with the likelihood of successful treatment may help patients to achieve better outcomes following the expensive and invasive treatment. The first chapter of this thesis will introduce the concepts of infertility, possible causes of infertility, and lay out a brief introduction to IVF treatment: its history, epidemiological facts, standard treatment procedures, endpoints, risks, and possible predictors of treatment success.

1.1 The incidence of infertility

Infertility, by definition, is a disease characterised by the failure to establish a clinical pregnancy after 12 months of regular, unprotected sexual intercourse due to an impairment in a person's capacity to reproduce, either as an individual or as a couple (Zegers-Hochschild et al., 2014). In research, subfertility is sometimes used interchangeably with infertility. According to the World Health Organisation (WHO, 2002), infertility is a global public health issue affecting up to 15% of couples of reproductive age worldwide, despite the different estimates of the infertile population reported in various studies. Boivin et al. (2007) reviewed a total of 25 population surveys that reported the prevalence of infertility in more and less developed countries, sampling 172,413 women aged 20–44 years. They estimated that in 2006, approximately 72.4 million women worldwide were infertile, among whom 40.5 million were seeking infertility medical care (Boivin et al., 2007). They also observed that the prevalence of infertility was similar between more and less developed countries (3.5%–16.7% and 6.9%–9.3%, respectively). Mascarenhas et al. (2012) estimated that in 2010, 48.5 million child-seeking women worldwide were infertile. Their estimate was lower than that of Boivin et al. (2007) due to a larger dataset and a different statistical approach. Mascarenhas et al. (2012) also provided more information on regional differences, e.g., the prevalence of infertility was the highest in South Asia, Sub-Saharan African, North Africa/Middle East, Central/Eastern Europe, and Central Asia out of the 190 countries and territories included in the study. In addition to studies that based estimations on data from only females, several studies also estimated the incidence of infertility based on data acquired from both genders. Louis et al. (2013) gathered data from a national representative sample of males aged 15–44 years and estimated an overall infertility prevalence of 12% in the US (Louis et al., 2013). This estimate was consistent with other estimates of 15.5%, using the same method, based on the female population in the US (Thoma et al., 2013). Similar figures

were reported in a recent investigation in the UK where Datta et al. (2016) collected data from 15,162 males and females aged 16–74 years and estimated the prevalence of infertility to be 12.5% in females and 10.1% in males.

1.1.1 Types and causes of infertility

Infertility is not a homogeneous condition with a single etiological cause, and it is usual to distinguish between primary and secondary infertility. Primary infertility refers to people experiencing difficulty conceiving who have not previously had a child: this is estimated to affect 1.9% of females of reproductive age seeking to conceive. In contrast, secondary infertility refers to people experiencing difficulty conceiving who have previously had a child: this is estimated to affect 10.5% of females of reproductive age seeking to conceive (Mascarenhas et al., 2012).

In terms of understanding the causes of infertility, it is helpful to first understand the processes involved in natural conception. Humans require gametes from both sexes, i.e., an oocyte (or occasionally oocytes) from the female and sperm from the male. Males can produce approximately 1,000 sperm every second (Johnson, 2018). However, females are born with a fixed number of follicles (approximately one million at birth and retaining 300,000 at puberty), which will grow and release oocytes (i.e., eggs), on a monthly basis during a female's reproductive lifespan (van Zonneveld et al., 2001). In other words, females can only generate a limited number of oocytes during their lives. In most cases, a female produces one oocyte per month (Johnson, 2018). The monthly preparation for pregnancy is termed the menstrual cycle, which includes the production of the oocyte and the preparation of the uterus for pregnancy (Coward and Wells, 2013). The majority of females have menstrual cycles of 23–33 days. Ovulation (i.e., the natural release of a mature oocyte from its ovarian follicle) generally occurs around day 16 of a 30 day menstrual cycle. In cases where no fertilisation occurs, a period will occur roughly 14 days following ovulation (Coward and Wells, 2013).

For successful conception to happen naturally, sufficient numbers of active sperm have to swim into the cervix and reach the fallopian tubes at the proper time, when the mature oocyte is travelling from the ovary through the fallopian tubes after ovulation. If the oocyte and sperm successfully meet and combine, the oocyte becomes fertilised, which is known as an embryo. Fertilisation normally takes place in the fallopian tubes, which are also involved in early embryo development and the transport of the embryo into the uterine cavity, where the embryo may implant into the lining of the womb and further develop into a foetus, resulting in pregnancy. For a successful pregnancy to occur, the male and female reproductive systems must both be anatomically and physiologically intact. A problem with any of the aforementioned steps may result in difficulties conceiving (Johnson, 2018). Infertility is usually not attributable to a single factor. One-third of all diagnosed cases are due to female factors only and one-fifth are due to male factors only, with disorders from both the male and female present in approximately 39% of all diagnosed cases (Thonneau et al., 1991).

Female factors

Female infertility accounts for approximately 35% of all cases (Coward and Wells, 2013). A number of underlying conditions may cause female infertility. Examples of these conditions include diminished ovarian reserve, ovulation disorders, tubal blockage or disease, and uterine abnormalities.

Ovarian reserve is often used to describe the number and/or quality of oocytes, reflecting the ability to reproduce. Ovarian reserve declines with increasing age (Gleicher et al., 2011). Most patients who receive a diagnosis of diminished ovarian reserve will proceed with IVF treatment or consider other alternatives, such as adoption (Seli et al., 2011).

Ovulation disorders refer to conditions in which follicular development and ovu-

lation functions are impaired, and hence, the oocyte is not released from its follicle (Brugo-Olmedo et al., 2001). Women with ovulation disorders often present with irregular menstrual periods (oligomenorrhoea) or absence of menstrual periods (amenorrhoea). Ovulation disorders can be caused by hypothalamic pituitary failure, dysfunction of the hypothalamic–pituitary–ovarian axis, or ovarian failure. Hypothalamic pituitary failure is often characterised by a deficiency of gonadotrophin or oestrogen, which are essential hormones for the reproductive system to function. Patients with dysfunction of the hypothalamic–pituitary–ovarian axis typically present with polycystic ovaries/polycystic ovarian syndrome (PCOS), which may result in imbalanced levels of sex hormones (e.g., raised levels of testosterone or androgen). Approximately 20% – 30% of the causes of infertility are due to ovulatory disorders (Seli et al., 2011).

Tubal factors, including blockage of a fallopian tube, tubal obstruction, and pelvic adhesions due to infection, endometriosis, and previous surgery, account for about 14% of cases of female infertility. Endometriosis is a disease characterised by the presence of endometrial tissue occurring outside the uterine cavity, most commonly in the pelvis (Coward and Wells, 2013). Endometriosis alone accounts for approximately 5% of female infertility (Seli et al., 2011).

The process of implantation, placentation, and embryo development are dependent upon the uterus. Uterine abnormalities, either congenital or acquired (e.g., intrauterine exposure to drugs), have been associated with infertility or recurrent abortions (Brugo-Olmedo et al., 2001).

Male factors

Male factors account for approximately 35% of all infertility cases (Coward and Wells, 2013). Common male factors include ejaculation dysfunction, abnormal semen, endocrine disturbances, and immunological factors. Ejaculation dysfunction

interferes with the capacity to deposit semen into the vagina during intercourse and is a frequent cause of male factor infertility. Inability to have or maintain an erection long enough to allow sexual intercourse to take place (erectile dysfunction), occurrence of ejaculation outside the vagina (premature ejaculation), and retrograde semen passing backwards into the bladder (retrograde ejaculation) are examples of ejaculation dysfunction. Abnormal semen can result from a low sperm count, lack of sperm, sperm with motor dysfunction, or sperm with abnormal shape. Damage to the testicles may also reduce the quality of semen. However, the majority of cases of abnormal semen remain unexplained (National Collaborating Centre for Women's and Children's Health (UK), 2013). It has been suggested that an inherently poor sperm production capacity of the testes with a likely genetic origin may be a cause of abnormal semen (National Institute for Clinical Excellence, 2016). Endocrine disturbances refer to abnormal levels of sex hormones in males. For example, abnormally low levels of testosterone or lack of gonadotrophin stimulation of the testes may lead to infertility in males. Immunological causes refer to the presence of antisperm antibodies in the seminal fluid or bound to motile sperm cells. Immunological factors account for approximately 3% of male factor infertility (Meniru, 2001). The exact mechanism by which antisperm antibodies cause male infertility is unclear. However, evidence suggests that antisperm antibodies negatively affect sperm concentration, motility, and liquefaction, which are indicators of sperm quality (Cui et al., 2015).

Unexplained infertility

Apart from known female or male factors, other diagnoses of infertility are "unexplained," where the infertility remains unexplained after the completion of the standard protocol of investigations (e.g., a semen analysis and chlamydia test for male factors and blood test and ultrasound scan for female factors (Templeton and Penney, 1982; National Collaborating Centre for Women's and Children's

Health (UK), 2013). According to various estimates, the incidence of unexplained infertility is 22.3%–38% (Collins and Rowe, 1989; Templeton and Penney, 1982; Wilkes et al., 2009).

1.1.2 Treatment options for infertility

Different treatment options are available to people with infertility, depending on their clinical diagnoses. Treatment options often include medical treatment (pharmacological treatment), surgical treatment, and assisted reproductive technology (ART). This section will briefly introduce several common treatment options for infertility.

Medical treatment

Medical treatment is often used for women with ovulatory dysfunction, including PCOS, and men and women with fertility issues related to hormonal disturbances. Fertility medications are usually used to stimulate the endocrine system to encourage ovulation in females and to improve sperm count in males (National Collaborating Centre for Women's and Children's Health (UK), 2013). Medical treatment for females with ovulatory disorders, with the intention of inducing normal ovulatory cycles, is termed ovulation induction. Examples of common fertility medications include clomiphene or tamoxifen (which both promote ovulation in women with ovulatory dysfunction), metformin (which can lower insulin and blood glucose levels and stimulate ovulation, particularly helpful for women with PCOS), and dopamine agonists (which encourage ovulation). Medications, such as pseudoephedrine and imipramine, are used for ejaculation failure in males. Sex hormones can also be given to both females and males to improve fertility. Examples include gonadotrophins, which can help stimulate ovulation in females and may also improve fertility in males, and gonadotrophin-releasing hormone,

which can encourage ovulation in females. Medical treatment is unlikely to increase the chances of pregnancy among women and men experiencing unexplained infertility; therefore, it is not recommended for this patient group (National Collaborating Centre for Women's and Children's Health (UK), 2013).

Surgical procedures

According to the National Health Service (NHS), surgical procedures are treatment options for females with blocked or scarred fallopian tubes, endometriosis, fibroids, or PCOS. Surgery can break up scar tissue in females' fallopian tubes, allowing the oocytes to pass through. For females with endometriosis, surgery may be indicated to excise the endometriotic tissue, release adhesions, and repair the associated damage to the pelvic organs. Submucosal fibroids (small growths in the uterus) can be surgically removed. A surgical procedure named laparoscopic ovarian drilling can be used as an alternative to medical treatment for women with PCOS. For infertile males, surgery can help in the following ways: (1) relieving an obstruction of the vas deferens, (2) curing a varicocele, or (3) bringing down ectopic testes (Meniru, 2001). The use of surgical procedures is highly dependent on the patient's diagnosis and bodily condition (National Collaborating Centre for Women's and Children's Health (UK), 2013). Although surgical procedures still play an important role in the management of infertility, the application of surgical treatment options is being re-evaluated due to developments in ART (Meniru, 2001).

Assisted reproduction technology (ART)

Apart from medication and surgical procedures, assisted reproduction technology (ART), including intrauterine insemination (IUI), IVF, and intracytoplasmic sperm injection (ICSI), can also help people who are having difficulty conceiving.

Intrauterine insemination (IUI)

Intrauterine insemination is one of the simplest and most commonly performed assisted fertility treatments (Meniru, 2001). It is performed using a very thin catheter inserted through the uterine cervix to inject washed sperm directly into the uterus where the highest quality sperm are left to fertilise the oocytes naturally. The purpose of the sperm wash in IUI is to separate the high-quality sperm from sperm that are sluggish, still, or abnormally shaped. The injection is timed to occur as close to ovulation as possible (Meniru, 2001). In some cases, fertility drugs are used to stimulate ovulation before IUI.

In IUI, either with a natural or stimulated ovulation, the process of fertilisation and pregnancy happen naturally, which requires reasonable quality sperm and an intact female reproductive system. Therefore, according to the National Institute for Health and Care Excellence (NICE; 2013), IUI should be offered as an alternative to conception via vaginal sexual intercourse in the following cases: (1) inability to have or difficulty with having vaginal sexual intercourse, (2) conditions that require specific consideration in relation to methods of conception (e.g., sperm washing for a man who is HIV positive), or (3) same-sex couples. However, IUI is not a preferred treatment option for people with unexplained infertility, blocked fallopian tubes, mild endometriosis, or male factor infertility (National Collaborating Centre for Women's and Children's Health (UK), 2013).

In vitro fertilisation (IVF)

As the name suggests, the process of fertilisation happens in glass in the laboratory for IVF. In brief, a treatment cycle usually starts with manipulation of the female menstrual cycle and stimulation of ovulation using fertility drugs. Following ovulation, mature oocytes are surgically collected from the woman's ovaries and fertilised with prepared sperm in a laboratory. The fertilised oocyte, i.e.,

the embryo, is allowed to develop in the laboratory for 2–6 days before being transferred back into the woman's uterus for further development.

The NICE guideline (National Collaborating Centre for Women's and Children's Health (UK), 2013) recommends that IVF should be offered according to the following criteria: (1) women with blocked or damaged fallopian tubes, (2) people with unexplained fertility problems or who have had other treatment (e.g., IUI) that failed, (3) men with low sperm count or high numbers of sperm with abnormal shape or movement, (4) women with difficulty ovulating who have not been successful with other treatments, and (5) older women (i.e., in their late 30s and older) who are less likely to be successful with more natural or less invasive treatments, such as IUI.

Intracytoplasmic semen injection (ICSI)

Intracytoplasmic semen injection (ICSI) is a procedure where a single sperm is injected into an oocyte. The technology of ICSI was developed in the 1990s and was introduced as an extension to IVF treatment in cases where men had poor semen quality (Palermo et al., 1992). According to NICE (2013), ICSI is performed as an optional part of IVF and is recommended for people with male factor infertility (for example severe deficits in semen quality) and azoospermia (absence of sperm in a semen sample). As no clear evidence existed, the NICE guidelines (National Collaborating Centre for Women's and Children's Health (UK), 2013) recommended that couples should be informed that ICSI improves fertilisation rates, but not pregnancy rates, when compared with IVF alone.

1.2 *In vitro* fertilisation

1.2.1 Development and facts

In vitro fertilisation was initially developed to help women who are infertile due to tubal diseases or the absence of fallopian tubes following surgery (Edwards et al., 1980). Robert Geoffrey Edwards, who was awarded the Nobel Prize in 2010 for the "development of human *In Vitro* fertilisation (IVF) therapy," described the intention of his and his colleagues' work as "to recover pre-ovulatory oocytes by laparoscopy, fertilise them *in vitro* using sperm from the husband, grow the embryos in culture through three or four cleavage divisions, and then place them in the mother's uterus" (Edwards et al., 1980). The efforts resulted in the first human IVF child in the world, Louise Brown, who was born in the UK on July 25th, 1978 (Edwards et al., 1980). This was followed by the first IVF child born in Australia in 1980, in the US in 1981, in continental Europe in 1982, and many other births around the world in 1982 and 1983 (Fauser and Edwards, 2005). The technology and treatment regimens of IVF treatment continued to improve. For example, in the early 1980s, the ovarian stimulation regimens using different compounds was developed, and the use of transvaginal ultrasound, which allowed the monitoring of the ovarian response and oocyte retrieval without general anaesthesia, was developed (Fauser and Edwards, 2005). In the 1990s, the development of ICSI allowed couples suffering from severely compromised male fertility to have their own genetic children with the help of IVF (Devroey and Van Steirteghem, 2004).

IVF has become a widely accepted and routinely performed fertility treatment (Pandian et al., 2015). The International Committee for Monitoring Assisted Reproductive Technology (ICMART) estimated that the total global utilisation of IVF, including ICSI, was more than one million cycles, resulting in approximately 400,000 babies, in 2011 (Adamson et al., 2018). In the UK, according to the

Human Fertilisation and Embryology Authority (HFEA; 2018), licenced clinics have carried out over one million IVF treatment cycles from 1991 to 2016. In 2016, over 68,000 IVF treatment cycles were carried out, a 4% increase over 2015 (Human Fertilisation and Embryology Authority, 2018). A population-based study estimated that in the UK, approximately half of the couples who were diagnosed with infertility were offered or received IVF or ICSI as a first-line treatment (Wilkes et al., 2009). In 2016, 42% of IVF patients across the UK were aged <35 years, and 58% were ≥35 years (Human Fertilisation and Embryology Authority, 2018). The average live birth rate from IVF was 26% for fresh treatment cycles and 28% for frozen treatment cycles in 2016, which increased from approximately 15% in 1990 for both fresh cycles and frozen cycles (Human Fertilisation and Embryology Authority, 2018).

1.2.2 Cost and cost-effectiveness

According to the HFEA, the average cost of an IVF cycle is £3,000–£5,000. A health economic analysis (National Collaborating Centre for Women's and Children's Health (UK), 2013) revealed that the cost effectiveness of IVF treatments varies across age groups and diagnoses, e.g., up to three cycles of IVF was cost effective for women aged <39 years, but one or more cycles of IVF was not cost-effective for women aged 40–42 years with unexplained infertility, male factor infertility, or mild endometriosis or for women aged ≥43 years. Although NICE recommended that women <40 years should be offered three cycles of IVF treatment by the NHS, the final decision about whether an individual can receive NHS-funded IVF in their local area may follow stricter criteria (NHS, 2018). According to the HFEA report Human Fertilisation and Embryology Authority (2018): only 41% of IVF treatments were funded by the NHS in 2016. Kennedy et al. (2006) collected data from 58% of the licensed IVF centres in England, which performed nearly 50% of all IVF cycles in the UK, and assessed the imple-

mentation status of the funding recommendations made by NICE in 2004. They discovered that 90% of the centres were being commissioned for one IVF cycle, only three centres (9%) reported routine commissioning of two cycles, and no centres were being commissioned for three treatment cycles (Kennedy et al., 2006). The latest data from the HFEA (2018) indicated a reduction in the number of service providers that followed the NICE guidelines, from 24% of English Clinical Commissioning Groups in 2013 to 12% in 2017.

1.2.3 A typical treatment cycle

Pre-treatment: consultation and tests

The recommendation for a couple to have IVF treatment comes from their GP who will refer them to a fertility specialist at an NHS hospital or fertility clinic after preliminary diagnosis (National Collaborating Centre for Women's and Children's Health (UK), 2013). At the hospital or fertility clinic, the couple will meet with a fertility specialist to discuss their fertility history and to undergo physical examinations (e.g., blood pressure, body mass index [BMI], serum hormone levels, functioning of the ovaries in females, and sperm quality in males). The IVF treatment regimen is highly individualised, and all patients have their own treatment protocols. If the specialist considers IVF the best option for the patients, and the patients accept, further examinations (e.g., blood tests for viruses and cervical screening) will take place in order to decide the treatment protocols.

Ovarian stimulation

Unlike the natural menstrual cycle in which one oocyte matures and is released from the ovaries under the effects of follicle stimulating hormone (FSH) and luteinising hormone (LH), IVF aims to produce many (5–20) oocytes within a sin-

gle menstrual cycle (Meniru, 2001) and to surgically retrieve the matured oocytes from the follicles before ovulation occurs. The development of many follicles is achieved by the administration of gonadotrophin injections. Gonadotrophin is the sex hormone that is secreted by the pituitary gland and acts on the gonads (i.e., the ovaries or testes). It regulates the secretion of FSH and LH and controls the growth of follicles.

In a natural menstrual cycle, the pituitary gland controls ovarian function and normally stimulates ovulation by suddenly increasing the production of LH towards midcycle. The sudden increase in LH, known as the LH surge, induces ovulation, when the ovarian follicle bursts and the mature oocyte is released from the ovary into the fallopian tube. In an IVF treatment cycle, the ovaries are stimulated with relatively large doses of FSH and LH, which are administered as daily injections, to enable many follicles to develop at the same time. This procedure is called ovarian stimulation. It is important to control the timing of the LH surge during ovarian stimulation to avoid a premature LH surge, which can disrupt normal follicle and oocyte development and cause the release of the oocytes before surgical retrieval, resulting in a cancelled cycle for the patient. Administration of the gonadotrophin-releasing hormone (GnRH) modulator, which affects the secretion of essential hormones of the reproductive system, including gonadotrophins, LH, and FSH, over several days, either by daily injection or nasal spray, can significantly reduce the incidence of a premature LH surge and result in more clinical pregnancies (National Collaborating Centre for Women's and Children's Health (UK), 2013). Therefore, a couple's IVF treatment journey usually starts with ovarian downregulation, which aims to suppress the pituitary's output of FSH and LH and take control of the stimulation process (Meniru, 2001).

Two main protocols (i.e. stimulation plans) are available, with different methods of ovarian stimulation. The "long protocol" with ovarian down-regulation involves the administration of hormone injections, which starts around day 21 of the

menstrual cycle preceding the treatment cycle, until the ovulation is triggered and the oocytes are retrieved. GnRH agonists are used in "long protocol" IVF to create a high output of FSH and LH and suppress the natural response of the pituitary gland (Meniru, 2001). However, the administration of GnRH agonists is associated with an increased risk of overstimulation of the ovaries (see section ovarian hyperstimulation syndrome [OHSS]), thus another treatment protocol may be applied for women with higher risks. The "short protocol" involves the administration of GnRH antagonists (a different type of medication) as hormone injections that start at day 1, 2, or 3 of a treatment cycle until the ovulation is triggered and the oocytes are retrieved.

During the stimulation phase of the treatment, women's ovarian responses are closely monitored over 3–7 days in order to (1) determine the best time for oocyte collection; and (2) reduce the incidence and severity of ovarian hyperstimulation syndrome (OHSS). The monitoring phase can last for several days and can vary slightly between individuals (National Collaborating Centre for Women's and Children's Health (UK), 2013).

Controlled ovulation

Controlled ovulation is the main event for oocyte production. The aim of this process is to take control of the natural ovulation process in order to produce a number of mature oocytes for surgical retrieval prior to fertilisation in the laboratory. Natural ovulation is triggered by the LH surge, which initiates final maturation of the oocyte and softening of the connective tissue elements that anchor the oocyte-cumulus complex to the wall of the follicle. In IVF treatment, alternative ovulation triggers are required. Human Chorionic Gonadotrophin (hCG) is a glycoprotein hormone that shares key structural similarities with LH and can induce similar oocyte maturation in developing follicles. Therefore, it is often administered as the ovulation trigger in IVF. Different types and doses of

drugs are assigned to individual patients as part of their treatment protocol. The NICE guidelines (2013) recommended that the decision of which medication to use and the starting dose should be based on individual conditions, with age, BMI, the presence of polycystic ovaries, and ovarian reserves being considered (National Collaborating Centre for Women's and Children's Health (UK), 2013). Usually, women are required to take the ovulation triggering injection approximately 36 hours prior to oocyte collection (Human Fertilisation and Embryology Authority, 2018).

oocyte and sperm collection

During oocyte collection, the mature oocytes (oocytes) are retrieved from women's ovarian follicles under direct vision laparoscopically or by ultrasound (De Los Santos et al., 2015). A transvaginal ultrasound scan is carried out and oocytes are aspirated from follicles in both ovaries through a needle that is used to pierce the vaginal wall and puncture the follicle (i.e., aspiration). Each tube of aspirated fluid is examined under magnification to identify the oocyte, which is then removed, washed, and placed in a culture dish containing culture medium for storage. The surgical procedure usually takes around 30 minutes, and the NICE guidelines (2013) recommended the use of sedation or local anaesthetic and pain relief to reduce the discomfort of the procedure for women (National Collaborating Centre for Women's and Children's Health (UK), 2013).

At about the time as oocyte collection, sperm samples are collected from the man. The semen sample is usually produced by masturbation following two to three days of abstinence from ejaculation. In cases where ejaculatory failure occurs, sperm can be collected surgically from both the epididymis and the testis. The semen sample is usually analysed and prepared before fertilisation, with the intention of removing seminal fluid, dead and abnormal sperm, and other cells from the sample, leaving only normal, motile sperm suspended in a clean culture

medium.

Fertilisation and embryo transfer

The next procedure is fertilisation, which usually happens on the day of oocyte and sperm collection. A measured volume of the prepared sperm suspension is added to the dish containing the oocytes 3–6 hours after the oocyte collection and is left for 16–18 hours to fertilise. In cases where there is a sperm defect (e.g., low sperm count, poor motility, or abnormal morphology) or surgically recovered sperm may lead to poor or failed fertilisation, ICSI may increase the chances of successful fertilisation and is recommended (National Collaborating Centre for Women's and Children's Health (UK), 2013). Instead of leaving the motile sperm in the culture medium with the oocyte, a single sperm is injected into the cytoplasm of the oocyte.

The IVF doctor and embryologist will discuss the fertilisation results and the number of embryos for transfer with the couple. Although single embryo transfer results in a 50% lower live birth rate in a fresh IVF cycle than double embryo transfer, it reduces the risk of multiple births by 400% and low birth weight by 30%, which makes the treatment safer for both the woman and the newborn (McLernon et al., 2010). Therefore, the NICE guidelines (2013) recommended the use of single embryo transfer where a top-quality embryo is available. The confirmed number of embryos will then be transferred into the woman's uterine cavity (National Collaborating Centre for Women's and Children's Health (UK), 2013).

Pregnancy test

The patients have to wait two weeks after the embryo transfer for the pregnancy test to find out if the embryo(s) has successfully implanted in the wall of the

uterus (Eugster and Vingerhoets, 1999). This is often referred to as the "two-week waiting period." After the waiting period, patients will be able to find out their pregnancy results by either a pregnancy test using a blood sample, or a pregnancy scan. If the pregnancy result is confirmed positive and the positive result persists, the fertility clinic will usually discharge the patients into antenatal care (NHS, 2018). If treatment fails, the IVF doctor will review the IVF treatment cycle and discuss options with the patients (Meniru, 2001).

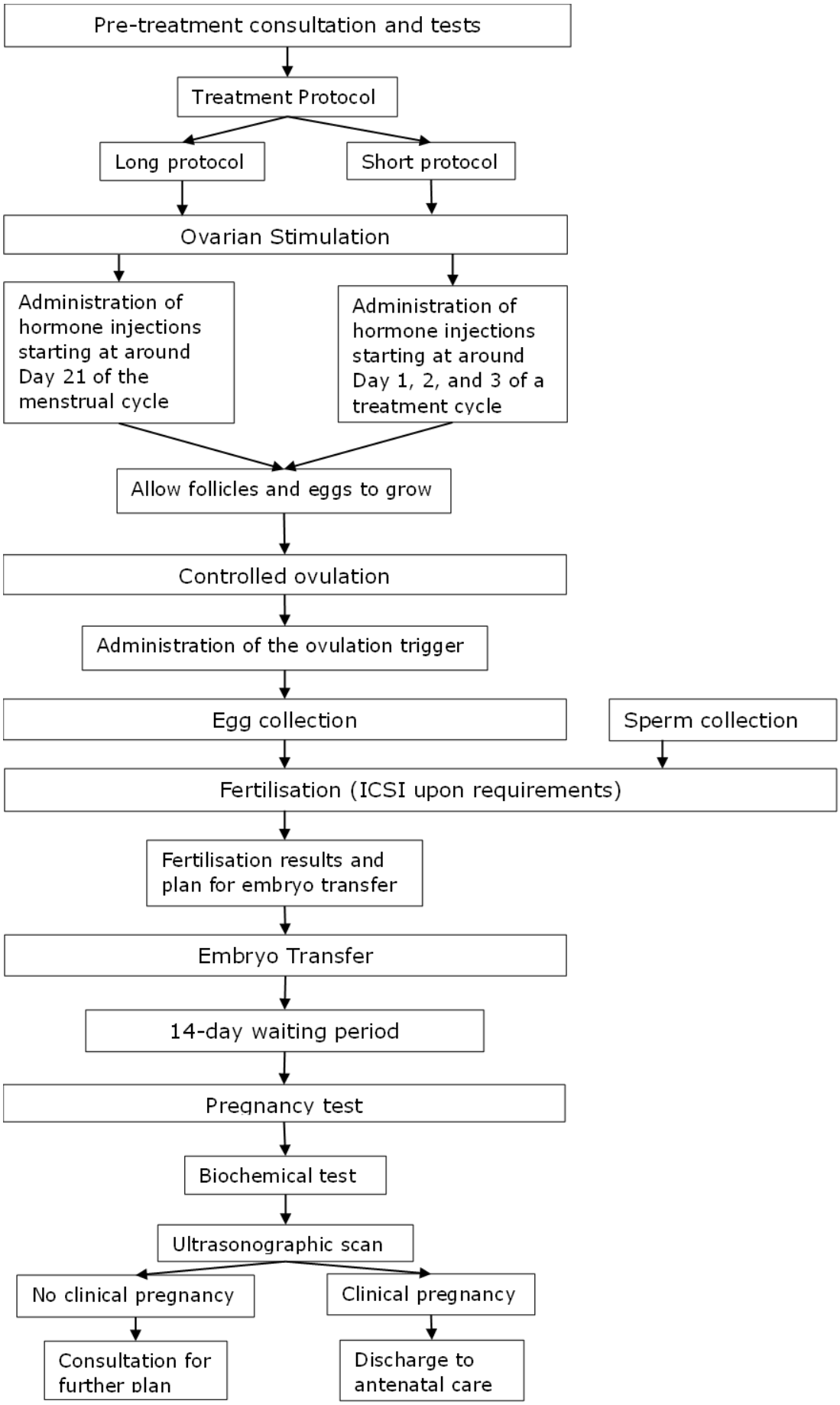


Figure 1.1 Flow of a typical cycle of IVF treatment. ICSI: intracytoplasmic semen injection

1.3 The success of an IVF cycle

1.3.1 Defining success

The WHO meeting report on the Medical, Ethical, and Social Aspects of Assisted Reproduction (2001) recommended that the principal outcome statistic of ART should be singleton and multiple live births per treatment cycle initiated. However, in daily practice, different options exist for how to define the success of an IVF cycle, from positive results on pregnancy tests to the birth of a healthy child. A paper contributed by Schieve et al. (2002) to the WHO meeting report discussed different outcomes of IVF that should be assessed and reported. Schieve et al. (2002) suggested that while assessing and reporting the outcomes of IVF treatment, both the relevance to the intended audience (e.g., patients, scientists, and clinicians) and the feasibility of the accurate collection of data should be taken into consideration. The goal of IVF for a couple is not just pregnancy but rather the birth or even the growth of a healthy child. Therefore, live birth is the more distal outcome that should be reported to patients so that they can make an informed decision about whether to undergo the treatment and what treatment plan they should follow. Live birth is defined as the birth of one or more infants that show signs of life (Schieve et al., 2002).

However, from a scientific and clinical point of view, the fundamental goal of IVF is to overcome a couple's infertility and achieve pregnancy. During treatment, the IVF physician intervenes to initiate the fertilisation and subsequent implantation of the embryos. Once pregnancy is achieved, the patients are usually discharged into antenatal care to maintain the pregnancy. Therefore, the live birth rate may not only be due to successful fertilisation and implantation but may also reflect the maintenance of the pregnancy. Assessing outcomes, such as pregnancy rate, the percentage of embryos successfully implanted in the uterus, or the percentage of embryos successfully fertilised, can provide more explicit information to allow

techniques and treatment protocols to develop (Schieve et al., 2002). In addition, it may be difficult for fertility clinics to keep track of patients after they are released from the clinic's care, which makes the feasibility and accuracy of collecting live birth data an issue. Thus, when assessing medical practice and advances in IVF, or comparing different treatment approaches, pregnancy is a proximate outcome to give an indication of the direct treatment outcome and is preferable to live births (Schieve et al., 2002). Clinical pregnancy is usually defined as the presence of a gestational sac in the uterus, which is visible on a sonogram, provides evidence of implantation, and coincides with the time a pregnancy becomes clinically recognisable instead of being purely chemically recognisable (i.e. transient increases in hCG in the woman's serum (Schieve et al., 2002)).

1.3.2 Typical success rate of IVF

For data monitoring studies on a large scale, due to incomplete data of the number of IVF cycles initiated; success rate is usually calculated based on the number of oocyte aspirations performed (i.e., the surgical procedure to retrieve oocytes) (Zegers-Hochschild et al., 2014). A global report on the utilisation, effectiveness, and safety of ART initiated by the International Committee for Monitoring Assisted Reproductive Technology calculated the success rate based on the available data from 2,560 ART clinics across 65 countries in 2011 (Adamson et al., 2018). For IVF without ICSI, the global pregnancy and delivery rate per aspiration were 24% and 17.6%, respectively. The pregnancy and delivery rate per aspiration were slightly higher for IVF with ICSI at 26.2% and 19%, respectively. In Europe, the clinical pregnancy rates per aspiration were 27.2% and 29.3% in IVF with and without ICSI, respectively, whereas the birth rates were 22.3% and 20.1%, respectively (De Geyter et al., 2018). According to the HFEA (2017), the typical live birth rate in the UK was 26% or 28% per treatment cycle, depending on whether the embryo transfer was performed soon after fertilisation (i.e., fresh

cycle), or using frozen embryos (i.e., frozen cycle), respectively. It is noteworthy that birth rates vary considerably if broken down by age (2%–29%), with women aged <35 years having the highest chance of conceiving (Human Fertilisation and Embryology Authority, 2018).

1.3.3 Risks of the treatment

As described previously, a typical IVF treatment cycle begins with ovarian stimulation and monitoring, and then progresses to oocyte retrieval, fertilisation, embryo transfer, and hopefully, implantation, pregnancy, and live birth. The optimal outcome of a treatment cycle is, of course, live birth, with both the mother and the newborn being healthy. However, a cycle may be discontinued at any of these stages for clinical reasons, such as failed oocyte retrieval or embryo transfer due to various medical complications and risks during treatment. Examples of treatment risks include multiple births, ovarian hyperstimulation syndrome (OHSS), ectopic pregnancy, and birth defects.

Multiple births

According to the HFEA, multiple births are the single biggest risk to the health of women and babies undergoing IVF. In 2008, around 25% IVF births were multiples, compared with about 2% in natural conception (Human Fertilisation and Embryology Authority, 2018). Multiple births can lead to a series of health problems for both the mother (e.g., miscarriage, pregnancy complications, postnatal infection, and bleeding) and the babies (e.g., prematurity, low birthweight, long-term disabilities, and health risks; Fauser and Edwards (2005)).

Ovarian hyperstimulation syndrome (OHSS)

Ovarian hyperstimulation syndrome is another common yet serious risk of IVF treatment: a potentially life-threatening complication of superovulation, due to the administration of ovarian stimulating medication (Claman et al., 2011). It is characterised by cystic enlargement of the ovaries and fluid leaking from the ovarian blood vessels (Kumar et al., 2011). The exact cause of OHSS is currently unknown, but it is associated with elevated responses to gonadotrophins administered during ovarian stimulation (Kumar et al., 2011). The incidence of OHSS varies between 0.6% and 10% in IVF cycles (National Collaborating Centre for Women's and Children's Health (UK), 2013). Symptoms of OHSS can range from mild abdominal swelling and discomfort to severe abdominal pain, vomiting, and dehydration. The severe form of the condition occurs in 0.5%–2% of IVF cycles (MacDougall et al., 1992). The symptoms may resolve naturally, but in severe cases, OHSS may cause treatment to be delayed or cancelled (National Collaborating Centre for Women's and Children's Health (UK), 2013).

Ectopic pregnancy

An ectopic pregnancy occurs when a foetus develops outside of the womb, usually in the fallopian tube. The incidence of ectopic pregnancy in IVF is 4.5%–11.1% in fresh treatment cycles across various studies (Dubuisson et al., 1991; Huang et al., 2014; Marcus and Brinsden, 1995). Although women with ectopic pregnancies cannot reach delivery (i.e., the pregnancy will be lost), ectopic pregnancy itself is regarded as a positive clinical pregnancy result (National Collaborating Centre for Women's and Children's Health (UK), 2013).

Neonatal risks: preterm birth, low birth weight, and birth defects

Preterm birth (PTB) and low birthweight (LBW) are recognised as the most important neonatal risks of ART treatment (Sazonova et al., 2013; Tomic and Tomic, 2011; Wen et al., 2012). Preterm birth is defined as a birth that takes place between 22 and 37 completed weeks of gestational age (Zegers-Hochschild et al., 2017). A meta-analysis revealed that PTB is 1.84 times more likely to occur in IVF singletons than in spontaneously conceived infants (McDonald et al., 2009). A similar risk of PTB was found in IVF twins when compared with spontaneously conceived twins, with a relative risk (RR) of 1.23 (McDonald et al., 2009). Low birthweight is defined as a birthweight <2,500 g (Zegers-Hochschild et al., 2017). A meta-analysis revealed an increased risk of LBW in IVF singletons (RR = 1.80; McDonald et al. (2009)) and twins (RR = 1.14; McDonald et al. (2009)). An observational study analysed neonatal outcomes of 65,868 singleton live births as results of 402,185 IVF cycles across the UK from 1991 to 2008 and found the incidence of PTB and LBW were 9.2% and 9.3%, respectively (Sunkara et al., 2015).

Birth defects are developmental disorders present at birth. A birth defect is a structural abnormality of any type, which may happen to any component of the body from cells and tissues to organs and systems (Moore et al., 2015). A meta-analysis showed that infants conceived via IVF (including ICSI) were 1.73 times more likely to suffer from birth defects across different systems (e.g., the nervous, genitourinary, digestive, and circulatory system), when compared with spontaneously conceived infants (Wen et al., 2012). However, birth defects are a controversial neonatal risk of ART, and whether the problems are caused by the IVF procedure or the infertility is not yet well understood (Hansen et al., 2005; Kamphuis et al., 2014).

1.3.4 Towards success: predictors of treatment outcomes

Studies have identified many predictors of treatment outcomes following IVF treatment, and various prediction models have been built. A systematic review of a meta-analysis performed by van Loendersloot et al. (2010) evaluated the available models and identified the most relevant predictors of success in IVF, which include female age, duration of infertility, basal FSH, and number of oocytes retrieved.

Female age

Compelling evidence shows that female age is the most influential factor in determining IVF outcomes. van Loendersloot et al. (2010) observed the same pattern in 13 studies, with female age being identified as the most important predictor of pregnancy in nearly every study included in the review (the sizes of the studies ranged from 144 to 36,412). The results of the meta-analysis based on the available data from ten studies revealed that increased female age led to decreased pregnancy rates (summary odds ratio [OR] = 0.95; van Loendersloot et al. (2010)). Several other studies have identified female age as a strong predictor of live birth following IVF treatment. For example, Roberts et al. (2010) extracted data from five IVF clinics in the UK between 2000 and 2005 and concluded that the live birth rate for IVF declined sharply from approximately 32 years of age (OR = 1.15) to a very low rate by 43–45 years (OR = 0.12). Nelson and Lawlor (2011) extracted data from the HFEA database and examined the predictors of live birth in all IVF cycles in the UK between 2003 and 2007. Similarly, their prediction model showed that the OR of live birth following IVF treatment decreased with increasing female age from 1.0 in the group 18–34 years to 0.15 in the group 45–50 years (Nelson and Lawlor, 2011). A possible explanation for the age effect may be that increasing female age is associated with decreasing quantity and quality

of oocytes (i.e., diminished ovarian reserve), which is clinically relevant in women from their mid-30s (Broekmans et al., 2007).

Duration of infertility

Increased duration of infertility is associated with a reduced likelihood of pregnancy following IVF treatment, with a summary OR = 0.99 based on two studies reporting results from 1,107 patients (van Loendersloot et al., 2010). The duration of infertility is also a predictor of live birth, even when results are adjusted for female age. Nelson and Lawlor (2011) reported that the odds of live birth decreased from 1.48 in women with <1 year of infertility to 0.71 in women with ≥12 years of infertility. Roberts et al. (2010) reported that the OR decreased from 1.31 in women with 0–2 years of infertility to 0.89 in women with ≥9 years of infertility.

Basal follicle stimulating hormone

Basal FSH, one of the sex hormones that controls the growth of follicles, is a reflection of women's ovarian reserve and is closely related to women's age (Broekmans et al., 2007). Conclusions from seven studies suggested that higher basal FSH values are associated with lower pregnancy rates after IVF treatment (OR = 0.94; van Loendersloot et al. (2010))

Number of oocytes retrieved

The number of oocytes retrieved is another predictor of pregnancy following IVF treatment. Results from six studies showed a positive association between a larger number of oocytes retrieved and greater likelihood of pregnancy following IVF

(van Loendersloot et al., 2010). A summarised OR of 1.04 for the effect of the number of oocytes retrieved on pregnancy after IVF was calculated using the available data from four studies. (van Loendersloot et al., 2010)

Chapter 2

Psychological Stress, Physiological Aspects of Stress, and Their Roles in IVF

Psychological stress is known to be related to health and disease outcomes, including, potentially, the outcomes of IVF treatment. This chapter will discuss the concepts and theories of stress, the evidence around stress in the context of infertility and IVF, and the physiological aspects of stress and how it might affect the outcomes of IVF treatment, with a focus on the 'stress hormone' cortisol.

2.1 Stress theories

2.1.1 The concept of stress and stressor

One of the pioneers in stress research Selye (1956) stated that stress is "the body's nonspecific response to a perturbation, either adaptive or not that occurs when an individual's perception of environmental demands exceeds their perceived capacity" (Selye et al., 1980). In this definition of stress, the occurrence of stress requires the presence of stressors (or stress stimuli), which have been further categorised into various types by other researchers. For example, Lazarus and Cohen (1977) suggested three types of stressors: major cataclysmic (changes affecting large numbers of people), such as war or earthquakes; major changes affecting one or a few people, such as loss or bereavement of important others or incapacitating illness; and daily hassles, such as feeling lonely or having an argument with others. Elliott and Eisdorfer (1982) categorised stressors into four different categories based on their duration: acute, time-limited stressors, such as parachute jumping; stressor sequences, such as job loss or bereavement, the results of which may last over an extended period of time; chronic intermittent stressors, such as sexual difficulties, which may occur repeatedly; and chronic stressors, such as disabilities or chronic work stress, which persist continuously over a long period. The work to understand stress by illustrating stressors and responses was referred to as the "stimulus–response approaches" (Lazarus, 1984). Limitations of such approaches include not being able to truly identify the stressor until the resultant responses have been observed, and to distinguish a stressful response from other responses (Lazarus, 1984).

2.1.2 The stress and coping theory

In 1966, Lazarus suggested that stress be treated as a systematic concept made up of many variables and processes. He introduced stress and coping theory, which emphasised the contribution of an individual's interaction with the environment (Cohen and Lazarus, 1973; Lazarus, 1975; Lazarus and Folkman, 1987). According to the transaction model, stress involves the occurrence of an event, which on its own does not necessarily induce stress, and an individual's cognitive appraisal of the event. The cognitive appraisal has two aspects: whether the event is personally significant and whether it exceeds the resources an individual has to cope. Three forms of appraisal were identified: primary, secondary, and reappraisal (Lazarus, 1984). Primary appraisal is defined as the initial appraisal of the event itself that individuals generate, which would fall into one of four categories: (1) irrelevant, (2) benign and positive, (3) harmful and a threat, or (4) harmful and a challenge. This involves the appraisal of the outside environment. Secondary appraisal involves the individual evaluating the feasibility, possible effectiveness, and consequences of their coping strategies. This involves an appraisal of the individual themselves regarding what may or can be done in the situation. Reappraisal refers to changes in the appraisal, based on new information from the environment, or the individual, or both, which may generate new emotions and behaviours (Folkman et al., 1991). The involvement of appraisal and reappraisal in stress highlights the role of individuals themselves: individuals actively interact with the external environment rather than passively responding to the stimulus. Therefore, the transaction model of stress views stress as a relationship between the person and the environment, which is constantly being appraised by the person. As a result, the responses to stressors vary between individuals (Ayers et al., 2007).

2.1.3 The allostasis model of stress

McEwen and colleagues (1988) proposed a more physiological explanation for stress by introducing the allostasis model, which suggests that stress is the disturbance of homeostasis. Homeostasis, initially introduced by Cannon (1929), refers to a steady internal state or systems that the human organism needs to survive (McEwen, 2004). Whenever changes occur in the environment that cause a mismatch between an individual's perceptions and experiences, homeostasis may be disrupted. The human organism then needs to make changes to cope with the disturbances (i.e., to respond to stressors) and regain homeostasis. This process of coping was referred to as "allostasis," which essentially refers to the systems that maintain the balance of homeostasis through constant changes (McEwen, 2005). The primary mediators of allostasis include hormones of the hypothalamic–pituitary–adrenal (HPA) axis (e.g., corticotrophin-releasing hormone and cortisol), catecholamines, and cytokines (McEwen, 2005). These mediators not only respond to challenges and demands in a stressful situation, but are also produced regularly as a basic function of bodily physiology. The effects of these mediators can be protective over the short term, under acute stress (e.g. the "fight-or-flight" response), when mediators such as glucocorticoids, are released to activate bodily responses in order to get out of a dangerous situation (McEwen, 2007). However, when allostasis is chronically imbalanced, excessive or insufficient production of mediators, for instance, chronically elevated or inadequate secretion of glucocorticoids, leads to wear and tear on the system, namely allostatic load (McEwen, 2004). Allostatic load reflects the cost to the organism of continuously adjusting to repeated or chronic stress induced by changes in the environment, which could either be physiological or psychological (McEwen, 1998). This is in line with the description of stressors in stress and coping theory (Lazarus, 1966). The effect of the allostatic load on health during chronic stress, unlike the adequate time-limited responses under acute stress, is often negative,

reflected by increased risks of diseases involving bodily systems (e.g., the cardiovascular and metabolic systems), where homeostasis is achieved through the primary mediators of allostasis, such as the hormones of the HPA axis (McEwen, 2004).

2.2 Physiology of stress

Responses to stressors can be physiological (e.g., activation of the nervous system and secretion of hormones), behavioural (e.g., smoking or drinking alcohol as coping mechanisms), cognitive (e.g., cognitive appraisal of the situation), or affective (e.g., anxiety and depression; (Ayers et al., 2007; Ogden, 2019)). To fully understand how stress influences health outcomes in this thesis, the physiology of stress and how it can be measured need to be described first. The effect of stress starts with the nervous system responding to stressors in the environment. For stress to influence bodily functions, two main bodily systems are required, namely the nervous system and the endocrine system (Brannon, 2018; Ogden, 2019). The nervous system and endocrine system work differently but in concert to form the physiological basis of allostasis (Brannon, 2018). They operate through two major pathways, namely the sympathetic–adrenal–medullary (SAM) and the HPA pathways.

2.2.1 The nervous system

The nervous system has two major divisions: the central nervous system, which consists of the brain and spinal cord, and the peripheral nervous system, which consists of all the other neurons. The peripheral nervous system is further divided into two parts: the somatic nervous system, which primarily serves the skin and the voluntary muscles, and the automatic nervous system (ANS), which primarily

serves the internal organs. The ANS is further divided into the sympathetic and parasympathetic branches. The sympathetic branch of the ANS is responsible for coping with acutely induced stress. When perceptions of stress occur, the sympathetic branch of the ANS is aroused and mobilises the body's resources to get out of the situation quickly (Brannon, 2018). This process is similar to the "fight-or-flight" response introduced by Cannon (1939) and McEwen (1998); Brannon (2018); Kemeny (2003). The parasympathetic nervous branch of the ANS is active under normal, non-stressful situations and is responsible for relaxation, digestion, and normal growth function. The sympathetic and parasympathetic nervous systems function reciprocally, with the activation of one increasing as the other is decreasing (Kemeny, 2003).

2.2.2 The neuroendocrine system

The neuroendocrine system consists of the endocrine glands distributed throughout the body, which are controlled by, and interact with, the nervous system. The neuroendocrine system functions through the secretion of hormones by the endocrine glands (e.g., the pituitary, thyroid, and adrenal glands; ovaries; and testes). During bodily responses to stress, the secreted hormones move into the bloodstream and travel to different parts of the body. Specialised receptors in target tissues or organs bind with these hormones, resulting in direct effects on the tissues or organs, or secretion of another hormone. Unlike neurotransmitters released by the nervous system, whose activation takes seconds and results in short-term effects, hormones take minutes or hours to reach the target tissues or organs and usually have prolonged effects (Brannon, 2018).

2.2.3 The Sympathetic-Adrenal-Medullary (SAM) pathway

When an individual appraises a certain event as stressful, the amygdala followed by the hypothalamus, which are both in the brain, are activated and send signals. These signals subsequently activate the sympathetic nervous system, releasing the neurotransmitter norepinephrine at various organ sites, including the adrenal medullary, to secrete epinephrine (Kemeny, 2003). Epinephrine, also known as adrenaline, a hormone, circulates immediately through the bloodstream to prepare the body for an acute response such as increasing the heart rate, increasing respiration rate, increasing blood glucose levels, and increasing circulation of blood to the exercising muscles, etc. Norepinephrine, also known as noradrenaline, is also a hormone that travels through the bloodstream and induces actions in the body, such as increasing the heart rate and blood pressure, reducing blood flow to non-exercising muscles, and increasing sweating to improve cooling (Ogden, 2019; Brannon, 2018; Cacioppo et al., 2007). This is the stress response through the SAM pathway.

Epinephrine and norepinephrine can be measured in the urine, blood, or saliva. Because of their unique roles in the SAM pathway, epinephrine and norepinephrine are indicators of the state of "fight-or-flight" (Cacioppo et al., 2007). However, epinephrine and norepinephrine have short half-life in blood (Cacioppo et al., 2007). This means that using epinephrine and norepinephrine in blood as indicators of acute stress requires rapid sample handling. However, it takes minutes for these hormones to enter the urine and approximately an hour to enter the saliva, which limits the usefulness of epinephrine and norepinephrine in reflecting the acute state of "fight-or-flight" (Cacioppo et al., 2007).

2.2.4 The Hypothalamic-Pituitary-Adrenal (HPA) pathway

The HPA axis is another physiological system involved in stress and has been the most commonly investigated pathway over the past decades (Tsigos and Chrousos, 2002). When an individual experiences a stressful event or perceives a situation as stressful, the HPA axis is activated, starting with the hypothalamus secreting corticotrophin-releasing hormone (CRH). The CRH travels to the anterior pituitary gland, which secrets a pulse of adrenocorticotrophin hormone (ACTH) in response. The adrenal gland receives the ACTH signal and produces its final hormonal product, cortisol. The function of cortisol, when released into the bloodstream in a stressful situation, is to help the body get out of the situation. For example, it can increase the utilisation of energy by actions such as promoting the conversion of protein and lipids to useable carbohydrates so that the body has enough energy to respond, decreasing inflammation by facilitating the movement of immune cells so that inflammation at a wounded area can be suppressed, and promoting the formation of memories of the event so that the individual can avoid the same situation in the future (McEwen, 2004). Deactivation of the HPA axis relies on a negative feedback system where the glucocorticoid receptors in the brain and the pituitary gland detect changes in levels of cortisol. When cortisol levels circulating in the bloodstream increase, the receptors will signal the hypothalamus to discontinue the secretion of CRH and the pituitary gland to discontinue the secretion of ACTH, which will ultimately decrease the levels of circulating cortisol and return the body to homeostasis (McEwen, 1998, 2004; O'Connor et al., 2021).

Because the HPA axis not only actively responds to stressors but also participates in regulating energy supplies for daily activities, it does not stay dormant when there is no perceived threat in the environment (Lightman, 2016). Therefore, the HPA axis regulates its hormones and follows a distinct circadian rhythm on a daily

basis. In addition to this circadian rhythm, hormones of the HPA axis are secreted in a pulsatile manner to ensure that the body is prepared to respond adequately if any stressful situations occur (Lightman, 2016). Animal studies have established the pulsatile secretion of CRH and ACTH (Ixart et al., 1991; Mershon et al., 1992; Sarnyai et al., 1995). Similar patterns of ACTH secretion have also been found in human studies (Horrocks et al., 1990). Animal studies have also shown that rats had greater ACTH and/or emotional responses to stressors when confronted with stressful events during the rising phase of hormones than during the falling phase, which supported the hypothesis that stress responses vary depending on the secretion stage of the HPA-axis hormones (Haller et al., 2000; Sarabdjitsingh et al., 2010).

As part of the HPA axis, the basal secretion of cortisol follows both the circadian rhythm and the ultradian rhythm (i.e., the pulsatile secretion pattern, see Figure 2.1). A typical 24-hour pattern of cortisol secretion is characterised by a surge in cortisol levels from the second half of the night, until a peak in the early morning, and a steady decline throughout the day, reaching its lowest levels during the first half of the night (Esposito and Bianchi, 2012; Fries et al., 2009). The most distinct aspect of the circadian rhythm of cortisol is the rapid increase in cortisol concentrations of 50%–100% within the 30–45 min after awakening (Clow et al., 2004). This phenomenon of cortisol secretion, also known as the cortisol awakening response (CAR), has only been observed in the morning as part of the process of awakening and not after short naps during the day (Federenko et al., 2004) or when sleep is disrupted by awakening in the middle of the night (Dettenborn et al., 2007). Therefore, the CAR is thought to wake the body and generate energy in the morning to prepare for the upcoming activities throughout the day (Clow et al., 2010).

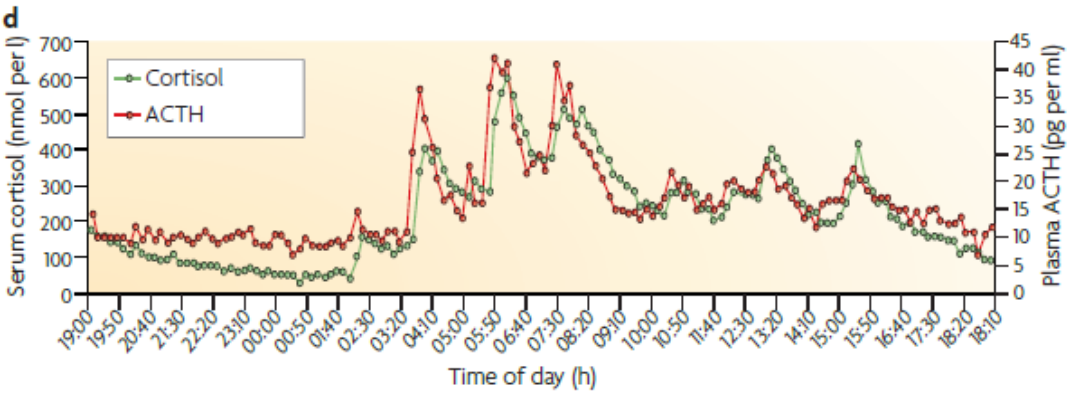


Figure 2.1 Adrenocorticotropic hormone (ACTH) and cortisol pulses in a 24-hour period. Reprinted by permission from **Springer Nature Customer Service Centre GmbH**: **Springer Nature NATURE REVIEWS NEUROSCIENCE** The crucial role of pulsatile activity of the HPA axis for continuous dynamic equilibration, Lightman and Conway-Campbell (2010)

The ultradian rhythm of cortisol secretion has been observed in animals, e.g., rats (Jasper and Engeland, 1991), sheep (Fulkerson and Tang, 1979), and monkeys (Holaday et al., 1977). A human study in which blood samples were taken automatically by a computerised system from healthy participants every 10 minutes for 24 hours showed an average number of 17.4 cortisol secretory events during the sampling period, with an average inter-pulse interval of 76.1 minutes (Henley et al., 2009). To investigate the significance of the ultradian rhythm of cortisol secretion in daily function, a randomised controlled trial on human participants blocked the natural secretion of cortisol using metyrapone and administrated a fixed daily dose of 20 mg of hydrocortisone (a cortisol replacement), using the following methods for five days: following the circadian rhythm but not ultradian rhythm (non-pulsatile), following both the circadian and ultradian rhythms, and not following either the circadian or ultradian rhythm. They examined participants' brain function using both functional magnetic resonance imaging and psychological tasks and questionnaires relating to sleep, memory, and responses to mood stimuli. The results showed that the absence of the ultradian rhythm in cortisol secretion was associated with poor working memory performance, represented by reduced accuracy and increased bias in tasks as task demands increased, and poor sleep quality (Kalafatakis et al., 2018). Prior evidence also

demonstrated the significance of the ultradian rhythm of cortisol in the stress response. For example, one study exposed rats to a mild stressor of white noise while automatically micro-sampling blood from them for 10 minutes (Windle et al., 1998) and found that physiological responses to environmental stressors at different phases of cortisol ultradian activity appeared to be different. Specifically, additional secretion of cortisol was observed immediately prior to the onset of the noise stress in the rising phase of endogenous basal cortisol, while little or no cortisol response to stressors was observed in the falling phase of endogenous basal cortisol levels (Windle et al., 1998). These results indicated that basal cortisol activity influences cortisol secretion during stress.

2.3 Stress in the context of fertility and IVF treatment

Infertility is an emotionally overwhelming and stressful condition for patients, and treatments for infertility are equally stressful (Matthiesen et al., 2011; Verhaak et al., 2007b). Notwithstanding the arguments around whether psychological stress affects the outcomes of IVF treatment, there tends to be agreement that IVF is a stressful experience for patients. In fact, evidence has shown that patients undergoing fertility treatments often experience considerable psychological stress related to their health, the uncertainty of whether the treatment will be successful, the financial implications of the treatment, or the treatment procedure itself (Redshaw et al., 2007). By researching the emotional experiences of patients undergoing IVF, healthcare professionals can better understand the patients, develop interventions to support their emotional well-being, and improve routine care. Although there is some inconsistency, evidence has shown that psychological stress may directly influence the outcomes of IVF treatment, particularly through the activity of the HPA axis. Evidence has also started

to emerge that suggests that the hormone cortisol may be a possible pathway through which psychological stress can influence pregnancy outcomes following IVF treatment. Therefore, understanding how cortisol is associated with IVF treatment outcomes could contribute to understanding the mechanism of how psychological stress influences IVF treatment outcomes.

2.3.1 Psychological stress among patients with fertility problems

Evidence from qualitative studies has illustrated the psychological burden that women with fertility difficulties experience. For example, in a qualitative study investigating the psychological consequences of infertility among 25 Iranian women with fertility difficulties, including those seeking treatment, Hasanpoor-Azghdy et al. (2014) showed that women commonly experienced fear, anxiety, worry, grief, and depression related to both the infertility itself and the fertility treatment. They also showed that these women expressed fatigue, helplessness, and hopelessness as a result of the fertility treatment process. This in-depth evidence is helpful for understanding the emotions that women with fertility difficulties experience but does not show the prevalence of psychological stress among women experiencing fertility difficulties. However, some cross-sectional studies have been able to present this evidence. For example, in a cross-sectional study involving 300 women who were diagnosed with primary infertility, Patel et al. (2016) showed that the prevalence of infertility-specific stress among women was 80%. In another cross-sectional study on 120 couples diagnosed with primary and secondary infertility, Zaidouni et al. (2018) showed that the prevalence of infertility-specific stress was 53% among women and 41% among men. These two studies were conducted on slightly different populations (i.e., women with primary infertility in the first study and couples with primary or secondary infertility in the second) and used different scales to examine stress (i.e., a clinical assessment in the first and

the Perceived Stress Scale in the second), which may account for the difference in the prevalence of stress in the results. Neither study compared the level of stress in women with fertility difficulties with population normative data, or included a comparison group of women or couples without fertility difficulties. Therefore, it is unclear whether the levels of stress among women with fertility difficulties were higher than the general population. Nevertheless, they showed that stress is a common issue among women experiencing fertility difficulties. Studies have also revealed evidence relating to the levels of anxiety and depression among women with fertility difficulties. For example, in a cross-sectional study of 225 women, including 134 women with primary infertility and 91 women without fertility difficulties, Lakatos et al. (2017) found greater anxiety and depression among women with fertility difficulties than women without and revealed that 45% of the women with fertility difficulties exhibited clinically significant depressive symptoms and 40% showed high levels of trait anxiety. They further reported that women who had previously undergone ART had significantly greater depressive symptoms then women with fertility difficulties and no fertility treatment history. However, whether the greater levels of depression in women with ART history were associated with the process of the treatment or the unfavourable outcomes of the treatment was not clear. Nevertheless, the findings may indicate that women with fertility difficulties generally experience more anxiety and depression than women without fertility difficulties, and ART may further increase the level of depressive symptoms among women with fertility difficulties.

2.3.2 Does stress influence IVF treatment?

Since IVF is an emotionally trying experience for patients, considerable research efforts have been made to understand whether stress affects IVF treatment in terms of both treatment outcomes and patient experiences. Examples of such research efforts include one systematic review investigating women's emotional

adjustment to IVF (Verhaak et al., 2007b), and three systematic reviews investigating the influence of psychological stress on IVF treatment outcomes (Boivin et al., 2011; Matthiesen et al., 2011; Nicoloro-SantaBarbara et al., 2018).

Review evidence from Verhaak et al. (2007b)

Verhaak et al. (2007b) conducted a systematic review focusing on women's emotional adjustment to IVF with regards to different phases of the treatment, including pre-treatment only, during treatment, post-treatment, and long-term post-treatment and reviewed evidence on risk factors associated with emotional adjustments to IVF e.g., personality characteristics, coping skills, and social support. The aim was to explore whether psychological stress influenced treatment outcomes of IVF; therefore, focus was placed on the pre-treatment, during treatment, and after-treatment outcomes.

For pre-treatment emotional adjustment, Verhaak et al. (2007b) reviewed nine prospective studies conducted between 1987 and 2003, including six that reported depression outcomes and nine that reported anxiety outcomes. In terms of depression outcomes, none of the six studies reported significant differences in the levels of depression between female patients undergoing IVF and population normative data or matched controls who were not undergoing any fertility treatment. The evidence for anxiety was less conclusive, with four of the nine studies reporting no differences in anxiety scores between patients undergoing IVF and population normative data and five studies reporting increased anxiety in patients undergoing IVF. One of the nine studies compared depression and anxiety levels in patients undergoing IVF with age-matched routine care gynaecology patients who were not undergoing any fertility treatment. The other eight studies compared depression and anxiety in patients undergoing IVF with population normative data, but the characteristics of the normative data were not clearly described. In addition, five of the studies measured depression or anxiety just before a treatment cycle,

one study measured depression and anxiety three months before treatment, and three studies did not specify when depression or anxiety were measured. These issues may have contributed to the mixed findings regarding anxiety. However, although it is difficult to conclude whether patients undergoing IVF exhibit increased anxiety before treatment, most of the evidence showed that they did not differ from the population normative data with respect to depression levels.

For during-cycle emotional adjustment, Verhaak et al. (2007b) reviewed five studies, of which only one measured subjective distress daily throughout a treatment cycle, and others varied in terms of when emotional adjustment was measured. Despite the heterogeneity in the study designs, these studies all showed a slight increase in distress amongst patients during the course of a treatment cycle, with oocyte retrieval day and pregnancy test day being the most stressful phases of the cycle.

For post-treatment emotional adjustment in comparisons, Verhaak et al. (2007b) reviewed six studies. Four of the studies reported an increase in depression after one or more unsuccessful treatment cycles. Two of these studies reported increased anxiety after an unsuccessful treatment cycle, but one showed no difference in anxiety pre- and post-treatment. One of the six studies also reported reduced anxiety and depression after successful treatment (Verhaak et al., 2005).

In summary, this review suggests that levels of anxiety may be only slightly higher in patients undergoing IVF at the pre-treatment stage than the population normative data, while unsuccessful treatment may increase women's levels of negative emotions (e.g., anxiety and depression).

Review evidence from Matthiesen et al. (2011)

The debate around whether psychological stress is associated with the likelihood of pregnancy following IVF is also ongoing due to the considerable heterogeneity

in the evidence. Matthiesen et al. (2011) conducted a systematic review and meta-analysis of prospective studies from 1992 to 2009, investigating the associations between stress, anxiety, and depression and various end points of ART treatment, including biochemical pregnancy, clinical pregnancy, number of oocytes, fertilisation rate, implantation rate, transfer rate, and live birth rate. For the clinical pregnancy outcome, they reviewed six studies that reported findings on stress and clinical pregnancy and found that four of these supported a significant association between greater stress and a reduced likelihood of achieving clinical pregnancy. The pooled association reached significance level; however, the effect size r (ESr) was small (ESr = 0.08, $p = 0.02$). They reviewed nine studies that investigated the association between trait anxiety and clinical pregnancy and revealed a significant pooled effect of state or trait anxiety on clinical pregnancy, with four studies supporting a significant association between greater trait anxiety and a reduced likelihood of clinical pregnancy (ESr = -0.14, $p = 0.02$). Seventeen studies investigating the association between state anxiety and clinical pregnancy were included in the review, with three studies supporting a significant association between greater state anxiety and a reduced likelihood of clinical pregnancy. The pooled association was significant but with a small ESr (ESr = 0.10, $p = 0.03$). Sixteen studies investigating the association between depression and clinical pregnancy were included in the review, with four studies reporting a significant association between greater depression and a reduced likelihood of clinical pregnancy. However, the pooled association was not significant (ESr = 0.11, $p = 0.06$ (Matthiesen et al., 2011)). They also reviewed evidence on the association between stress, anxiety, and depression and biochemical pregnancy and live birth. Only eleven studies investigated the association between stress ($k = 1$), state anxiety ($k = 5$), trait anxiety ($k = 3$), and depression ($k = 2$) and biochemical pregnancy. Three out of five studies reported significant associations between state anxiety and biochemical pregnancy, but the pooled association was not significant (ESr = -0.01, $p = 0.89$). None of the studies reported significant

associations between trait anxiety and biochemical pregnancy, and the pooled association was not significant (ESr = -0.09, $p = 0.28$). Only eight studies included in the review assessed the associations between stress ($k = 2$), state anxiety ($k = 4$), trait anxiety ($k = 1$), and depression ($k = 1$) and live birth. Two of the four studies reported significant associations between state anxiety and live birth rate, but the pooled association was not significant (ESr = -0.09, $p = 0.19$) Matthiesen et al. (2011) also noted significant between-study heterogeneity.

Review evidence from Boivin et al. (2011)

Another systematic review conducted by Boivin et al. (2011) assessed 14 prospective studies conducted between 1992 and 2006 that examined whether pre-treatment emotional distress, namely anxiety and depression, influenced the pregnancy outcomes in women undergoing ART. They defined pregnancy as being evidenced by a positive pregnancy test based on an β-hCG urine or blood test $\leq$21 days after embryo transfer (i.e., biochemical pregnancy), a positive foetal scan based on ultrasonographic visualisation of foetal heart activity (i.e., clinical pregnancy), or a live birth. Results from 14 studies showed that both the overall effect of distress (0.04, 95% confidence interval [CI]: 0.11–0.03) on the likelihood of pregnancy and the heterogeneity between studies ($I^2 = 14\%$, $p = 0.30$) were not significant. They also reported that there were no differences between subgroups in the timing of the emotional assessment (e.g., within one month of the treatment cycle or more than one month before treatment; $p = 0.66$), patients' previous experience of ART ($p = 0.66$), or whether or not the pregnant group in the studies included all patients initiating the cycle ($p = 0.39$) (Boivin et al., 2011). However, their subgroup analysis also found that of the four studies that used biochemical pregnancy as the endpoint, women who did not achieve biochemical pregnancy showed greater pre-treatment distress than women who subsequently achieved biochemical pregnancy (pooled effect = 0.42, 95%CI: 0.71–0.14, $I^2 = 0\%$, and

$p = 0.97$). However, this result was not observed in studies ($k = 10$) that used clinical pregnancy or live birth as the endpoint (pooled effect = 0.01, 95% CI: 0.09–0.06, $I^2 = 0\%$, and $p = 0.59$).

Review evidence from Nicoloro-SantaBarbara et al. (2018)

A more recent meta-analysis conducted by Nicoloro-SantaBarbara et al. (2018) assessed evidence on the association between women's anxiety, depression, and stress pre-treatment and during treatment and the likelihood of successful IVF or ICSI treatment. A total of 20 studies conducted between 1992 and 2011 were included in the meta-analysis. They defined successful treatment outcomes as pregnancy, as determined by clinical evidence of hCG levels (i.e., biochemical pregnancy), ultrasound confirmation of a heartbeat (i.e., clinical pregnancy), or live birth. Evidence from 11 studies showed that the pooled association between pre-treatment anxiety and the likelihood of treatment success was not significant (Cohen's $d = 0.16$, 95% CI: 0.33–0.02). Likewise, the pooled associations between depression ($k = 13$ studies, $d = 0.15$, 95% CI: 0.33–0.03) or stress ($k = 5$, $d = 0.13$, 95% CI: 0.47–0.21) and the likelihood of treatment success were not significant. Similarly, no significant overall effects were found for during-treatment anxiety ($k = 7$, $d = 0.10$, 95% CI: 0.38–0.18) or depression ($k = 3$, $d = 0.05$, 95% CI: 0.31–0.21) on the likelihood of treatment success. They also reported that previous fertility treatment, age, and duration of infertility were not moderators of pre-treatment anxiety, depression, or stress.

These three systematic reviews with similar aims (Boivin et al., 2011; Matthiesen et al., 2011; Nicoloro-SantaBarbara et al., 2018) seem to have reached different conclusions, and the evidence regarding whether stress affects IVF treatment outcomes seems mixed. The discrepancies could result from several issues. First, both Matthiesen et al. (2011) and Nicoloro-SantaBarbara et al. (2018) synthesised the evidence by stress, anxiety, and depression, whereas Boivin et al. (2011)

only focused on psychological distress (i.e., anxiety and depression). Matthiesen et al. (2011) reported a small but significant pooled effect for the association between anxiety and stress and clinical pregnancy, whereas the pooled association between depression and clinical pregnancy was not significant. However, Boivin et al. (2011) combined anxiety and depression to represent psychological distress in their review and did not find a significant pooled association between distress and pregnancy. Interestingly, with synthesised evidence relating to different aspects of psychological stress, Nicoloro-SantaBarbara et al. (2018) did not find any significant effect of anxiety, depression, or stress on treatment success. The reason may be that in the review, different treatment endpoints including biochemical pregnancy, clinical pregnancy, and live birth were combined to indicate treatment success.

Similarly, Boivin et al. (2011) defined successful treatment as a combination of biochemical pregnancy, clinical pregnancy, and live birth, whereas Matthiesen et al. (2011) reported the effect of stress on these different endpoints of ART treatment separately. These treatment outcomes represent different stages of the treatment (i.e., different periods of time into the treatment), thus it is possible that pre-treatment stress or distress may affect each treatment stage differently, which may have contributed to the discrepancies in the findings of the three reviews. For example, Matthiesen et al. (2011) revealed a statistically significant pooled association between trait anxiety and the clinical pregnancy rate, but the association between trait anxiety and biochemical pregnancy was not significant. Boivin et al. (2011) found a significant pooled association between distress and biochemical pregnancy but not between distress and clinical pregnancy or live birth. When combining these different treatment endpoints, neither Boivin et al. (2011) nor Nicoloro-SantaBarbara et al. (2018) found significant effects of stress or distress on treatment outcomes. This implies that the association between stress and the various treatment endpoints may be different and should therefore be investigated separately.

However, these three reviews share several limitations. Firstly, the majority of the included studies focused primarily on psychological distress, namely anxiety and depression, but only a few studies investigated the prevalence of stress or the influence of stress *per se* on ART outcomes (Matthiesen et al., 2011; Boivin et al., 2011; Nicoloro-SantaBarbara et al., 2018). Stress is a different phenomenon to psychological distress, which involves the occurrence of stressor(s), the cognitive appraisal of the stressor(s), and the resources individuals have to cope with the situation, which may lead to distress (Ebbesen et al., 2009). Therefore, the association between stress and IVF treatment outcomes should not only be determined by the association between distress and IVF treatment outcomes. Secondly, these three reviews focused on psychological stress measured by self-report tools. As discussed in the previous sections, psychobiological pathways, such as the HPA axis, and biomarkers involved in these pathways, such as cortisol secretions, are activated during bodily responses to stress. Therefore, to fully understand whether stress influences the outcomes of IVF treatment, it is critical to also consider the role of these psychobiological pathways and biomarkers in addition to assessing stress using self-reported measures with established validity and reliability.

2.4 Does cortisol influence pregnancy outcomes in IVF treatment?

2.4.1 Cortisol and health

Before looking into evidence on the role of cortisol in IVF, it is important to first address relevant evidence on the association between cortisol and health in general. As described in Section 2.2.4, cortisol is an essential glucocorticoid steroid hormone secreted by the adrenal cortex (Chourpiliadis and Aeddula,

2020). Receptors for glucocorticoids, such as cortisol, are present in almost every cell and organ in the human body (Adam et al., 2017). These receptors can act as transcription factors and can regulate gene expression in the cells, subsequently affecting the function of the organs that they are part of (de Kloet and Sarabdjitsingh, 2008). Cortisol can increase the utilisation of energy and activate anti-inflammatory responses under acute stress by binding with those receptors (de Kloet and Sarabdjitsingh, 2008; McEwen, 2004). However, during chronic exposure to stress, the prolonged secretion of cortisol and repeated activations of this process can potentially have long-lasting effects on the functioning of the organs, which contributes to a variety of adverse health outcomes (Lupien et al., 2009; McEwen, 2004).

In a meta-analysis, Adam et al. (2017) conducted a systematic review on the association between diurnal cortisol slopes and mental and physical health outcomes. The diurnal cortisol slope captures the degree of change in cortisol from morning to evening since awakening. They reviewed 80 studies involving a total of 26,167 individual participants and investigating the association between diurnal cortisol slopes and a variety of mental and physical health outcomes, including depression, anxiety, externalising symptoms, obesity, inflammatory or immune outcomes, cardiovascular diseases, fatigue, cancer, and mortality outcomes. They reported a significant overall association between a flatter diurnal cortisol slope, reflecting a disrupted circadian rhythm of cortisol, and poor health outcomes across all studies. When subtypes of health outcomes were examined, a flatter diurnal cortisol slope was significantly and independently associated with inflammation and other immune outcomes, cancer diagnosis and disease progression, fatigue, internalising and externalising symptoms, depression, and obesity. They also reported that flatter diurnal cortisol slopes predicted a significantly increased risk of mortality. However, they did not find a pooled effect of the diurnal cortisol slope on cardiovascular symptoms and diagnoses or anxiety (Adam et al., 2017).

Although this review provided relatively robust evidence that the diurnal cortisol slope is associated with a number of health outcomes, it is noteworthy that there are other indices of cortisol that are widely measured, which reflect other aspects of cortisol activities and may have different associations with the health outcomes. In fact, five different types of diurnal cortisol slope were examined by Adam et al. (2017): (1) Wake-to-bed slopes capture the absolute change or rate of change in cortisol from immediately upon waking to late evening or bedtime. (2) Peak-to-bed slopes capture the absolute change or rate of change in cortisol from the peak of the CAR to late evening or bedtime. (3) Short daytime slopes capture the slopes over a shorter period of time between waking and bedtime. (4) Fixed time point slopes capture the slopes over a given time frame (e.g., between 6 am and 6 pm), which reflects within-individual changes in the slopes. (5) Amplitude measures reflect the degree of variation in the diurnal cortisol rhythm from intensive repeated measures of cortisol values throughout the day. By combining all the physical health outcomes and only looking at the associations by slope type, Adam et al. (2017) found different effect sizes for the association between different slope types and health outcomes. The five studies using amplitude measures showed stronger associations (larger effect sizes) with a variety of health outcomes (e.g., major depressive disorder, allergic rhinitis, and extensive nasal polyposis); the fixed time point measures ($k = 27$), wake-to-bed slopes ($k = 115$), and peak-to-bed slopes ($k = 23$) showed mild to moderate effect sizes; and the short daytime slopes ($k = 9$) did not show significant effect sizes. However, Adam et al. (2017) did not conduct further subgroup analysis by health outcomes to show whether these differences in effect sizes would further differ by health outcomes. Nevertheless, these findings suggest that while investigating the association between cortisol and health outcomes (either combined health outcomes or a single health outcome), it is critical to consider the heterogeneity of cortisol indices and the measures associated with them.

2.4.2 Cortisol: fractions, indices and measures

Free and bound cortisol

Cortisol circulates in the blood both freely and in bound forms (bound to binding proteins, such as corticosteroid-binding globulin [CBG], sex hormone-binding globulin, and albumin) (Levine et al., 2007). Total cortisol (the sum of free and bound cortisol) reflects the hormonal status of the subject with combined information of both cortisol and the CBG (Levine et al., 2007). Free cortisol is the fraction of cortisol that is not bound to binding proteins (Turpeinen and Hämäläinen, 2013) and is crucial for the consequent signalling response (Lightman and Conway-Campbell, 2010). The reason for this is most likely because only free cortisol is biologically active and able to move out of capillaries and into cells (Levine et al., 2007), according to the free hormone hypothesis (Mendel et al., 1989). Therefore, the effect of cortisol on the target tissues is related to the free bioavailable fraction rather than the level of total cortisol in circulating blood (Lightman and Conway-Campbell, 2010). However, in human plasma, approximately 80%–90% of cortisol is bound to CBG and 10%–15% is bound to albumin, leaving approximately 5% unbound cortisol in the blood (Lewis et al., 2005). Corticosteroid-binding globulin is saturated at 400–500 nmol of cortisol, which is approximately the peak concentration of each pulse; therefore, plasma cortisol levels above this concentration will result in higher levels of free cortisol available for diffusion across cell membranes (Cameron et al., 2010). It is generally assumed that cortisol measured in saliva is also free cortisol (Kirschbaum and Hellhammer, 1994) but the concentration of cortisol in saliva is less than 10% of that in blood (Levine et al., 2007).

Cortisol indices and how they are measured

Different cortisol indices can reflect different characteristics of cortisol activity and are often examined through different sampling methods. Commonly examined cortisol indices include the momentary reactivity of cortisol, CAR, area under the curve (AUC) with respect to ground (AUCg) and increase (AUCi), diurnal cortisol slope (diurnal slope), cortisol levels at a single time point, and cortisol levels over a longer period.

The momentary cortisol reactivity reflects individuals' cortisol reactivity to laboratory-induced acute stress or everyday situations. It is typically used to explore the intra-individual relationship between cortisol, mood, and experiences (Saxbe, 2008). Immediate blood or saliva samples are typically taken to capture the momentary cortisol activity. As previously described, CAR represents the start of daily cortisol activity and is a distinct phenomenon of the circadian rhythm of cortisol (Lightman and Conway-Campbell, 2010). Due to the sharp rise in cortisol levels within an hour of awakening, CAR is usually measured in blood or saliva, and samples are typically collected immediately upon awakening and 15–45 minutes after awakening. The former is to reflect the pre-awakening cortisol secretion, and the latter is to inform the characteristics of the increase (Clow et al., 2010). Some researchers have suggested that the measure of CAR should also include an estimate of overall cortisol secretory activity, which is usually captured by the AUC. The AUC is widely used to determine the overall secretion of cortisol over a particular time frame (Vedhara et al., 2006). Two further formulae for the AUC, namely AUCg and AUCi, were proposed to capture not only the overall hormone output but also the reactivity of the system (Pruessner et al., 2003). The AUCg indicates the area between ground and cortisol values taken across the day. The AUCi, which is less frequently used by researchers, stands for the area between the increase and cortisol values taken throughout the day (Pruessner et al., 2003). Both AUCg and AUCi can be measured in blood

or saliva by collecting one sample immediately at awakening and several samples throughout the day (Vedhara et al., 2006). The diurnal cortisol slope captures the line resulting from the regression of cortisol values collected across the day onto hours since awakening, excluding the morning awakening response. Other single-time-point indices, for example waking cortisol (i.e., cortisol levels immediately upon waking), bedtime cortisol (i.e., cortisol levels at the end of the day when levels approach a nadir), and resting cortisol (i.e., random, non-fasting cortisol in the morning, excluding CAR) are also used in many studies (Saxbe, 2008; Rector et al., 2019). The diurnal cortisol slopes are usually measured by collecting multiple blood or saliva samples throughout the day (Adam et al., 2017). The aforementioned cortisol indices are often examined in blood and saliva samples. Cortisol levels over a longer period of time are usually examined in urine or hair samples. One urine sample can capture the accumulated cortisol levels since the last urination, thus collecting urine samples for 12–24 hours can reflect the total cortisol levels over that time. When collected for the entire 24 hours, urine samples can be assayed for cortisol levels that are not affected by the circadian rhythm of cortisol secretion (Turpeinen and Hämäläinen, 2013). Hair samples can capture cortisol levels over a longer time frame. It was originally hypothesised that circulating cortisol is incorporated into the hair shaft during its formation (Meyer and Novak, 2012), and because human hair grows approximately 1 cm per month, a hair sample of 1 cm would reflect the cortisol levels over the past month, thus hair cortisol is a promising method to investigate long-term levels of cortisol (Saxbe, 2008). However, a meta-analysis of stress-related and basic determinants of hair cortisol in humans reported a 29% decline in hair cortisol concentrations (HCCs) from the first proximal 3 cm hair segment to the second proximal 3 cm hair segment (Stalder et al., 2017). It is now widely agreed that the first 5–6 cm of hair can reliably reflect the accumulated cortisol levels over the past 5–6 months (Russell et al., 2012). This gives researchers the advantage of being able to measure the systematic levels of cortisol over several months by

non-invasively collecting a single hair sample. It is now recognised as a useful tool to examine the potential mechanism through which chronic stress influences health outcomes (Staufenbiel et al., 2013; Russell et al., 2012).

Table 2.1 Summary of the characteristics and measures of cortisol indices

Cortisol Indices	Sample	Time frame captured	Feature of cortisol secretion reflected	Demand for participants
The momentary reactivity	Blood, Saliva	Immediate reactivity	Changes in cortisol over hours, days, or weeks within individuals.	Repeated sampling of both cortisol and experience measures for several days
CAR	Blood, Saliva	0-45min upon awakening	The start of the circadian rhythm of cortisol	One sample immediately upon awakening, other samples during 15-45min after awakening
AUC_g	Blood, Saliva	Typically throughout a day since awakening	The total hormonal output or basal activity of the HPA axis.	One sample immediately upon awakening, other samples throughout the day
AUC_i	Blood, Saliva	Typically throughout a day since awakening	Reactivity of the HPA axis across the day	One sample immediately upon awakening, other samples throughout the day
Diurnal slope	Blood, Saliva	Throughout a day since awakening	Circadian fluctuations of cortisol across the day	At least 4 to 6 saliva samples across each day for at least 2 or 3 days
Cortisol at a single time point	Blood, Saliva	Specific time points	Cortisol levels at a specific time point	One sample at a specified time point
12/24h Urinary cortisol	Urine	12/24h	Cortisol secretion over the period of urinary sampling	Urine sample from each urinary void over the sampling period
Hair cortisol concentration	Hair	Systematic cortisol levels for up to 6 months per sample	Cortisol levels over the past 1-6 months from the sampling time point	One sample taken from the root end of hair

Note. CAR: cortisol awakening response, AUC_g: area under the curve with respect to ground, AUC_i: area under the curve with respect to increase, HPA axis: hypothalamic-pituitary-adrenal axis

2.4.3 The association of cortisol and IVF treatment: physiological pathway

As mentioned in Chapter One, Section 1.2.3, the female reproductive system functions under the regulation of the hypothalamic–pituitary–gonadal (HPG) axis. The hypothalamus secretes GnRH, which travels to the pituitary to stimulate the release of FSH and LH, which then stimulate the secretion of oestrogen and progesterone from the ovaries. A feedback loop exists by which oestradiol exerts negative feedback on both the hypothalamus and pituitary and inhibits the secretion of GnRH, and subsequently FSH and LH production. Glucocorticoids, such as cortisol, can directly inhibit the release of GnRH from the hypothalamus and the release of gonadotrophins from the pituitary (Whirledge and Cidlowski, 2017). Through this pathway, elevated cortisol levels could lead to reproductive dysfunction. This pathway has been studied in animals. For example, in a study on rhesus monkeys, Hayashi and Moberg (1989) showed that increasing cortisol levels suppressed the follicular rise of oestrogen, and subsequently the surge of LH and FSH and the luteal release of progesterone. Breen and Karsch (2004) reported that cortisol suppressed pulsatile LH secretion by inhibiting the responsiveness of the pituitary of sheep to GnRH. In addition to acting on the hypothalamus and pituitary, glucocorticoids can also directly bind with their receptors in ovarian cells, hence directly affecting ovarian function by suppressing the secretion of LH and FSH (Whirledge and Cidlowski, 2010). In an animal study, Yuan et al. (2016) reported that female mice exhibited significantly decreased oocyte developmental potential after cortisol injections.

As it regulates the reproductive system, the success of IVF treatment depends on the functioning of the HPG axis, and dysfunction of the HPG axis could lead to treatment failure. Because cortisol can directly affect the functioning of the HPG axis, it can theoretically affect the success of an IVF cycle.

2.4.4 The association of cortisol and IVF treatment: systematic review by Massey et al. (2014)

Although there is relatively robust support from theories and laboratory data that cortisol may influence the outcomes of IVF treatment, the observational evidence is less clear. In a systematic review of the association between cortisol and IVF treatment outcomes, including clinical pregnancy, oocyte number, fertilisation rates, cleavage, and miscarriage rates, Massey et al. (2014) reviewed 16 studies from eight countries, with a total sample size of 1,647 female patients. Of a total of 12 studies investigating the association between cortisol and clinical pregnancy, three studies reported higher levels of cortisol in patients who achieved clinical pregnancy, four studies observed higher cortisol levels in patients who failed to achieve pregnancy, and five studies reported no difference in cortisol levels between the pregnant and non-pregnant groups. The three studies reporting a positive association between cortisol and clinical pregnancy took either one blood or follicular fluid sample, or both from participants around the time of oocyte retrieval or recovery. Among the four studies reporting a negative association between cortisol and clinical pregnancy, one collected seven blood samples from pre-treatment to post-embryo-transfer, one collected one blood and one follicular fluid sample at pre-treatment and at oocyte retrieval, one collected four blood samples from pre-treatment to pregnancy test and gestation following a successful pregnancy, and one did not describe details of the cortisol samples. Among the five studies reporting a null association between cortisol and clinical pregnancy, one took 16 blood samples throughout a natural menstrual cycle, one took a blood and a follicular fluid sample at oocyte retrieval, one took 24-h urine samples on the day after gonadotrophin administration, one took three urine samples at pre-treatment, oocyte retrieval, and embryo transfer, one study took six saliva sample during day 1-3 of the treatment cycle. Of the three studies investigating the association between cortisol and oocyte numbers, only one found

that lower levels of cortisol were associated with an increased number of oocytes retrieved (seven blood samples were collected from pre-treatment to post-embryo-transfer), while the other two did not find any significant association (one took a blood and a follicular fluid sample at oocyte recovery and one took six saliva samples during day 1-3 of the treatment cycle) . Of the four studies investigating the association between cortisol and the fertilisation rate, two reported that higher cortisol levels were associated with oocytes that did not fertilise (one collected seven blood samples and one collected a blood and a follicular fluid sample, both measured the time-frame from pre-treatment to embryo-transfer), one reported the opposite (one follicular fluid sample between ovarian stimulation and oocyte retrieval), and one found no association (one follicular fluid and one blood sample at oocyte retrieval). Of the three studies that examined cortisol and oocyte cleavage, one found that lower cortisol levels were associated with oocytes that cleaved (seven blood samples collected from pre-treatment to post-embryo-transfer) and two found no association (one collected both blood and follicular fluid samples at pre-treatment, and between oocyte retrieval and embryo transfer during treatment, one collected both blood and follicular fluid samples at oocyte recovery). Of the two studies that examined cortisol levels and oocyte maturity, one found an association between higher cortisol levels and follicles containing mature oocytes (one blood and one follicular fluid sample collected at pre-treament and between oocyte retrieval and embryo transfer during treatment), while the other one did not find a significant association (one blood and one follicular fluid sample collected at oocyte recovery). One study examined cortisol and miscarriage rates in IVF and did not find a significant association. This study collected three saliva and three blood samples post-embryo-transfer and in early pregnancy. Massey et al. (2014) reported that the quality of available evidence was poor, with 11 out of 16 studies being considered low quality. They also addressed a number of methodological limitations in the current studies investigating the influence of cortisol on IVF treatment outcomes. First, the reliance on acute (saliva or blood)

or short-term (urine or follicular fluid) measures of cortisol, which can only provide information about cortisol levels over minutes (saliva or blood) and hours (urine or follicular fluid), which does not reflect the effects of exposure to stress over the long term. Second, the majority of studies failed to control for variables known to influence cortisol levels, for instance, time of day, caffeine consumption, and BMI. Third, the studies measured cortisol at different stages of IVF treatment, which may lead to different study findings. These three methodological factors may together have contributed to the equivocal study findings. Massey et al. (2014) concluded that future research would benefit from measuring cortisol over longer periods and prior to commencing treatment.

2.4.5 The role of hair cortisol in pregnancy in IVF treatment: Massey et al. (2016)

Following their systematic review of the role of cortisol in IVF treatment, Massey et al. (2016) conducted a cohort study to explore the relationship between acute (saliva) and chronic (hair) levels of cortisol measured in the 1–2 weeks prior to commencing IVF and the likelihood of achieving clinical pregnancy (defined as a pregnancy diagnosed by ultrasonographic visualisation of foetal heart activity at six weeks post embryo transfer). A total of 118 female patients were recruited into their study from a single English fertility clinic. The inclusion criteria included non-smoking women with a BMI of 19–35 undergoing a long GnRH agonist protocol. To determine HCCs over the 3 months prior to treatment, participants were instructed to take a 3cm hair sample close to the scalp from the posterior vertex of the head within 1–2 weeks prior to commencing treatment. To determine salivary cortisol, all participants were asked to collect saliva samples upon awakening, 30 min after awakening, and at 10 pm each day for two consecutive days immediately prior to commencing downregulation medications. The complete data (n = 135) were analysed for different cortisol indices, including CAR,

AUC_g, AUC_i, diurnal slope, and HCC (n = 88). The results showed that after controlling for covariates of treatment outcomes and cortisol, namely age, number of oocytes, fertilisation rate, BMI, and the season of hair sampling, CAR (p = 0.485), AUC_g (p = 0.527), AUC_i (p = 0.731), and diurnal slope (p = 0.889) were not significantly associated with clinical pregnancy, whereas higher HCCs were significantly associated with a reduced likelihood of clinical pregnancy (p = 0.017). Post-hoc analyses revealed that there was no significant relationship between salivary cortisol measures and clinical pregnancy in the subsample who provided both saliva and hair samples (n = 88). Higher HCCs were still significantly associated with a reduced likelihood of clinical pregnancy (p = 0.034), accounting for between 26.7% (Cox and Snell) and 36.4% (Nagelkerke R^2) of the variance in pregnancy outcomes after accounting for the variance explained by salivary cortisol (Massey et al., 2016). This was the first study to explore the effect of both acute (saliva) and chronic (hair) cortisol on the clinical outcomes of IVF treatment, and the first study to show a significant association between higher levels of hair cortisol and a reduced likelihood of clinical pregnancy, as the outcome of IVF treatment. This suggests a new angle to investigate the influence of stress on IVF treatment outcomes by examining cortisol.

However, despite its unique contributions, this cohort study had several methodological limitations. First, it was conducted on a small sample (n = 88), which was considered the minimum standard for assessing cortisol concentrations in such research (Massey et al., 2016). Second, several potential covariates (e.g., hair characteristics such as natural hair colour and frequency of hair washing) of hair cortisol (Wosu et al., 2013) were not accounted for. These factors may impact the generalisability of the study findings. Therefore, replication of this study with a larger cohort and more detailed investigations of possible covariates is required. Third, the probable effects of lifestyle factors that may be related to both cortisol and pregnancy results (e.g., sleep quality, physical activity, and diet) were not investigated in the study. Lifestyle factors are modifiable contributors

to fertility and outcomes of IVF treatment (Homan et al., 2007); therefore, it is important and necessary to consider the possible role of lifestyle factors when investigating the effect of stress and cortisol on IVF treatment outcomes.

2.4.6 Recent evidence of the association of hair cortisol with IVF treatment outcomes

Since the publication of the study by Massey et al. (2016), a few new studies have updated the literature on the association between hair cortisol and IVF treatment outcomes. In a prospective study of 43 female patients undergoing IVF or ICSI, C Santa-Cruz et al. (2020) collected data at two time points: at the consultation with the reproductive endocrinologist before commencing the IVF cycle (T1) and after embryo transfer but before the hCG pregnancy test (T2, approximately 12 weeks after T1). None of the patients were aware of their pregnancy (biochemical) outcomes at T2. Participants completed self-report questionnaires to assess anxiety, depression, and resilience at T1 and T2. Participants were also asked to provide a 3 cm hair sample from close to the scalp at T1 and T2 to assess HCCs. The first hair sample (T1) was to measure HCC during the three months prior to starting IVF treatment. The second hair sample (T2) was used to assess HCC during the three months (12 weeks) from starting the treatment cycle to after the embryo transfer. C Santa-Cruz et al. (2020) found that HCCs significantly increased from T1 to T2 in all patients. This increase was also found separately in women who achieved biochemical pregnancy and those who did not. However, HCCs were not significantly different between women who achieved biochemical pregnancy and those who did not at either T1 or T2. C Santa-Cruz et al. (2020) also conducted logistic regression to assess the association between HCC and the likelihood of a positive biochemical pregnancy test. The results showed that HCCs at T2 significantly predicted 46% of the positive pregnancy tests ($\beta = 0.11$, $p < 0.05$). With regard to the self-reported anxiety, depression, and resilience,

C Santa-Cruz et al. (2020) found that women who did not achieve biochemical pregnancy had significantly higher levels of depression at T2 when compared with women who achieved pregnancy ($t = 3.13$, $p = 0.01$). Women who achieved biochemical pregnancy also had significantly higher levels of resilience at T2 when compared with those who did not achieve biochemical pregnancy ($t = 4.89$, $p = 0.01$), and resilience at T2 predicted 34% of the positive pregnancy tests ($\beta = 0.10$, $p < 0.05$). However, there was no significant difference in the levels of anxiety at T2 between women who achieved biochemical pregnancy and those who did not. The between-group differences in anxiety, depression, and resilience were also not significant at T1. Both this study and the study conducted by Massey et al. (2016) showed a significant correlation between HCC and IVF treatment outcomes but at different stages of treatment (i.e., during treatment and pre-treatment, respectively) and with different treatment outcomes (i.e., biochemical pregnancy and clinical pregnancy, respectively). These findings together support a relationship between HCC and IVF treatment outcomes.

Recent evidence of the association between other cortisol indices and IVF treatment outcomes

Several other recent observational studies have also investigated the associations between stress, cortisol, and IVF treatment outcomes using other cortisol indices. For example, Cesta et al. (2018) conducted a study on 485 female patients undergoing IVF. Self-reported perceived stress, infertility-specific stress, and salivary cortisol were measured before patients started their IVF cycles. Both morning and evening salivary cortisol were measured. They reported that neither the self-reported stress nor the salivary cortisol was significantly associated with the likelihood of subsequently achieving oocyte aspiration, embryo transfer, or clinical pregnancy after embryo transfer. Self-reported stress and cortisol were also not significantly associated with the oocyte maturity rate or the embryo utilisa-

tion rate. However, they found a significant association between increased levels of evening cortisol and decreased average quality of the embryos.

Miller et al. (2019) conducted a study on 72 female patients undergoing IVF. Self-reported stress and anxiety and salivary cortisol were measured at three time points: pre-treatment, on the day of oocyte retrieval, and on the day prior to embryo transfer. Follicular cortisol concentrations were also measured on the day of oocyte retrieval. With regard to the self-reported anxiety and stress, they reported no significant difference in anxiety scores of women pre-treatment, on the oocyte retrieval day, and on the day of embryo transfer. However, self-reported stress significantly increased from pre-treatment to oocyte retrieval day and showed a non-significant decreasing trend by embryo transfer day. No significant differences in self-reported anxiety and stress were found between women who subsequently underwent embryo transfer and those who did not. No significant differences in self-reported anxiety and stress were found between women who subsequently became biochemically pregnant and those who did not. Women undergoing IVF for the first time and those who had previous IVF treatment also did not differ in self-reported anxiety and stress. With regards to the hormonal results, this study reported that salivary cortisol levels significantly increased from pre-treatment to its highest level on the oocyte retrieval day and then significantly decreased by the embryo transfer day. Higher cortisol levels were found in women who were undergoing their first cycle than women who had previously undergone IVF treatment. However, the difference was only significant on the embryo transfer day. They further examined the relationship between self-reported anxiety, stress, cortisol, and IVF treatment outcomes. No significant difference in cortisol levels were found between women who subsequently underwent embryo transfer and those who did not or between women who subsequently became biochemically pregnant and those who did not. However, they found a significant positive correlation between follicular cortisol concentrations measured on the oocyte retrieval day and the fertilisation rate. They reported no significant association

between salivary cortisol and self-reported anxiety or stress at any time point.

Both studies showed that pre-treatment, self-reported stress, and anxiety were not associated with IVF treatment outcomes. It is difficult to synthesise other findings from these studies because they had slightly different treatment endpoints (e.g., clinical pregnancy in Cesta et al. (2018) and biochemical pregnancy in Miller et al. (2019)), and measured different biomarkers of stress (e.g., salivary evening cortisol in Cesta et al. (2018) and follicular cortisol in Miller et al. (2019)). However, Miller et al. (2019) reported significant increases in cortisol levels during treatment and Cesta et al. (2018) suggested that elevated pre-treatment evening cortisol levels were associated with lower embryo quality. These results again suggest that cortisol may be associated with IVF treatment outcomes.

In summary, the influence of psychological stress on health is well discussed in theory, including the potential mechanisms of such effects. The HPA axis is one of the most researched mechanisms through which psychological stress may affect physical health. However, the evidence of whether psychological stress plays a role in achieving favourable outcomes of IVF treatment (e.g., clinical pregnancy) is still inclusive. This is particularly shown by the different conclusions from three systematic reviews investigating the influence of psychological stress on IVF treatment outcomes (Matthiesen et al., 2011; Boivin et al., 2011; Nicoloro-SantaBarbara et al., 2018). One outstanding limitation of the studies examined by these reviews is the lack of evidence on the potential pathways through which psychological stress may influence IVF treatment outcomes. Efforts have been made to resolve this issue by examining the role of the HPA axis, in particular its final product the hormone cortisol, in IVF treatment. However, due to the heterogeneity in the type of cortisol samples (e.g., blood, saliva, urine) and indices (e.g., CAR, AUC, diurnal slop) examined across studies, a conclusive conclusion was not reached. Traditional methods of measuring cortisol (e.g., in blood, saliva, urine, and follicular fluid) are limited by the circadian and ultradian nature of

cortisol activities (Saxbe, 2008) and may not well indicate the long-term effect of stress (Russell et al., 2012; Massey et al., 2014). The approach of using hair samples, however, might provide more insights into the association between long-term levels of cortisol and IVF treatment outcomes that could not previously be answered by other approaches alone (Massey et al., 2014, 2016).

Chapter 3

The Role of Lifestyle

Lifestyle factors, such as physical activity, sleep, and diet, are known to affect health and wellbeing in humans. Fertility, as an aspect of health, is influenced by a variety of lifestyle factors. *In vitro* fertilisation, as a fertility treatment, can also be influenced by these factors. This chapter will describe current evidence on a variety of lifestyle factors that are believed to play important roles in fertility, infertility, and IVF treatment.

3.1 The role of lifestyle in fertility, infertility, and IVF treatment

3.1.1 Behavioural determinants of reproductive health

Lifestyles are patterns of behavioural choices in day-to-day living (Dean, 1989; Kickbusch, 1986). Diet, physical activity, sleep, alcohol consumption, and smoking are all common examples and elements of lifestyle choices. Individuals make their behavioural choices from the available alternatives under the influence of specific social, cultural, psychological, and economic factors (Organization et al., 1986). Although the options available to individuals are limited by their socioeconomic circumstances and cultural environment (Organization et al., 1986), lifestyles are formed by individual choices and are amenable to change (Kickbusch, 1986; Olsen, 2017). The individual involvement in lifestyle has made it a behavioural determinant of health that is associated with morbidity and mortality (Dean, 1989; Olsen, 2017). Reproductive health, as an essential component of health, is influenced by lifestyle in many ways. For example, lifestyle factors may affect male fertility through semen quality, female fertility through ovulation, and both male and female fertility through sexual behaviours and hormonal functions (Homan et al., 2007; Watson, 2015). Understanding the influence of specific lifestyle factors on fertility and the underlying mechanisms can provide crucial information for individuals to make healthier decisions and obtain optimal reproductive outcomes (Watson, 2015). Evidence has suggested that lifestyle factors may also influence fertility treatment outcomes and that cortisol may be related to lifestyle factors (e.g., sleep, physical activity, and diet) (Fukuda and Morimoto, 2001). Therefore, it is important to include lifestyle factors in this thesis while investigating the association between stress and IVF outcomes. The association between lifestyle and cortisol will be discussed in Chapter 4, and the role of lifestyle factors in IVF treatment will be discussed in this chapter.

3.2 Lifestyle behaviours among patients undergoing IVF treatment

Before discussing the possible roles of a range of lifestyle factors, such as physical activity and sleep, in IVF treatment, it is important to first understand patients' beliefs about which lifestyle factors may affect their fertility and the lifestyle choices they make in preparation for IVF treatment. Several studies have investigated patients' understanding of the importance and/or the influence of lifestyle choices (Hawkins et al., 2014; Rossi et al., 2016). In a cross-sectional study on 138 couples undergoing IVF treatment Hawkins et al. (2014) found that the majority of patients undergoing IVF (both male and female) believed exercising, avoiding stress, and maintaining a normal weight were beneficial, and yet surprisingly, 63% of patients believed in the benefit of rigorous exercise prior to transfer and complete bedrest following the transfer. They also observed that patients with higher levels of income and education were less likely to consider lifestyle influential (Hawkins et al., 2014). In another observational study on 208 patients undergoing IVF, Rossi et al. (2016) investigated patients' beliefs regarding lifestyle factors and IVF outcomes. They observed that the majority of patients considered women's lifestyle choices influential (e.g., stopping smoking [72%], drinking alcohol [69%], consuming caffeine [62%], and taking vitamins [88%]). However, a key limitation of these studies was that they only examined women's and men's views rather than their actual behaviours.

The above evidence suggests that patients undergoing IVF treatment do believe that certain lifestyle factors (e.g., physical activity and smoking) may influence the likelihood of treatment success, and some patients may change their behaviours in preparation for their treatment. In order to give appropriate advice to patients about whether they should adjust their lifestyle and what lifestyle factors they should adjust, it is important to consider the evidence on the association

between a variety of lifestyle factors, fertility, and IVF outcomes.

3.2.1 Diet and body mass index

Diet is the most important lifestyle attribute to health (Katz and Meller, 2014). Unhealthy diets contribute to malnutrition or obesity and are directly related to diseases such as heart disease, cancer, and diabetes (Who and Consultation, 2003) through an individual's weight (Ogden, 2019). In terms of reproductive health, extremes of body weight, either being obese or underweight, can disrupt the hormonal balance and lead to ovulatory dysfunction in women and abnormal semen in men (Sermondade et al., 2013; Watson, 2015).

Body mass index is a reflection of an individual's body weight and is calculated as follows:

$$\text{BMI} = \text{weight in kilograms} \div \text{height in meters}^2$$

Although BMI cannot definitively indicate whether a person is underweight, overweight, or of normal weight, it is the most useful indicator of a person's obesity status (Everitt, 2016). According to the WHO, for most adults, a healthy BMI is within the range 18.8–24.9. A BMI < 18.5 indicates underweight, 25–29.9 indicates overweight, and ≥ 30 usually indicates obesity (Who and Consultation, 2003). Compelling evidence has shown that BMI is closely related to disease, morbidity, and mortality. For example, in a systematic review of 97 studies, Flegal et al. (2013) suggested that a BMI at the level of obesity is significantly associated with higher all-cause mortality. The majority of cases of unhealthy weight are caused by an unhealthy lifestyle (Everitt, 2016). Therefore, as the major indicator of weight, BMI is a measure of lifestyle that reflects other lifestyle factors, such as diet and physical activity (Farhud, 2015; Who and Consultation, 2003).

Body mass index is a well-studied predictor of fertility. It has been suggested that the association between BMI and the risks of infertility follows a J shape (Bolumar et al., 2000). Studies have revealed an association between increases in BMI and higher risks of infertility (Rich-Edwards et al., 1994), higher risks of delayed conception (Hassan and Killick, 2004; Bolumar et al., 2000), and a reduction in sperm quality (Jensen et al., 2004). Being underweight is a risk factor for menstrual dysfunction and amenorrhoea, which are related to fertility difficulties in women (Knuth et al., 1977; Bates et al., 1982; Van der Spuy et al., 1988). The influence of BMI on infertility treatment has often been shown by changes in weight. For example, weight-loss programs were found helpful in achieving better results for overweight patients with fertility difficulties who were undergoing fertility treatments (Clark et al., 2000). Body mass index also plays an important role in IVF treatment. According to National Collaborating Centre for Women's and Children's Health (UK) (2013), a healthy BMI sometimes forms part of the eligibility criteria for NHS-funded IVF treatment. With regard to research evidence, in a systematic review based on a meta-analysis of the findings of 33 studies covering 47,967 treatment cycles, Rittenberg et al. (2011) found that women with BMI $\geq$ 25–29.9 had significantly lower clinical pregnancy and live birth rates and significantly higher miscarriage rates following IVF than women with a BMI $<$ 25.

The mechanism of the influence of BMI on IVF may be through the regulation of hormones. For example, a higher BMI is associated with increased fasting insulin concentrations, insulin resistance, increased levels of testosterone, and decreased sex hormone-binding globulin (Balen et al., 2006). Because sex hormone-binding globulin binds with androgens, such as testosterone, and oestradiol (an oestrogen), dysregulation in sex hormone-binding globulin and testosterone may lead to dysregulated levels of oestrogen, which can further inhibit follicular maturation (Zain and Norman, 2008). Evidence has shown that increased BMI is also associated with a significantly higher threshold dose and higher total dose of go-

nadotrophins required for the stimulation, and a longer duration of stimulation for IVF treatment (Balen et al., 2006; Fedorcsák et al., 2004). However, because cortisol plays an important role in energy utilisation, it is also possible for BMI to be associated with cortisol through its effect on hormones maintaining the energy metabolism, such as insulin and leptin (Zain and Norman, 2008; Purnell et al., 2004). Therefore, BMI is an important lifestyle covariate to consider when investigating the association between stress, cortisol, and IVF outcomes.

3.2.2 Physical activity and exercise

Physical activity is another lifestyle factor that is closely associated with BMI (Watson, 2015). Physical activity refers to bodily movement that requires movement of the skeletal muscles (especially the legs and arms) and results in energy expenditure above the resting level (Caspersen et al., 1985). Exercise is physical activity that is planned, structured, repetitive, and aims to improve or maintain physical fitness (Caspersen et al., 1985). In general, all physical activities are classified into three categories based on their intensity: moderate-intensity activities, such as walking or swimming at a low, moderate, or brisk pace; vigorous endurance activities, such as jogging, running, or cycling at a fast pace; and activities that increase strength and flexibility, such as weight training, heavy lifting, or carrying tasks (Owen et al., 2007). Compelling evidence has shown that regular and adequate physical activity can bring a variety of health benefits, including improving muscular and cardiorespiratory fitness; reducing the risk of diseases, such as hypertension, diabetes, coronary heart diseases, and depression; and maintaining energy balance and healthy weight (Organization et al., 2009; World Health Organization et al., 2010).

In a systematic review, Hakimi and Cameron (2017) examined the effect of physical activity on ovulation in reproductive-age women. They analysed findings from 10 interventions and four cohort studies and found an association between

an increased risk of ovulation dysfunction and extremely heavy exercise (>60 min/day) and between a reduced risk of ovulation dysfunction and vigorous exercise of 30–60 min/day.

Evidence has also suggested a positive effect of physical activity on IVF treatment outcomes. In a meta-analysis based on a systematic review to evaluate the effect of maternal physical activity prior to IVF treatment on the treatment outcomes, Rao et al. (2018) reviewed eight studies, including observational ($k = 7$) and interventional ($k = 1$) studies, with a population of 3,683 couples undergoing IVF, reporting on the association between physical activity and IVF outcomes, including implantation, clinical pregnancy, miscarriage, and live birth rates. They showed that women who regularly engaged in physical activity prior to IVF treatment had an increased chance of clinical pregnancy ($OR = 1.96$, 95% CI: 1.40–2.73) and live birth ($OR = 1.95$, 95% CI: 1.06–3.59) when compared with women who did not regularly participate in physical activity. However, no significant associations were found between physical activity and implantation rate or miscarriage rate (Rao et al., 2018).

Physical activity can also be regarded as a stressor because it causes many bodily systems, especially the cardiopulmonary system, to adapt to a new state (Mastorakos et al., 2005). This change of state causes changes in homeostasis, which then require primary mediators of allostasis (see Chapter 2), including hormones of the HPA axis, to maintain or restore the balance (McEwen, 2005). Robust evidence has shown that acute physical activity is associated with the secretion of pituitary hormones, such as growth hormone, ACTH, β-endorphin, and β-lipotropin (Elias and Wilson, 1993). As was discussed in Chapter 2, as part of the HPA axis, the secretion of ACTH will induce the secretion of cortisol. At the same time, the secretion of ACTH, β-endorphin, and β-lipotropin may inhibit the secretion of GnRH through the negative feedback mechanism (i.e., an increase in hormones produced by the pituitary, such as ACTH, travels back

to the hypothalamus and suppresses the secretion of GnRH) (Elias and Wilson, 1993). The reduced secretion of GnRH then further reduces the secretion of other hormones, such as LH and FSH, which are essential hormones regulating the female reproductive system (Elias and Wilson, 1993). Therefore, physical activity, as a lifestyle, can theoretically influence both cortisol and female reproductive function, which may then have an impact on IVF treatment outcomes.

3.2.3 Smoking

Worldwide, smoking is one of the biggest behavioural health risks that can cause premature death and a variety of health problems (Reitsma et al., 2017; Forouzanfar et al., 2016). Smoking damages nearly all bodily systems, especially the respiratory and the cardiovascular system (Warner, 2015). It has been identified as being among the top five behavioural risk factors for a wide range of diseases (e.g., cardiovascular and circulatory diseases, cancer, and chronic respiratory conditions), early death, and disability (Reitsma et al., 2017; Forouzanfar et al., 2016).

Smoking is a well-demonstrated risk factor for infertility in both males and females, with compelling evidence from previous studies. Population-based studies have shown that exposure to smoking may be related to an increased risk of infertility and delays in conception (Hull et al., 2000). Several systematic reviews have also reported on the pooled negative effect of smoking on fertility. For example, in a meta-analysis review of 34 studies reporting on the association between male semen quality and socio-behavioural factors, including smoking, Li et al. (2011) showed that smoking was a significant risk factor for five classic sperm parameters, including semen volume (pooled mean difference = 0.25); the density (pooled mean difference = 7.07), normality (pooled mean difference = 4.92), and motility (pooled mean difference = 1.85) of the sperm; and total sperm count (pooled mean difference = 32.20). In another systematic review on

the association between smoking and the risk of infertility in women of reproductive age, Augood et al. (1998) reviewed 12 studies and found that overall, women who smoked were 60% more likely to experience infertility when compared with non-smokers (OR = 1.60, 95% CI: 1.34–1.91). In addition, in cohort studies (k = 9) women who smoked were 42% more likely to experience delayed conception (over one year) then non-smoking women (OR = 1.42, 95% CI: 1.27–1.58). In case-control studies (k = 4) women who smoked were 2.27 times more likely to experience infertility than non-smokers (OR = 2.27, 95% CI: 1.28–4.02) (Augood et al., 1998).

The influence of smoking on fertility can extend to fertility treatment. In a meta-analysis based on a systematic review of 21 studies reporting on the outcomes of ART (e.g., clinical pregnancy, live birth, ectopic pregnancy, and spontaneous miscarriage) in cigarette smoking patients and non-smoking patients, Wallach et al. (1996) showed that patients who smoked had significantly lower odds of clinical pregnancy (OR = 0.56, 95% CI: 0.43–0.73) and live birth per cycle (OR = 0.54, 95% CI: 0.30–0.99) and significantly higher odds of spontaneous miscarriage (OR = 2.65, 95% CI: 1.33–5.30) and ectopic pregnancy (OR = 15.69, 95% CI: 2.87–85.76). In a meta-analysis based on another systematic review that synthesised evidence from 21 studies on the successful ART outcomes of patients who actively smoked during treatment and those who did not, Waylen et al. (2008) found that smoking in female patients was significantly related to a reduced likelihood of positive IVF treatment cycle outcomes, including live birth and pregnancy rate, and increased risks, including miscarriage and ectopic pregnancy. Similarly, in another systematic review of 22 studies, Klonoff-Cohen (2005) concluded that smoking negatively influenced IVF outcomes, for instance, fertilisation, pregnancy, and live birth rates.

The exact mechanism of the influence of smoking on IVF outcomes is not well understood (Waylen et al., 2008), but the HPA axis could be involved. Smoking

is a stressor that can acutely increase the plasma levels of prolactin, ACTH, and growth hormone (Kapoor and Jones, 2005). As described earlier, increases in these hormones may induce the secretion of cortisol and inhibit other hormones involved in the female reproductive system. Robust evidence has demonstrated significant increases in cortisol levels immediately after smoking (Chiodera et al., 1997) and significantly higher levels of cortisol secretion throughout the day in smokers than in non-smokers (Badrick et al., 2007). However, the effect of habitual smoking on long-term levels of cortisol has not been well investigated (Badrick et al., 2007; Wosu et al., 2013).

3.2.4 Alcohol consumption

A WHO estimate published in 2016 revealed that approximately 43% of the global population aged 15 years or over were current drinkers. More than 59% of the current drinkers in 2016 were from only three regions: the Americas, Europe, and the Western Pacific (Organization, 2019). Alcohol consumption is another factor that may threaten fertility and successful IVF treatment. In male populations, evidence has shown that alcohol consumption may negatively affect semen quality. In a meta-analysis of four studies based on a systematic review, Li et al. (2011) showed that alcohol can significantly reduce semen volume (pooled effect = 0.30). In a prospective cohort study that followed 430 Danish couples who were trying to conceive over six menstrual cycles, Jensen et al. (1998) found that after adjustment for menstrual cycle, smoking, diseases of the female reproductive organs, female BMI, sperm concentration, and duration of the menstrual cycle, women who consumed one to five drinks per week were 40% less likely to conceive than women with no alcohol intake. Women who consumed more than ten drinks per week were 66% less likely to conceive than women with no alcohol intake.

Evidence has shown that alcohol consumption may influence the outcomes of ART. For example, Klonoff-Cohen et al. (2003) conducted a prospective cohort

study (n = 221 couples) to investigate the influence of female and male alcohol consumption on a variety of endpoints of ART treatment, e.g., live birth, sperm count, number of oocytes aspirated, number of oocytes fertilised, number of embryos transferred, biochemical pregnancy, clinical pregnancy, multiple gestations, spontaneous miscarriage, and infant characteristics (i.e., gestational age and infant birth weight). Overall, they reported that both the amount and timing of female and male alcohol consumption may affect the endpoints of IVF treatment. The closeness of the timing of women's alcohol consumption to ART may have had a negative effect on achieving a pregnancy. Because the average alcohol content of one drink is usually 12 g, Klonoff-Cohen et al. (2003) did comparisons in ART outcomes between women who drank none alcohol, $\leq$ 12 g, and $>$ 12 g. They found that during the year before starting ART, women who drank 12 g more of alcohol per day were more likely to have significantly decreased number of oocytes aspirated by 13%. During the one month before starting ART, compared with women who drank 12 g fewer per day, Women who drank 12 g more of alcohol per day were at a significant increased the risk of not achieving pregnancy by 2.86 times. During the one week before starting ART, compared with women who drank 12 g fewer per day, Women who drank 12 g more of alcohol per day were at a significant risk of miscarriage by 2.21 times (Klonoff-Cohen et al., 2003). For males, every 12 g more alcohol consumption per day increased the the risks of not achieving a live birth by 2.28–8.32 times depending on the time of drinking (i.e., from one year prior to the attempt to the week before sperm collection) (Klonoff-Cohen et al., 2003).

In another prospective cohort study on 2,545 couples and 4,729 IVF cycles, Rossi et al. (2011) investigated the association between alcohol consumption at the beginning of IVF treatment and live birth following IVF and showed that women who drank at least four drinks per week had a 16% lower chance of live birth than those who drank fewer than four drinks per week. For couples who drank at least four drinks per week, the odds of a live birth decreased by 21%. In addition to

the influence on live birth, Rossi et al. (2011) found a negative effect of alcohol consumption on fertilisation and implantation. Female and male patients who consumed more than four drinks per week were 48% more likely to experience failed fertilisation (female: $OR = 1.48$; 95% CI: 1.13–1.95; male: $OR = 1.36$; 95% CI: 1.06–1.74) than those who consumed less than four drinks per week. Female patients who drank one to seven drinks peer week were 18% more likely to experience failed implantation ($OR = 1.18$; 95% CI: 1.03–1.35). No significant associations were found between alcohol consumption in female patients and the number of oocytes retrieved. Rossi et al. (2011) also observed that men who drank beer daily had a 27% lower likelihood of having poor sperm motility, 36% increased risks of failed implantation during treatment, and reduced likelihood of achieving a live birth by 35%, compared with men who drank less than one beer per week. However, men who drank wine weekly had a 23%–43% higher likelihood of having poor sperm morphology, depending on the type of wine they consumed (Rossi et al., 2011).

Animal models have shown that acute and chronic alcohol consumption can contribute to a disruption in the levels of hormones that are essential for female reproductive function, such as oestradiol, oestrogen, LH, and testosterone, which has been evidenced by consistent findings of disrupted menstrual cycles in female alcohol consumers (Emanuele et al., 2002). In particular, acute and chronic alcohol consumption have been shown to reduce the secretion of insulin-like growth factor 1 and its ability to induce the secretion of LH (Les Dees et al., 1998). As explained earlier, because of the close relationships between the hormones involved in the HPG axis and the HPA axis, the effects of alcohol on the female sex hormones may also influence the secretion of cortisol.

3.2.5 Caffeine consumption

Caffeine is a natural active ingredient found in more than 60 plants, including coffee beans, tea leaves, cola nuts, and cocoa pods (Mejia and Ramirez-Mares, 2014). Beverages, such as coffee, tea, energy drinks, and some soft drinks; chocolate and cocoa-containing products; some dietary supplement (e.g., weight management tablets); and certain medications (e.g., headache treatment) are common products that contain caffeine. Although it is one of the most routinely ingested ingredients worldwide (Mejia and Ramirez-Mares, 2014), the effects of caffeine on health are not well understood (Wikoff et al., 2017). Epidemiological and cohort studies suggest that caffeine may play a protective role in chronic diseases, including type 2 diabetes, Parkinson's disease, Alzheimer's disease, and liver diseases, and may help individuals to achieve weight loss and management (Higdon and Frei, 2006; Mejia and Ramirez-Mares, 2014; Grosso et al., 2017). However, studies have also suggested that caffeine consumption is associated with increased blood pressure and disturbed heart rhythms, which are risk factors for cardiovascular diseases (Mejia and Ramirez-Mares, 2014). Excessive caffeine consumption may also contribute to other issues, such as bone loss, lower bone density or fractures, headaches, nausea, and hypertension (Mejia and Ramirez-Mares, 2014).

There is evidence suggesting that caffeine consumption may influence sex hormones in women. For example, a study on 728 postmenopausal women found that higher caffeine consumption was significantly associated with lower levels of testosterone, greater plasma oestrone, and higher plasma sex hormone-binding globulin levels (Ferrini and Barrett-Connor, 1996). In two studies on premenopausal women, greater caffeine consumption was significantly associated with higher levels of sex hormone-binding globulin (London et al., 1991; Oba et al., 2010). A more recent study found that higher caffeine consumption was significantly associated with increased levels of oestradiol (Lucero et al., 2001).

Although the above evidence suggests a caffeine consumption–female sex hormone–female reproductive function pathway, the effect of caffeine consumption on fertility is inconclusive. In a systematic review on the association between caffeine consumption and natural fertility, spontaneous abortion, and fertility treatment outcomes, Lyngsø et al. (2017) reviewed eight studies reporting the association between caffeine consumption and natural fertility. A meta-analysis of three of the eight studies showed no significant association between caffeine consumption of up to 400 mg per day and the probability of a couple conceiving naturally in one menstrual cycle. Of the other five studies, two reported a significant association between increased caffeine consumption and a reduced probability of achieving natural pregnancy, and three reported no association (Lyngsø et al., 2017). Lyngsø et al. (2017) reviewed four studies reporting on the association between caffeine consumption and longer waiting times to pregnancy. A meta-analysis of two of the four studies showed no significant pooled effect of caffeine on longer waiting times to pregnancy, which was also suggested by the other two studies that were not included in the meta-analysis (Lyngsø et al., 2017). In a meta-analysis of the association between caffeine consumption and spontaneous abortion based on 27 studies, Lyngsø et al. (2017) found that higher caffeine consumption was significantly associated with increased risks of spontaneous abortion, which was supported by four out of five other studies that were not included in the meta-analysis. Unlike the relatively abundant evidence for the association between caffeine consumption, natural fertility, and spontaneous abortion, only three studies investigated the association between caffeine consumption and fertility treatment outcomes. Two of the studies were included in the meta-analysis and showed no significant association between caffeine consumption and live birth rates following IVF treatment. The other study reported no association between caffeine consumption and pregnancy rates following IVF treatment. Given the significant heterogeneity and risk of bias, Lyngsø et al. (2017) suggested that there is still a lack of evidence on the association between

caffeine and fertility, and it is necessary to investigate whether caffeine consumption can affect natural fertility or, most especially, ART treatment outcomes.

Published in the same year, Abadia et al. (2017) conducted a prospective cohort study on 300 patients undergoing IVF to investigate the association between pre-treatment caffeine intake and IVF outcomes, including implantation success, clinical pregnancy, and live birth rates. They found that low to moderate levels (<200 mg/day) of caffeine intake in the year prior to IVF treatment were not associated with the implantation success, clinical pregnancy, or live birth rate following the treatment. They also found no association between pre-treatment caffeine intake and ovarian stimulation, fertilisation, or embryo quality. The current evidence does not seem to support an association between caffeine consumption and IVF treatment outcomes. However, only two studies reporting on the live birth rate and one reporting on pregnancy outcomes following IVF were included in the systematic review (Lyngsø et al., 2017). In addition, Abadia et al. (2017) measured caffeine intake in the year prior to IVF treatment; therefore, it was unclear whether the patients undergoing IVF had changed their habitual consumption of caffeine shortly before the treatment, and whether short-term (e.g., one month) caffeine consumption prior to treatment was associated with pregnancy outcomes following the treatment.

3.2.6 Sleep

Sleep is a fundamental, yet often underestimated, lifestyle factor that is closely related to health (Schlarb, 2015). In adults, sleep duration typically varies between 6.5 and 9 hours of sleep within a 24-hour period. Sleep is thought to influence hormone regulation, metabolism, and the circadian rhythm, which are essential for maintaining homeostasis (Ellis, 2007). Although the full mechanism of how health is influenced by sleep remains undetermined (Ellis, 2007), evidence has revealed that poor sleep is associated with a variety of health outcomes, such

as mortality, cardiovascular disease, type 2 diabetes, hypertension, mood disturbance, depression, and obesity (Cappuccio et al., 2010; Ferrie et al., 2007). Apart from the direct association between sleep and health-related outcomes, other lifestyle factors, such as physical activity, alcohol and caffeine consumption, and diet, may be associated with sleep (Ellis, 2007; Schlarb, 2015).

Evidence has shown an association between sleep and fertility treatment outcomes. A previous systematic review investigated the impact of sleep on female and male reproductive functions, including the association between sleep and ovarian function, spermatic function, natural fertility, and IVF treatment outcomes (Caetano et al., 2021). A total of 33 studies were included in the review. Six studies reported on the association between sleep quality or length and ovarian function, of which one concluded that poor sleep quality was a risk factor for premature ovarian failure, two reported that poor sleep quality or insomnia was associated with the incidence of menstrual cycle problems, two reported that a sleep duration shorter than five hours was associated with an increased risk of menstrual cycle irregularities, and one reported no association between shorter sleep duration and menstrual cycle irregularities (Caetano et al., 2021). Four studies included in the review reported associations between shift work and ovarian function: One study found no association, while the other three reported that rotating shifts were associated with risks of irregular menstrual cycles and abnormally short or long cycle duration (Caetano et al., 2021). For the association between sleep and sperm parameters, such as sperm concentration and total sperm count, 12 studies were included in the review. Nine studies showed a significant association between sleep quality and sleep duration and sperm concentration, total sperm count, and normal sperm morphology (Caetano et al., 2021). The other three studies found no significant associations between sleep quality, sleep duration, or shift work and sperm parameters (Caetano et al., 2021). Five studies investigated the association between sleep and natural fertility with mixed findings. Two studies found no significant associations between sleep duration, sleep

difficulties, or shift work and difficulty achieving natural pregnancy. One study found that women who were classified as intermediate chronotypes (i.e., did not prefer being physically and mentally active in the mornings or in the evenings) experienced more difficulties in achieving natural pregnancy than women who preferred being physically and mentally active in the mornings. One study reported a significant association between shorter sleep duration and infertility in couples. One study reported a significant association between shift work and infertility in men, but there were significant differences in the study groups in terms of demographic characteristics.

For the association between sleep and IVF outcomes, Caetano et al. (2021) reviewed six studies, of which four reported on associations between sleep disorders and IVF outcomes, and one reported on sleep quality and emotional distress among patients undergoing IVF. However, the findings of these studies were not conclusive. Three reported impaired sleep quality among women undergoing fertility treatments, both pre-treatment and during treatment. One study did not find a significant association between sleep disturbances during ovarian stimulation and the number of oocytes or embryos preserved. One study reported that evening, night, or rotating shift work were significantly associated with poor oocyte retrieval outcomes. One study reported a significant association between sleep disorders and poor ovarian response or poor oocyte retrieval outcomes. One study did not find a significant association between sleep duration and the number of oocytes retrieved or the fertilisation rates but found a significantly higher pregnancy rate in women with a sleep duration of 7–8 hours than in women with a sleep duration of 9–11 hours (Caetano et al., 2021). Due to the heterogeneity in the measurement of sleep and IVF outcomes, Caetano et al. (2021) did not reach a clear conclusion.

However, current evidence suggests that women undergoing fertility treatments are likely to experience poor sleep quality. For example, an observational study in

Denmark followed 163 women and 132 partners undergoing IVF or ICSI throughout their treatment cycle and surveyed their sleep quality (Philipsen et al., 2021). They collected data pre-treatment, during treatment (prior to the pregnancy test), and 9–11 weeks after the second survey (after treatment). The results showed that the average scores on the sleep quality measure (Pittsburgh Sleep Quality Index (Buysse et al., 1989)) in women undergoing IVF or ICSI were 8.3 pre-treatment (*standard deviation [SD]* = 2.4), 8.4 during treatment (*SD* = 2.3), and 8.1 after treatment (*SD* = 2.3), which were all higher than the cut-off for poor sleep quality (i.e., scores >5). In addition, at least 90% of the women surveyed reported poor sleep quality at all three time points (Philipsen et al., 2021). Similar results were found among the partners of the women undergoing treatment: Pittsburgh Sleep Quality Index scores were 8.0 pre-treatment (*SD* = 2.3), 7.8 during treatment (*SD* = 2.3), and 7.9 after treatment (*SD* = 2.2), and 85%–89% of the partners reported poor sleep quality (i.e., scores >5) at all three time points. No significant differences were found between the total sleep scores at different time points (Philipsen et al., 2021).

One hypothesis relating to the effect of sleep on IVF treatment outcomes is that sleep dysregulation can dysregulate the activation of the HPA axis, both directly and indirectly, by increasing stress, which then activates the HPA axis (Kloss et al., 2015). A previous systematic review evaluated evidence on the association between sleep and HPA axis reactivity, as measured by cortisol, and found that of the four studies reporting the influence of self-reported sleep quality and the cortisol responses to stressors, one study reported greater cortisol responses to a physical stressor in poor sleepers than in good sleepers, two studies reported no association between reduced sleep quality and cortisol reactivity to a psychosocial stressor or cognitive-interference task, and one study reported increased cortisol responses to tasks in male but not in female participants with poor sleep quality when compared with good sleepers (van Dalfsen and Markus, 2018). They also reported on the association between sleep duration and acute cortisol responses

to stressors. Of the seven studies included in the review, five studies found no significant effect of variations in sleep duration on acute cortisol responses to stressors, one study found increased cortisol responses for those with longer sleep duration when compared with those with a shorter sleep duration, and one study found increased cortisol responses after a night of total sleep deprivation when compared with a normal night of sleep (van Dalfsen and Markus, 2018). The results revealed that self-reported sleep quality may be associated with greater acute cortisol responses to stressors, but variations in sleep duration may have no influence on cortisol responses (van Dalfsen and Markus, 2018). However, this review only synthesised data on the influence of sleep on acute cortisol levels in response to tasks rather than long-term levels of cortisol.

Other evidence has shown that sleep may be associated with psychological stress in patients undergoing IVF. In a longitudinal study to investigate the relationship between self-reported somatic symptoms, sleep disturbances, and psychological distress in 100 female patients undergoing oocyte retrieval or IVF, Lin et al. (2016) showed that female patients with sleep disturbances were 3.13 times more likely to experience psychological distress during oocyte retrieval treatment than patients without sleep disturbances. Female patients with sleep disturbances undergoing IVF were 2.16 times (95% CI: 1.20–3.12) more likely to experience psychological stress during treatment than patients without sleep disturbances. Based on the current evidence, it is difficult to conclude whether sleep influences IVF treatment outcomes owing to the different treatment endpoints, measures of sleep, and sleep measurement time points in the relevant studies. In addition, evidence of whether sleep is associated with long-term levels of cortisol is lacking. Therefore, it is necessary to investigate the association between sleep as a lifestyle factor and a variety of IVF outcomes, such as biochemical pregnancy and clinical pregnancy, both directly and indirectly through examining cortisol levels.

3.3 An individual pathway to fertility

Homan et al. (2007) reviewed existing evidence on the impact of lifestyle factors on fertility in both the general population and those undergoing fertility treatment, with a focus on specific lifestyle factors, including smoking habits, weight, diet, exercise, caffeine consumption, alcohol consumption, and exposure to environmental pollutants. They reported that there was strong evidence for adverse effects of smoking, obesity, and low weight on fertility in both the general population and patients undergoing ART. There was also strong evidence for positive effects of a well-balanced diet and moderate levels of exercise on fertility. However, evidence supporting these conclusions was mostly from the general population and not patients undergoing ART. The effect of other lifestyle factors, including caffeine consumption, alcohol consumption, and exposure to environmental pollutants, on fertility was less conclusive due to inconsistency in the evidence. Specifically, whether there were dose–response relationships between caffeine consumption, alcohol consumption, and exposure to environmental pollutants and fertility outcomes was unclear (Homan et al., 2007). They reported that most of the studies investigating lifestyle choices, such as caffeine consumption and alcohol intake, relied heavily on participants' recall of the amount and timing of their consumption, which could potentially introduce recall errors into the study. Homan et al. (2007) also reviewed evidence on the effect of psychological stress on fertility in both populations and reported that the evidence was inconclusive. As discussed in the previous chapter, the inconsistency in the evidence may be related to how stress was measured, when stress was measured, and how fertility and treatment outcomes of ART were defined in the studies. Similarly, in studies on the impact of lifestyle factors on fertility and outcomes of ART, there were considerable differences in how and when lifestyle choices were measured.

Drawing on their findings, Homan et al. (2007) proposed a model of "individualized pathways to conception and a healthy child" and suggested that when a

couple faced with infertility attends a clinical consultation, modifiable lifestyle factors should be identified in conjunction with the medical assessment. These lifestyle factors and associated behaviours are modifiable and may assist couples in achieving natural conception or optimise their chances of conceiving through ART treatment (Homan et al., 2007). Treatment plans should be tailored to patients based on both the infertility diagnoses and the identified lifestyle factors, with the aim of optimising treatment outcomes. For example, if a patient is diagnosed with ovulation dysfunction and the associated lifestyle factors are obesity, poor diet, and physical inactivity, the patient could be advised to participate in exercise, improve their diet, and achieve a healthier weight in order to assist in the treatment of the ovulation dysfunction. Homan et al. (2007) suggested that this individualised pathway to reproductive health should be considered an initial strategy in fertility healthcare for two main reasons. First, with the individual management of lifestyles, some patients may achieve conception and live birth without the need for high-cost fertility treatments. Second, with the assistance of lifestyle modification, the efficacy of fertility treatments will be improved.

Homan et al. (2007) also pointed out a noteworthy limitation in the evidence available: a lack of evidence of the underlying biomedical mechanisms though which lifestyle factors affect IVF treatment outcomes, e.g., changes in hormone levels. Indeed, the evidence discussed above mainly relates to the direct associations between lifestyle factors and IVF outcomes, with inconclusive findings on the effects of several lifestyle factors (e.g., caffeine consumption and sleep) on IVF outcomes. A popular hypothesis is that these lifestyle factors may influence IVF outcomes through their effects on female sex hormones (e.g., oestrogen, LH, and other hormones involved in the HPG axis). As explained in Chapter 2, the HPA axis and the HPG axis are closely associated; therefore, the hormone of interest in this thesis, cortisol, may also be involved in the pathway through which these lifestyle factors influence IVF outcomes. It is also possible that since some of these lifestyle factors are regarded as stressors (e.g., physical activity and poor

sleep) (Mastorakos et al., 2005; van Dalfsen and Markus, 2018), they may influence IVF treatment outcomes through stress. However, current research either investigated the association between lifestyle factors and IVF outcomes, the association between stress (sometimes including the measurement of cortisol) and IVF outcomes, or the association between lifestyle factors and stress. Very few studies brought the evidence together or comprehensively examined the associations between lifestyle, stress, and cortisol and IVF outcomes (Massey et al., 2014; Nicoloro-SantaBarbara et al., 2018). Therefore, a study investigating all three aspects will provide valuable insights into how modifiable lifestyle and stress-related factors are associated with IVF outcomes and whether the hormone cortisol is a pathway through which these associations occur.

To summarise, although it is widely believed that certain lifestyle factors such as sleep, alcohol consumption, and physical activities may influence the outcomes of IVF treatment (Homan et al., 2007; Hawkins et al., 2014; Rossi et al., 2016), research evidence is less clear on how and why these factors may have such effects. It is therefore important to examine the possible associations between a wide range of lifestyle factors and IVF treatment outcomes, both directly and indirectly through potential pathways such as the HPA axis and the hormone cortisol.

Chapter 4

Lifestyle, Psychological Factors and Cortisol

The relationships between infertility, infertility treatment, stress, and lifestyle are complex. In Chapter Two, the major theories relating to stress, psychobiological pathways of stress (especially cortisol), and ways in which stress may influence fertility and IVF treatment outcomes were discussed. In Chapter 3, lifestyle factors were introduced that have been thought to influence IVF treatment outcomes and the available evidence on these factors was discussed. In the model that was built as the theoretical basis for this thesis (see Figure 1), one of the missing pieces of the puzzle that requires further exploration is the relationship between cortisol and lifestyle behaviours. Studies have investigated the association between lifestyle factors, stress, and the functioning of the HPA axis over time using chronic measures or repeated sampling of cortisol as a biomarker. What follows in Chapter 4 is a systematic review that was carried out as part of this thesis to synthesise and evaluate the research evidence from the past 20 years (Jia et al., 2022).

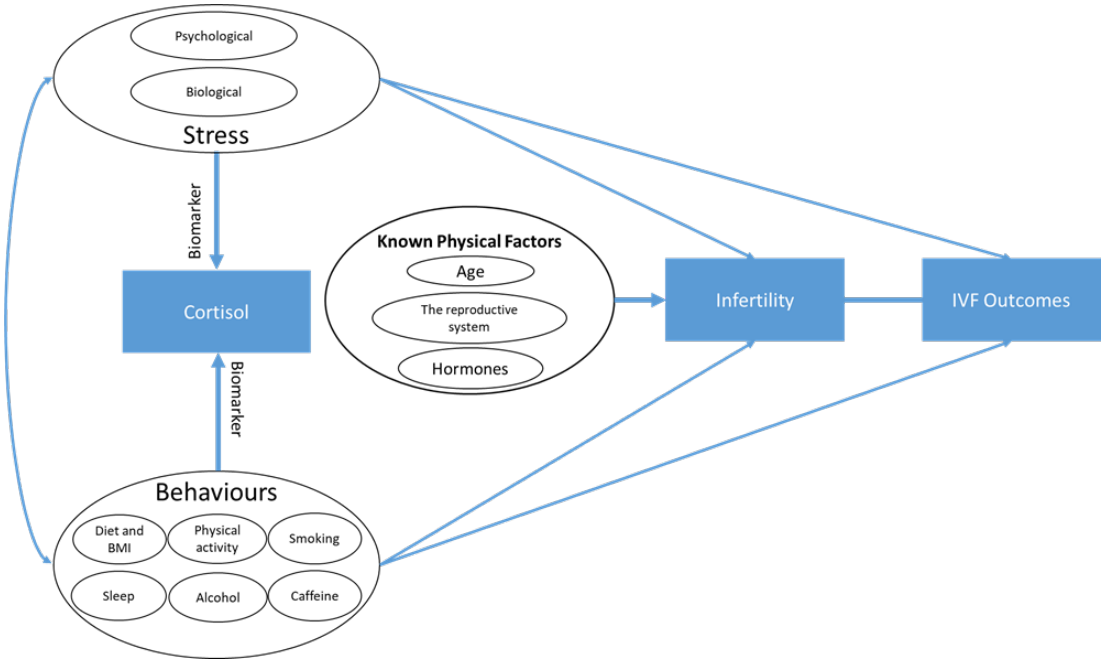


Figure 4.1 Thesis conceptual model

4.1 Associations between modifiable lifestyle and psychological factors and long-term levels of cortisol: a systematic review

As discussed in Chapter 2, the final product of the HPA axis, cortisol, is a glucocorticoid steroid hormone secreted by the adrenal cortex that is often characterised as a "stress hormone" because of its critical role in the physiological response to psychological stress. However, when stressors occur, both physiological (e.g., hormone secretions) and behavioural (e.g., alcohol consumption and smoking) responses are triggered, which (in the case of an adaptive response) contribute to the restoration and maintenance of homeostasis. For example, stress-related food consumption can contribute to changes in cortisol levels, which, in turn, impact glucose homeostasis and energy deposition (McEwen and Wingfield, 2003). Therefore, cortisol is associated not only with mood but also with a range of lifestyle factors that respond to and are affected by stress-related changes in homeostasis and circadian rhythms, e.g., sleep, physical activity, and diet (Fukuda and Morimoto, 2001).

Several reviews have examined the relationship between cortisol, mood, and lifestyle factors. However, these have been limited in a variety of ways, e.g., by focusing on one type of cortisol measure (hair cortisol), a specific cortisol index (diurnal cortisol slopes), or a specific population. For example, Adam et al. (2017) examined the association between diurnal cortisol slopes and mental and physical health outcomes and reviewed 179 associations reported in a total of 80 studies. Significant pooled associations were found between flatter cortisol diurnal slopes and worse outcomes for 10 out of 12 mental and physical health outcomes that were assessed: depression, internalising disorders, externalising disorders, fatigue, immune or inflammatory outcomes, BMI or obesity, cancer, other mental health outcomes, other physical health outcomes, and mortality

(Adam et al., 2017). In contrast, Lopresti et al. (2013) conducted a narrative review of lifestyle factors that contribute to pathways associated with major depression. They described findings from several animal- and human-based studies on the relationship between diet, sleep, and exercise and major depressive disorder and the role of these three lifestyle factors in physiological pathways, including neurotransmitter processes, immune–inflammatory pathways, HPA axis disturbances, oxidative stress and antioxidant defence systems, neuroprogression, and mitochondrial disturbances (Lopresti et al., 2013). Regarding the contribution of diet to major depression through the HPA axis, four studies revealed that cortisol diurnal curves with medium or low variability were associated with diets high in fat and saturated fatty acids and that omega-3 supplements may lower cortisol activity. For the relationship between sleep and the HPA axis, two studies suggested that cortisol levels were higher in individuals with insomnia. A further seven studies experimentally manipulated the effects of sleep deprivation or restriction on the HPA axis, with inconsistent results. Two studies found elevated levels of cortisol after acute sleep deprivation, and two studies found either no change or decreased levels of cortisol after sleep deprivation. The third topic included in this review is the effect of exercise on the HPA axis. Ten studies showed that acute exercise elevated acute levels of cortisol. In general, the relationship between exercise and HPA activity is complex and appears to be influenced by the type, duration, intensity, and chronicity of the exercise; the measure of HPA activity; and the population in the study (Lopresti et al., 2013). Based on the evidence, Lopresti et al. (2013) concluded that the studies mainly illustrated the effect of these three lifestyle factors on acute activity of the HPA axis, including acute cortisol secretion, but the long-term effects of these lifestyle factors on the HPA axis were not well understood.

Although both reviews helped to enhance our understanding of the relationship between cortisol and mood and cortisol and lifestyle, the focus on a specific population and a limited number of indices of cortisol, mood, and lifestyle necessarily

limited the conclusions. This was further compounded by the fact that both reviews focused on acute or short-term measures of cortisol, which, in part, reflects the history of research in this field, which has focused on short-term measures of cortisol (such as awakening cortisol, evening cortisol, or immediate cortisol levels after an intervention or challenge). However, the pulsatile nature of the hormone (Young et al., 2004) and concerns regarding the health relevance of single, isolated measures has increased interest in capturing levels over longer periods of time (Adam et al., 2017).

This has been facilitated by innovations in the field and, in particular, the development of methods to assess cortisol levels in hair, with HCCs being able to provide a description of cortisol levels over periods of one to six months (Stalder and Kirschbaum, 2012). Wosu et al. (2013) and Stalder et al. (2017) conducted separate systematic reviews to investigate correlates of HCC. Wosu et al. (2013) examined the association between HCC and socioeconomic factors (e.g., income and education), occupational stressors (e.g., shift work), age, sex, ethnicity, psychiatric symptoms and disorders (e.g., bipolar disorder and posttraumatic stress disorder [PTSD]), adrenocortical conditions (e.g., Cushing syndrome), pregnancy, early life adversity (e.g., childhood maltreatment), BMI, waist-to-hip ratio, alcohol use, smoking habits, oral contraceptive use, and physical activity. Focusing on the findings of direct interest to this review (i.e., indices of mood and lifestyle), they found associations between HCC and psychiatric symptoms and disorders, occupational stressors, physical activity, BMI, waist-to-hip ratio, and alcohol intake. They also reported that the available evidence did not support an association between HCC and smoking and found insufficient evidence to examine the association between diet, sleep quality, caffeine consumption, and HCC.

However, while comprehensive, Wosu et al. (2013) only narratively synthesised evidence on possible correlates of HCC. In contrast, Stalder et al. (2017) quantitatively investigated the correlation between HCC and age, sex, BMI, waist-

to-hip ratio, hair washing frequency, hair treatment, smoking, oral contraceptive use, blood pressure, previous and current chronic stressors, social support, and depressive symptoms. With a focus on mood and lifestyle indices alone, the meta-analysis showed that HCC was significantly and positively associated with previous and current chronic stress, BMI, waist-to-hip ratio, and blood pressure and negatively associated with anxiety disorders. However, no significant associations were found between HCC and other mood disorders, social support, or smoking. In summary, previous reviews have been limited by their focus on a restricted range of indices for cortisol, mood, and/or lifestyle factors, and, to date, none have included evidence from interventional studies. Therefore, this review will address these gaps and provide an updated synthesis of the evidence pertaining to the nature and direction of the associations between long-term levels of cortisol, lifestyle, and mood factors.

4.2 Methods

4.2.1 Eligibility criteria

(1) Population

This review focused on adults aged 18–65 years who were not on any medications and had no medical conditions (e.g., cancer or auto-immune diseases) or mental disorders because conditions (e.g., Cushing syndrome and major depressive disorder) and the use of certain medications (e.g., steroids) are associated with changes in cortisol levels (Casals and Hanzu, 2020; Rothe et al., 2020), which may confound the relationship between lifestyle or mood and long-term cortisol. Patients undergoing any type of medical treatment were excluded for similar reasons.

(2) Cortisol parameters

This review focused on long-term levels of cortisol only, including HCC and longitudinal changes in cortisol measured in blood, saliva, or urine. Hair cortisol is usually measured in 1–3 cm hair segments representing cortisol levels in the past 1–3 months (Stalder and Kirschbaum, 2012). Although other cortisol indices, such as cortisol reactivity, CAR, AUCg, and diurnal slope, are typically considered acute cortisol indices (Saxbe, 2008), studies investigating longitudinal changes in these cortisol indices were included in this review. The minimum time frame for studies measuring cortisol in hair was one month (four weeks). Therefore, to assess evidence that could be considered to be comparable, repeated measurements of cortisol using any of the aforementioned acute measures (i.e., blood, saliva, urine, or follicular fluid) (Stalder and Kirschbaum, 2012) had to cover a minimum period of at least four weeks to be included in this review. Studies involving the manipulation of cortisol levels through pharmacological agents were excluded.

(3) Intervention

Interventional studies targeting lifestyle factors (e.g., physical activity, diet, caffeine consumption, alcohol consumption, smoking, and sleep), surrogates for lifestyle (e.g., BMI), or measures of mood (e.g., stress, anxiety, depression, and PTSD) were included.

(4) Outcomes

Reported associations between long-term levels of cortisol and lifestyle and/or mood factors served as outcomes. For interventional studies where acute levels of cortisol were measured repeatedly, reported changes in cortisol levels at different time points within and between groups were also regarded as outcome variables.

Comparisons of cortisol indices between patients and healthy participants were excluded where relationships between lifestyle factors and cortisol indices were not measured and reported. Intervention effects on both lifestyle/mood and cortisol were also reported as outcomes for interventional studies. For studies reporting associations between cortisol and multiple lifestyle or mood factors, each association was included based on the lifestyle or mood factors reported. However, if a study reported an association between cortisol and combined lifestyle or mood factors as the outcome, these outcomes were removed due to difficulties inherent in separating the influence of different parameters.

(5) Study design

Both observational and interventional studies were included in this systematic review. To meet the inclusion criteria, interventional studies had to show evidence of change in at least one lifestyle factor, including BMI, physical activity, diet, caffeine consumption, alcohol consumption, smoking, or sleep, or at least one mood factor, including stress, anxiety, depression, or PTSD. Non-interventional studies had to investigate the relationship between any of the aforementioned lifestyle or mood factors and cortisol.

4.2.2 Exclusion criteria

Reviews and conference abstracts were excluded from the review. Editorials, newspaper articles, and other forms of popular media were excluded. Case studies and qualitative studies were excluded. Non-English papers were excluded. Failure to meet any one of the above eligibility criteria resulted in exclusion from the review. The number of excluded studies and reasons for exclusion after a review of the full text were recorded at each stage.

4.2.3 Information source

The systematic review was registered with PROSPERO (No. CRD42019112339), and the protocol can be accessed at https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=112339. Searches were conducted in the following electronic databases to identify relevant articles: MEDLINE, EMBASE, PsycINFO, Web of Science, and CINAHL. The initial search was conducted on 2018, January 23 and updated on 2018, October 22, which identified one recently published study. Another update was conducted on 2020, November 21, which did not identify any eligible studies. The reference lists of eligible papers were also searched to locate additional studies that were not identified by the database search.

Medical Subject Heading terms combined with free-text searches (see Appendix Table 1) were used and adjusted for each database as necessary. The search was limited to full text and English articles. The title and abstract of each article were checked manually against the inclusion/exclusion criteria to exclude irrelevant articles or studies that did not fit the objectives of this review. Given the large number of results returned from the initial search, title and abstract screening was conducted by one reviewer only. Full-text screening and assessment were conducted by two reviewers with strong inter-rater reliability (Cohen's kappa = 0.83) on inclusion/exclusion. Disagreement on inclusion or exclusion for six out of 157 studies was resolved by discussion.

4.2.4 Data extraction and meta-analysis

Data extraction was conducted independently by two reviewers. Data extraction for each study included information on the publication year, country or region where the study was conducted, sample size, age of the study sample, population description, design of the study (observational or interventional), and protocol for measuring cortisol (type of cortisol, cortisol indices calculated, time period

measured, times of samples, and number of samples per day).

Due to the high degree of heterogeneity in the designs of the interventional studies (e.g., type of lifestyle, length of follow-up, and type of cortisol measure), meta-analyses were only conducted for observational studies. Where data were sufficient and available, meta-analyses using random effects models were conducted for the correlation between long-term levels of cortisol and BMI, physical activity, perceived stress, anxiety, and depression. Effect sizes were expressed using Pearson's r for all studies. If r values were available from the study, they were entered into the meta-analysis. Otherwise, the r values were calculated from the available statistics (e.g., p-values and sample sizes). When long-term cortisol levels were measured at multiple points in a study (e.g., different trimesters of pregnancy), reported correlations at each time point were treated as independent. When the correlations between cortisol and lifestyle or mood factors were reported separately for subgroups (e.g., men and women), the results for each group were treated as independent. Comprehensive meta-analysis software (version 13.0) was used to calculate the pooled correlations (ESr).

4.2.5 Assessment of risk of bias

The risk of bias was assessed using the Downs and Black checklist (Downs and Black, 1998). A systematic review of quality assessment checklists included an evaluation of the Downs and Black checklist and recognised it as being suitable for use in systematic reviews (Deeks et al., 2003). The Downs and Black checklist has been used in several recent systematic reviews of randomised controlled trials and non-randomised studies (Babu et al., 2016; Donnelly et al., 2016). The checklist contains 27 items in total that assess five domains of study reporting and design: reporting, external validity, internal validity (bias), internal validity (confounding), and power. For each item, a score of 1 was given if the answer was yes and 0 if the answer was no or if it was unable to determine an item. Items

4, 8, 14, 15, 19, 23, and 24 were intervention-related and were recorded as "Not Applicable" (N/A) for all observational studies. Item 27 was designed to assess whether negative findings from a study could be due to chance (Downs and Black, 1998). The original item, with a score of 0–5, can only be used when the sample size required for a clinically and statistically significant result has been reported. Therefore, item 27 was simplified to a score of 0 (no sample size calculation) or 1 (sample size calculation reported) (Huffer et al., 2017). The assessment of the risk of bias was independently conducted by two reviewers, with 93% agreement on the items. Discrepancies were recorded and discussed between the reviewers until agreement was achieved. Categorisation of the total risk of bias scores was adapted from Huffer et al. (2017) and is presented in Supplementary Appendix Table S5 (A.2.2).

4.3 Results

4.3.1 Study characteristics

The database search resulted in 4,971 full titles. By 2020, November 21, after title and abstract screening and full-text assessment, a total of 52 publications, including 25 observational and 27 interventional studies, were included in the review and the results reported. The flow diagram of the screening and assessment process is presented in Figure 4.2. Studies included in the review were conducted between 1999 and 2018 across 16 countries, including the US, Brazil, China, and several European countries. The total sample size for observational studies was 4,087 (ranging from n = 23 to n = 768; median sample size = 164). The total sample size for interventional studies was 1,505 (ranging from n = 10 to n = 135; median intervention size = 30). A summary of study characteristics is presented in Supplementary Appendix A (A.1.1)

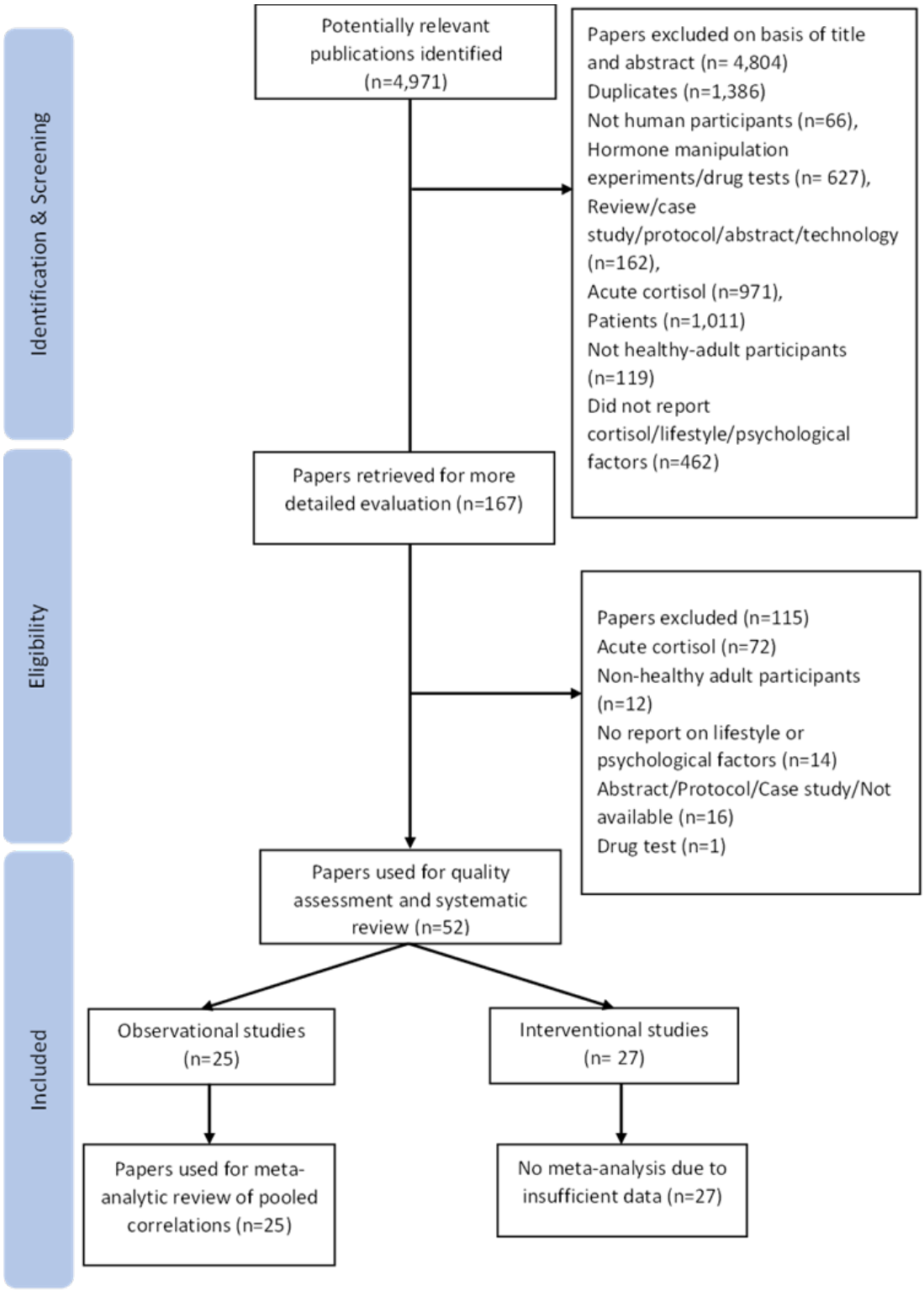


Figure 4.2 The PRISMA flow diagram of systematic review search and study selection

Results: risk of bias

For the current review, the maximum score on the checklist was 20 for observational studies and 27 for interventional studies, including four subsections: quality of reporting (a maximum score of 8 for observational studies and 10 for interventions), external validity (a maximum score of 3 for all studies), internal validity bias (a maximum score of 4 for observational studies and 7 for interventions), and confounding and selection bias (a maximum score of 4 for observational studies and 6 for interventions). Not applicable was recorded on the checklist where the item did not correspond well with the study design. The average total score was 9 out of 20 for observational studies and 11 out of 27 for interventional studies. The risk of bias score for observational studies was in the range 5–13. 17 observational studies scored >8 and were considered of moderate quality (Booij et al., 2015; Braig et al., 2016; Byun, 2013; Caparros-Gonzalez et al., 2017; Chen et al., 2015; Cieslak et al., 2011; Faresjö et al., 2013; Geng et al., 2016; Gerber et al., 2013; Grassi-Oliveira et al., 2012; Heinze et al., 2016; Huang et al., 2012; Kramer et al., 2009; Stoppelbein and Greening, 2015; Tsiouli et al., 2014; Ullmann et al., 2017; Wang et al., 2016). The risk of bias scores were 7–17 for interventional studies. Four interventional studies scored ≥ 15 and were considered of moderate quality (Chen et al., 2017; Daubenmier et al., 2011; Limm et al., 2011; Tsiouli et al., 2014). Therefore, the overall quality was considered moderate for observational studies and poor for interventional studies. The summary of the Downs and Black checklist score for each study and the detailed breakdown of the score's components are shown in Supplementary Appendix A2.

4.3.2 Population

The 25 observational studies were conducted on a population of 4,087 participants, including 3,269 females and 818 males. Eight studies were carried out on

the general population, for example, university students, teachers, and employees (Faresjö et al., 2013; Garcia-Leon et al., 2018; Gerber et al., 2013; Manenschijn et al., 2011; O'Brien et al., 2013; Qi et al., 2014; Sumra and Schillaci, 2015; Ullmann et al., 2017). Seven studies measured women's cortisol levels during pregnancy (Braig et al., 2016; Caparros-Gonzalez et al., 2017; Giesbrecht et al., 2013; Hoffman et al., 2016; Kramer et al., 2009; Larsen et al., 2016; Wikenius et al., 2016). Three studies were conducted on caregivers of children or other close relatives who had or were at risk of developing medical conditions, such as cancer or diabetes (Byun, 2013; Chen et al., 2015; Stoppelbein and Greening, 2015). Two studies were conducted on alcohol- or drug-dependent populations (Geng et al., 2016; Grassi-Oliveira et al., 2012). Five studies were conducted on individuals with psychological problems, such as depression or PTSD (Booij et al., 2015; Cieslak et al., 2011; Heinze et al., 2016; Schalinski et al., 2015; Wang et al., 2016).

The 27 interventional studies were conducted on a total of 1,505 participants, including 715 females and 600 males. The majority of studies ($k = 17$) were interested in a specific population, including people dependent on alcohol or cigarettes (Coiro et al., 2007; Goldberg et al., 2014; Sinha, 2011), caregivers (Brown et al., 2016; Goldberg et al., 2014), physically active populations (e.g., soldiers and athletes) (Diment et al., 2012; Mangine et al., 2016; Papacosta et al., 2013; Tanskanen et al., 2011), people with high stress levels (Bergen-Cico et al., 2014; Daubenmier et al., 2011; Huang et al., 2012; Marcus et al., 2003; Sannes et al., 2015), overweight populations (Tam et al., 2014), people with insomnia (Miller et al., 2015), and pregnant women (Chen et al., 2017). Ten studies focused on the general population, for instance, healthy students, couples, and community samples (Batista et al., 2015; Brown et al., 2002; Holt-Lunstad et al., 2008; Jacobs et al., 2013; Levine et al., 2007; Limm et al., 2011; Lynch et al., 2011; Timon et al., 2013; Torres et al., 2008).

4.3.3 Type of cortisol samples and time frame investigated

The majority of the observational studies relied on a single cortisol collection method ($k = 24$), including 19 studies using only hair samples (Braig et al., 2016; Caparros-Gonzalez et al., 2017; Chen et al., 2015; Faresjö et al., 2013; Garcia-Leon et al., 2018; Geng et al., 2016; Gerber et al., 2013; Grassi-Oliveira et al., 2012; Heinze et al., 2016; Hoffman et al., 2016; Kramer et al., 2009; Larsen et al., 2016; Manenschijn et al., 2011; O'Brien et al., 2013; Qi et al., 2014; Schalinski et al., 2015; Ullmann et al., 2017; Wang et al., 2016; Wikenius et al., 2016), five studies using only saliva samples (Booij et al., 2015; Byun, 2013; Cieslak et al., 2011; Giesbrecht et al., 2013; Stoppelbein and Greening, 2015), and one study using both hair and urine samples (Sumra and Schillaci, 2015).

Of the 25 observational studies, two studies measured cortisol levels over one month (four weeks) using hair samples (Geng et al., 2016; Qi et al., 2014), one measured HCC over one to two months (Larsen et al., 2016), one measured HCC over one to three months (Wikenius et al., 2016), 12 measured HCC over three months (Braig et al., 2016; Chen et al., 2015; Faresjö et al., 2013; Garcia-Leon et al., 2018; Gerber et al., 2013; Grassi-Oliveira et al., 2012; Manenschijn et al., 2011; O'Brien et al., 2013; Schalinski et al., 2015; Sumra and Schillaci, 2015; Ullmann et al., 2017; Wikenius et al., 2016), including one study using 15 morning urine samples (one sample taken every other day) to capture the cortisol profile over 30 days (Sumra and Schillaci, 2015), and one study measured HCC in pregnant women across all three trimesters of pregnancy (Garcia-Leon et al., 2018). For the study that measured cortisol in both hair and urine (Sumra and Schillaci, 2015), the urine results were not included in this review because the time frame captured by the urine samples was less than four weeks. One study measured HCC over six months (Heinze et al., 2016), and three studies measured HCC over 9–10 months (i.e., over the three trimesters of pregnancy) (Caparros-Gonzalez

et al., 2017; Hoffman et al., 2016; Kramer et al., 2009). Of the five studies measuring salivary cortisol, two studies measured salivary cortisol over one month (Booij et al., 2015; Byun, 2013), one measured over three months (Cieslak et al., 2011), one measured over the three trimesters of pregnancy (approximately 40 weeks) (Giesbrecht et al., 2013), and one measured over 12 months (Stoppelbein and Greening, 2015).

Most of the interventional studies ($k = 19$) used saliva samples to examine the changes in cortisol levels over time. One study used both blood and saliva samples to measure cortisol levels over eight weeks and four months, respectively (Daubenmier et al., 2011). Several studies examined salivary cortisol levels over a varying number of weeks: Three studies examined salivary cortisol levels over four weeks (Bergen-Cico et al., 2014; Holt-Lunstad et al., 2008; Lovallo et al., 2005), two studies over five weeks (Leivadi et al., 1999; Papacosta et al., 2013), one study over six weeks (Batista et al., 2015), and six studies over eight weeks (Huang et al., 2012; Jacobs et al., 2013; Manenschijn et al., 2011; Marcus et al., 2003; Tsiouli et al., 2014). One study measured salivary cortisol over three months (Brown et al., 2016). Three studies measured salivary cortisol levels for 10, 14, and 20 weeks, respectively (Torres et al., 2008; Sannes et al., 2015; Chen et al., 2017), and one study measured salivary cortisol levels for approximately one year (Limm et al., 2011). Six studies measured blood samples and reported changes in serum cortisol levels over time. Other studies measured serum cortisol levels over various time frames: one study for four weeks (Coiro et al., 2007), two for approximately five weeks (Miller et al., 2015; Sinha, 2011), one for seven weeks (Tanskanen et al., 2011), one for eight weeks (Diment et al., 2012), and one for 11 weeks (Brown et al., 2002). Two other studies used different sampling methods: One study used urine samples to measure cortisol levels for eight weeks (Timon et al., 2013), and the other measured cortisol levels over three months using hair samples (Goldberg et al., 2014).

Details of how cortisol samples were collected in these studies are presented in Supplementary Appendix A (A.1.2).

4.3.4 Lifestyle correlates of cortisol

A total of 13 observational studies reported the relationship between long-term levels of cortisol and lifestyle factors or variables indirectly related to lifestyle, namely BMI (Chen et al., 2015; Garcia-Leon et al., 2018; Geng et al., 2016; Heinze et al., 2016; Jacobs et al., 2013; Larsen et al., 2016; Manenschijn et al., 2011; Qi et al., 2014; Wang et al., 2016), physical activity (Booij et al., 2015; Garcia-Leon et al., 2018; Gerber et al., 2013; Sumra and Schillaci, 2015), sleep (Wang et al., 2016), smoking (Booij et al., 2015; Chen et al., 2015; Faresjö et al., 2013; Heinze et al., 2016), alcohol consumption (Booij et al., 2015; Heinze et al., 2016), caffeine consumption (Heinze et al., 2016), and diet (Heinze et al., 2016). Of the 13 studies, one study treated lifestyle factors as confounders while analysing the relationship between depression and cortisol levels instead of directly assessing the relationship between lifestyle factors and cortisol levels; therefore, it will not be discussed in detail in this section. However, the findings of this study will be reported in the next section (see Mood Correlates of Cortisol) (Booij et al., 2015). All of the observational studies measured long-term cortisol levels in hair.

A total of 15 interventional studies examined the effects of different lifestyle interventions on cortisol. These lifestyle interventions covered physical activity or exercise (Batista et al., 2015; Chen et al., 2017; Mangine et al., 2016; Papacosta et al., 2013; Tanskanen et al., 2011; Timon et al., 2013); diet (Brown et al., 2002; Diment et al., 2012; Tam et al., 2014; Torres et al., 2008), including one weight-loss diet intervention; sleep (Miller et al., 2015); smoking (Goldberg et al., 2014); caffeine consumption (Lovallo et al., 2005); and alcohol consumption (Coiro et al., 2007; Sinha, 2011). Of the 15 interventions, seven studies measured cortisol in blood (Brown et al., 2002; Coiro et al., 2007; Diment et al., 2012; Mangine et al.,

2016; Miller et al., 2015; Sinha, 2011; Tanskanen et al., 2011), six measured cortisol in saliva (Batista et al., 2015; Chen et al., 2017; Lovallo et al., 2005; Papacosta et al., 2013; Tam et al., 2014; Torres et al., 2008), one measured cortisol in urine (Timon et al., 2013), and one measured cortisol in hair (Goldberg et al., 2014). A summary of the studies investigating long-term cortisol levels and lifestyle factors is presented in Supplementary Appendix A (A.1.2).

BMI and body weight

A total of eight observational studies examined the association between BMI or bodyweight and long-term cortisol levels using HCC. In four studies, participants' BMI or bodyweight were measured by researchers during their research visit (Chen et al., 2015; Larsen et al., 2016; Manenschijn et al., 2011; Wang et al., 2016), two studies asked participants to self-report their BMI (Garcia-Leon et al., 2018; Qi et al., 2014), and the two remaining studies did not describe how BMI was measured (Geng et al., 2016; Heinze et al., 2016). Two studies measuring HCC over 1–2 months (Geng et al., 2016; Qi et al., 2014) did not find a significant association between BMI and HCC (n = 39–51), but one study with a larger sample size (n = 532) found a positive correlation between HCC and BMI (Larsen et al., 2016). Of the four studies measuring HCC over three months, three studies did not find significant associations between either self-reported or measured BMI and HCC (n = 62–85) (Chen et al., 2015; Garcia-Leon et al., 2018; Wang et al., 2016), but one study found a positive correlation between elevated HCC and increased BMI, which was objectively measured in a larger cohort (n = 122) (Manenschijn et al., 2011). One study (n = 58) measured HCC over six months and did not find a significant association between BMI and HCC (Heinze et al., 2016) but did not report how BMI was measured.

A meta-analysis (see Figure 4.3), which included data from six of the eight observational studies (Garcia-Leon et al., 2018; Geng et al., 2016; Larsen et al., 2016;

Manenschijn et al., 2011; Qi et al., 2014; Wang et al., 2016), showed a significant but small pooled correlation between BMI and chronic cortisol levels measured in hair (ESr = 0.148, $p < 0.001$, 95% CI: 0.213–4.301). However, the two studies that were not included in the meta-analysis because of insufficient data both reported no association between HCC and BMI (n = 85, (Chen et al., 2015) and n = 58, (Heinze et al., 2016)).

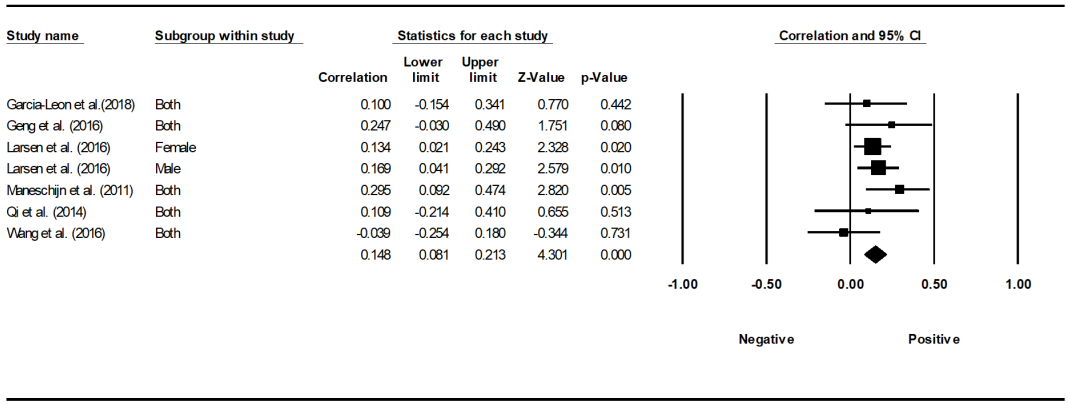


Figure 4.3 Random effect meta-analysis for the correlation between long-term levels of cortisol and BMI

An intervention examining the effect of a weight-loss programme on the HPA axis in overweight individuals was reviewed for the association between weight and cortisol. This intervention aimed to investigate whether a 24-week moderate caloric restriction by diet alone or a calorie-restricted diet with an aerobic exercise intervention could decrease salivary cortisol levels among overweight individuals (Tam et al., 2014). Intervention participants were randomised to either the diet group, with a structured 25% reduction in energy intake; the diet and exercise group, with a structured 12.5% reduction in energy intake and 12.5% increase in exercise energy expenditure; or the control group with a healthy weight-maintenance diet for six months. Saliva samples were collected eight times a day at baseline and at the end of the intervention. Body composition was measured, and designed food was provided to participants. Tam et al. (2014) reported that in the diet groups (both with and without exercise), significant weight loss was observed (approximately 10% in both intervention groups). However, the morning and

diurnal cortisol levels were significantly higher at week 24 when compared with baseline in both intervention groups. Similar increases in morning cortisol levels were seen in the control group at week 24. No significant group, time, or sex effects on morning and diurnal cortisol levels were found. Therefore, they concluded that although body composition improved in the intervention groups, a 10% weight loss with a 25% calorie-restricted diet alone or with exercise did not impact morning or diurnal salivary cortisol levels (Tam et al., 2014).

In summary, out of eight observational studies, only two studies reported a significant positive association between cortisol measured in hair and BMI (Larsen et al., 2016; Manenschijn et al., 2011), which was supported by the results of the meta-analysis. However, the intervention showed that although a 25% calorie-restricted diet alone or with exercise contributed to significant weight loss, it did not induce changes in cortisol levels (Tam et al., 2014).

Diets

Three other studies examined diet interventions, including one study with diet or diet and exercise as the interventions (Brown et al., 2002; Diment et al., 2012; Torres et al., 2008). The intervention diets and time frames were different across these studies. Two of the studies measured cortisol in blood (Brown et al., 2002; Diment et al., 2012) and one measured cortisol in saliva (Torres et al., 2008).

In premenopausal women, Brown et al. (2002) compared the hormonal effects of a Western diet high in saturated fat (control diet) with two other diets: the control diet plus soy protein (soy diet) and the control diet with polyunsaturated fat (PUFA diet). The aims of the study were to 1) examine the effects of specific dietary fatty acids on hormones related to breast cancer risk in premenopausal women and 2) examine whether soy can moderate the effects of certain fatty acids on these same hormonal factors. The hypothesised interventional effect on cortisol

was not specified. The participants were randomised to one of the three diets for two menstrual cycles followed by a washout time of two menstrual cycles before crossing over to the next diet–washout treatment period of four menstrual cycles. A total of 14 out of 26 participants who completed the intervention underwent three different treatments for a total length of 11 menstrual cycles. Blood samples were collected for two consecutive days during the second menstrual cycle of each treatment period and during the baseline month pre-intervention. Designed meals were provided to participants with one self-selected meal during each weekend. Records for the free meal were acquired. Brown et al. (2002) found that the PUFA diet induced a decrease in cortisol when compared with the control diet, but no significant difference in cortisol was found between the control diet and soy diet. Changes in BMI were not observed in either of the intervention groups (Brown et al., 2002).

Diment et al. (2012) compared the effects of daily nutritional supplements on immune indices in soldiers undertaking an eight-week training programme. The purpose of the study was to investigate whether providing additional energy intake to reduce the energy deficit would prevent a decrease in immune function during military training. Several immune indices and mediators were investigated, including plasma cortisol. They hypothesised decreased levels of circulating cortisol. The control group was asked to follow a habitual diet while the intervention group received a daily mixed nutritional supplement in addition to the habitual diet to reduce the daily energy deficit during arduous training. Blood samples were collected three times a day at baseline, six weeks later at a pre-field exercise, and eight weeks later at a post-field exercise. All food packaging was collected from the participants and uneaten food was weighed. They found that the intervention significantly prevented training-induced decreases in several immune indices and mediators. Body mass loss, lean mass loss, and fat mass loss induced by energy deficiencies were significantly lower in the intervention group than the control group. However, although they hypothesised decreased cortisol

levels post-exercise following the intervention when compared with baseline, they did not find an intervention effect on plasma cortisol (Diment et al., 2012).

Torres et al. (2008) compared the effects of a low sodium, high potassium diet, a high calcium, moderate sodium diet (HC), and a high potassium, high calcium diet (OD) on mood over ten weeks. All participants received two four-week dietary interventions (OD, low sodium, high potassium, or HC) preceded by a two-week control diet. On the day before the participants' visits with the study staff, the staff acquired a 24-hour dietary record from each participant. They assessed both self-reported mood, as measured by the Profile of Mood States, and cortisol, as a biomarker of stress and HPA axis activity. Participants were asked to rate their mood states every week. Saliva samples were collected three times a day at the end of each dietary period. The hypothesised changes in cortisol were not specified. Torres et al. (2008) reported that improvements in depression, tension, vigour, and mood were seen in the low sodium, high potassium diet when compared with the OD diet. Therefore, they concluded that a low sodium, high potassium diet had a positive effect on mood. However, the effect of this diet on cortisol levels was not reported. They only reported that higher cortisol levels were positively associated with greater vigour, lower fatigue, and higher levels of urinary potassium and magnesium. However, as higher potassium was a feature of both the low sodium, high potassium diet, and the OD diet, it is difficult to interpret these findings. A significant reduction in weight was observed during the low sodium, high potassium diet period, and a significant increase in weight was observed during the HC diet period (Torres et al., 2008).

In summary, two out of the three dietary interventions were considered successful in achieving their primary purposes. However, these dietary interventions seemed to have no effect on long-term cortisol levels. Therefore, definitive conclusions cannot be drawn due to the significant heterogeneity of these studies.

Physical activity

Three observational studies examined the association between physical activity and HCC, with HCC measured over three months in all three studies (Garcia-Leon et al., 2018; Gerber et al., 2013; Sumra and Schillaci, 2015). However, none of the studies measured physical activity for three months (i.e., the same period as cortisol). Two of the studies measured physical activity in participants for a typical week, including one study (n = 46) that assessed participants' general activity levels using accelerometers for seven consecutive days (Gerber et al., 2013) and one study (n = 31) that asked participants to report their frequency of exercise per week (Sumra and Schillaci, 2015). Physical activity was found to be positively correlated with HCC in the study using accelerometers (Gerber et al., 2013). Garcia-Leon et al. (2018) did not measure levels of physical activity among participants. Instead, participants were asked to identify themselves as regular or non-regular participants in physical exercise in general. This study had a large sample size (n = 591) and found a positive correlation between physical activity and HCC in the general population. However, no association was found in a subgroup of pregnant women. A meta-analysis of two of the three studies (Gerber et al., 2013; Garcia-Leon et al., 2018) showed a significant but small positive pooled correlation between physical activity and HCC (ESr = 0.160, $p < 0.001$, 95% CI: 0.235–4.043; see Figure 4.4).

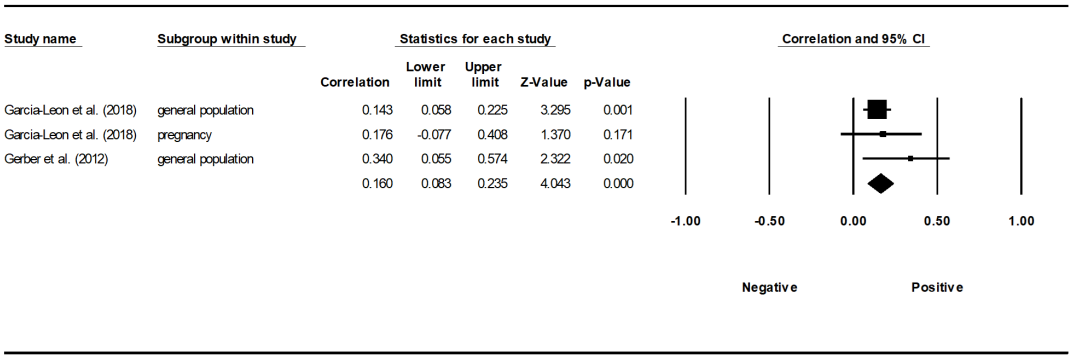


Figure 4.4 Random effect meta-analysis for the correlation between long-term levels of cortisol and physical activity

Six interventions examined the effects of physical activity or exercise on cortisol (Batista et al., 2015; Chen et al., 2017; Mangine et al., 2016; Papacosta et al., 2013; Tanskanen et al., 2011; Timon et al., 2013). However, these interventions all targeted different populations (e.g., students, pregnant women, athletes, or soldiers) with different intervention lengths (5–20 weeks) and examined cortisol in different forms (i.e., saliva, blood, and urine). Two studies examined the effects of yoga interventions (Batista et al., 2015; Chen et al., 2017) and reported similar outcomes for immediate changes in cortisol levels after yoga interventions but different results on the changes to long-term cortisol levels. Batista et al. (2015) conducted a six-week yoga programme (50 minutes per session twice a week). The objective was to evaluate the effects of acute and chronic yoga practice based on the hypothesis that yoga practice would reduce stress. The hypothesised changes in cortisol were not clearly described. Saliva samples were collected to measure cortisol four times a day before and after the yoga intervention over six weeks. Self-reported perceived stress was also measured pre- and post-intervention. The results showed a decrease in cortisol levels immediately after the yoga intervention, indicating the short-term effect of the intervention. Over the long term, cortisol levels in saliva were higher in the morning and decreased during the day, reaching the lowest levels before sleep, which was different to baseline cortisol levels where participants experienced an increase in cortisol before sleep. They also found reductions in perceived stress among participants after the intervention (Batista et al., 2015).

Chen et al. (2017) evaluated the effects of prenatal yoga on stress and immunity salivary biomarkers from 16–36 weeks' gestation. The study was based on the hypothesis that yoga would affect stress, which may be reflected by reductions in biomarkers of stress, such as cortisol. Cortisol was the only measure of stress used in the study. The yoga intervention consisted of 70-minute sessions twice per week for 20 weeks. Saliva samples were collected before and after yoga sessions every four weeks from baseline. They found decreased levels of salivary cortisol

immediately after yoga and higher levels of long-term cortisol in the intervention group. However, they did not find a significant long-term effect of yoga on salivary cortisol levels.

Findings from both studies suggest that yoga practice could decrease cortisol levels over the short-term. Although it may reduce individuals' distress (Batista et al., 2015) and enhance pregnant women's immune function (Chen et al., 2017), the effects of yoga on long-term cortisol levels were unclear.

Four studies examined the effects of exercise or training programmes over 5–8 weeks (Mangine et al., 2016; Papacosta et al., 2013; Tanskanen et al., 2011; Timon et al., 2013). Two of the studies measured cortisol in blood, one in saliva, and one in urine. Mangine et al. (2016) conducted a resistance exercise training of four sessions per week for eight weeks to examine the relationship between the endocrine response to resistance exercise and muscle hypertrophy in resistance-trained men. Participants were asked to complete at least 28 sessions throughout the study period. The hypothesised changes in cortisol were not clearly described. Blood samples were collected before and after exercise at baseline and eight weeks after training. After week one of training, when compared with the high-intensity, low-volume group, higher cortisol levels in response to exercise in the high-volume, moderate-intensity training group were observed. The difference in cortisol levels between groups was still significant after eight weeks of training. Therefore, the intervention had a significant impact on muscle size and growth, but cortisol was not significantly involved in the pathway (Mangine et al., 2016).

Papacosta et al. (2013) conducted a study to identify the change in selected salivary hormonal and immunological markers, mood state, and performance capacity related to judo in order to investigate physiological responses during tapering (i.e., gradual reduction in the training load) in judo. One of the aims was to examine the diurnal variation in cortisol during the tapering phase. However, the expected direction of the effects on cortisol was not reported. Judo athletes who

took part in the study trained as normal in week one, trained at a higher intensity in weeks 2–3, and tapered to half the normal training intensity in weeks 4–5. Two saliva samples were collected on the same day at the end of each week during weeks 1–3 and every day during weeks 4–5. They observed that evening cortisol fell below baseline at the beginning of the tapering week and then returned to baseline by the end of the tapered training. They also reported increased morning and evening salivary immunoglobulin A secretion rate; enhanced mood, muscle soreness, and overall performance among athletes after taper training; and decreased evening cortisol levels (Papacosta et al., 2013).

Tanskanen et al. (2011) examined the effects of an eight-week military training programme and the differences in training responses among participants based on whether they were overreaching or not. Overreaching can indicate that training is too strenuous (Tanskanen et al., 2011) and was defined as an accumulation of training and/or non-training stress resulting in short-term decrements in performance capacity that may be accompanied by related physiological and psychological signs and symptoms of maladaptation (Meeusen et al., 2006). They did not clearly describe the hypothesised changes in cortisol. Blood samples were taken three times a day (i.e., basal, pre-exercise, and post-exercise) at baseline and at the end of the training. They found a significant effect of exercise on several hormones and physiological indices, including serum cortisol. Specifically, basal cortisol decreased from week four to week seven of the training. However, in participants who were overreaching, there was evidence of significantly higher basal serum cortisol after eight weeks of training when compared with non-overreaching participants (Tanskanen et al., 2011).

Timon et al. (2013) examined the effects of an eight-week strength training programme for women, with sessions three times a week and weekend rest, to evaluate the effects of strength training on urinary steroid concentrations across the menstrual cycle phases. No specific hypothesis for changes in cortisol were described.

The first morning urine samples were collected on two consecutive days during the menstrual, follicular, and luteal phases of the menstrual cycle for both the month before training and the month after training. They found that strength training may affect oestrogen and progesterone metabolism in women by decreasing their levels throughout the menstrual cycle. However, they did not observe any significant changes in cortisol levels in urine following the training (Timon et al., 2013).

In summary, the findings from the observational studies were somewhat supported by the results of the interventions. Four out of six trials involving physical activity reported significant changes in cortisol levels after intervention. However, results from the trials should be interpreted with caution. Although both yoga interventions reported reduced cortisol levels immediately after yoga practice, only one reported higher morning cortisol levels and lower evening cortisol levels in participants over the long term (Batista et al., 2015). Three trials reporting significant intervention effects were conducted among athletes or physically trained participants and either reported changes in cortisol levels or differences between groups related to different training intensities (Mangine et al., 2016; Papacosta et al., 2013; Tanskanen et al., 2011). Due to the variety of physical activities involved, the way they were measured, and the way cortisol levels were measured across the trials, it is hard to draw conclusions regarding whether physical activity increases or decreases long-term cortisol levels. Based on these findings, it is concluded that although physical activity is likely to be positively associated with cortisol levels, the causal relationship remains unclear.

Sleep

Only one observational study investigated the relationship between sleep and HCC over three months among female veterans (Wang et al., 2016). Sleep quality was measured using the PSQI (Buysse et al., 1989). The time frame over which

sleep quality was measured was not fully described. However, Wang et al. (2016) did not observe a significant relationship between self-reported sleep quality and HCC.

Only one intervention assessed the association between sleep and cortisol levels (Miller et al., 2015). Miller et al. (2015) conducted a five-week sleep restriction intervention for insomnia using assumptions based on previous evidence of the association between insomnia, increased cortisol secretions, and core body temperature measured overnight (Riemann et al., 2010; Bonnet and Arand, 2010; Lack et al., 2008). Therefore, the primary purpose of this study was to evaluate whether the intervention would result in changes in arousal, as measured by overnight plasma cortisol concentrations and core body temperature. Specifically, they hypothesised that the sleep-restriction therapy would reduce early night plasma cortisol and core body temperature. Minimum time in bed and a "go-to-bed" threshold time were set for each participant at the beginning of the intervention. All participants were instructed not go to bed until after the threshold time and only go to bed if they felt "sleepy tired". Up to five weekly phone calls were made to participants to review the minimum time in bed, which was modified weekly for each individual. To measure cortisol levels, blood samples were taken from each participant every half hour from 10 pm to 6 am the next day on two consecutive days, both at baseline and at the end of the intervention. Miller et al. (2015) found that eight out of eleven patients who completed the intervention experienced a decrease in insomnia severity after the treatment. However, contrary to the hypothesis, cortisol results from six patients showed that cortisol levels in the early morning were significantly higher post-intervention when compared with baseline, but no changes were observed in the pre-sleep phase or early part of the night post-intervention. They also observed a significant reduction in core body temperature during sleep post-intervention, which was in line with the original hypothesis (Miller et al., 2015).

In summary, there is currently insufficient evidence to report on the association between sleep and long-term cortisol levels from both observational and interventional studies. However, increased morning plasma cortisol levels were observed in a small group post-intervention in tiral.

Smoking

Three observational studies examined the association between smoking and HCC. Of the three studies, two measured HCC over three months and asked participants whether or not they were smokers (n = 85, Chen et al. (2015)) or daily smokers (n = 236, Faresjö et al. (2013)). Both studies did not find a significant association between smoking status and HCC. One other study measured HCC over 3-6 months and measured smoking and smoking-related problems in the past three months using a questionnaire: Alcohol, Smoking and Substance Involvement Screening Test (ASSIST, Humeniuk et al. (2008)).(Heinze et al., 2016) With a sample size of 58, they also did not find a significant association between smoking and HCC either over the past 3 month or the 3 month before. Meta-analysis for the association between smoking and HCC was not conducted due to insufficient data (Heinze et al., 2016).

Only one trial assessed the association between smoking and cortisol levels (Goldberg et al., 2014). Goldberg et al. (2014) conducted a mindfulness-based smoking cessation intervention among healthy smokers, to examine the effects of two behavioural smoking cessation treatments (i.e., mindfulness training for smokers and a cognitive-behavioural therapy) on cortisol levels. They hypothesised that negative affect and increased cortisol output under stress were related to increased risk of smoking relapse and that stress reduction interventions designed to cultivate emotion regulation skills and target social-emotional processes could aid smoking cessation and may be marked by lower levels of cortisol in hair. Participants were randomised to one of two intervention groups: mindfulness instruction (group

one) or relaxation and CBT strategies targeted at smoking cessation (group 2). Both groups had eight intervention meetings over seven weeks, made quit attempts approximately one month after the first meeting, and received nicotine patches for two weeks. To measure cortisol levels, a 3cm hair sample was taken from each participant at one-month post-quit study visit, to determine the HCC (1) during the month before the intervention, and (2) during the month post-quit. Goldberg et al. (2014) found that participants' HCC was lower at 1-month post-quit compared to the month before the intervention. Intervention participants who stayed abstinent at the 1-month post-quit study visit had significantly lower levels of HCC compared with that in participants who had relapsed post-quit. They also reported that participation in mindfulness training and smoking cessation itself may each be independently associated with a reduction in hair cortisol. Differences were not found between two intervention groups for either decreases in HCC or the prevalence of smoking abstinence at post-quit. Another variable measured by Goldberg et al. (2014) was negative affect (mood) over the past week at baseline and post-quit, with the Depression Anxiety Stress Scale (Lovibond and Lovibond, 1995). They reported that larger decreases in HCC were positively associated with greater reductions in negative affect at post-intervention (Goldberg et al., 2014).

To summarise, evidence from observational and interventional studies seem contradictory. None of the three observational studies reported a significant association between smoking and cortisol measured in hair, regardless of the sample size, measures of smoking, or whether HCC was measured over three or six months. However, the intervention included in the review showed a slightly different result. This trial found reduced HCC 1-month post-quit among participants of a smoking cessation intervention. They also reported an association between smoking cessation and reductions in HCC (Goldberg et al., 2014).

Alcohol Consumption

Only one observational study investigated the association between alcohol consumption and HCC and reported that there was no significant correlation between self-reported use of alcohol and HCC measured over six months (Heinze et al., 2016). In this study, the use of alcohol and alcohol-related problems over the preceding three months were measured using the Alcohol, Smoking and Substance Involvement Screening Test (Humeniuk et al., 2008). Two hair samples were taken from participants to measure HCC over the preceding three months and the three months prior to that (i.e., six months in total) (Heinze et al., 2016). With insufficient evidence, the association between alcohol consumption and long-term levels of cortisol is unclear.

Alcohol withdrawal and cortisol responses to tasks

The two interventional studies involving alcohol did not directly assess the effect of alcohol abstinence treatment on long-term levels of cortisol instead they examined alterations in the HPA axis in alcoholic patients in experimental conditions, after long-term alcohol withdrawal (Coiro et al., 2007; Sinha, 2011). Specifically, Coiro et al. (2007) examined cortisol response to exercise tasks in alcoholic patients who underwent an alcohol abstinence program for 8 weeks. Details of the intervention were not described. No specific hypothesis regarding the changes in cortisol was reported. At four, six, eight weeks after abstinence, all participants underwent an exercise task. Eight blood samples were taken from each participant on each test day before and after the exercise task. Their results showed that serum cortisol levels significantly rose during exercise in the healthy controls and peaked after 30 minutes after the exercise task. However, the cortisol levels among abstinent alcohol patients did not significantly change during the exercise task, after four weeks of alcohol abstinence. This indicated a significantly blunted cortisol response to exercise tasks. However, at eight weeks post-abstinence, the

cortisol responses to exercise tasks for the abstinent alcoholic patients were significantly higher than two weeks previously and was no longer distinguishable from non-alcoholic participants (Coiro et al., 2007).

Sinha (2011) examined cortisol responses to stress and alcohol cues among alcoholic participants who underwent a one-month in-hospital abstinence treatment. Details of the treatment were not fully described. One of their study aims was to examine whether participants undergoing treatment would show alterations in cortisol levels at 1-month abstinence. The direction of the effects on cortisol was not specified. Each participant underwent the stress and alcohol cue tasks twice: at baseline and at one-month-abstinence. Blood samples were taken from participants each time before and after the tasks. Their results showed that although there were no significant differences at baseline, alcoholic participants showed a lack of cortisol response to both stress and alcohol cues compared with non-alcoholic controls at 1-month-abstinence. This indicated impaired responsiveness of the HPA axis in alcoholic patients (Sinha, 2011).

In summary, the overall evidence for the association between alcohol consumption and long-term levels of cortisol is insufficient. The only observational study did not find a significant relationship. Both trials, on the other hand, reported evidence of blunted cortisol responses in alcoholic participants compared with non-alcoholic participants, indicating an adverse influence of alcohol on the HPA axis which can still be observed one-month after alcohol abstinence. However, whether these adverse effects can be observed pre-abstinence or after 2 months of alcohol abstinence is less clear (Coiro et al., 2007; Sinha, 2011).

Caffeine

The association between caffeine consumption and long-term levels of cortisol was not investigated by observational studies included in this review. Only one in-

terventional study examined the effect of a caffeine intervention on cortisol levels (Lovallo et al., 2005). This study was based on the assumption that the effects of caffeine on the HPA axis (e.g.,increasing cortisol) may diminish with daily consumption of caffeine (Lovallo et al., 1996, 2005). To investigate this assumption, Lovallo et al. (2005) examined cortisol responses to a caffeine challenge after controlled levels of caffeine intake. The intervention was conducted over 4 weeks. On each week, participants were instructed to abstain for 5 days from dietary caffeine but took caffeine capsules, followed by a test day, and a crossover day to buffer sudden changes in intake between study weeks. On the test day, each participant was administrated caffeine capsules three times and underwent a stress challenge between the first two administrations. In order to investigate the effect of different levels of caffeine intake on cortisol responses to caffeine challenges, three different maintenance doses were administrated to participants, to mimic a range of consumption commonly found in the US diet: none (0 mg/day), moderate (300 mg/day), high (600 mg/day). Eight saliva samples were collected from each participant on the test day each week, before and after the stress challenge and each administration of caffeine capsules. This study observed that as tolerance to caffeine developed with regular daily intake, the cortisol responses to caffeine on test day were reduced. Specifically, taking caffeine capsules on the test day after five days of abstinence resulted in a significant elevation of cortisol throughout the entire day. On the other hand, a significant elevation of cortisol was only seen in the group that had moderate caffeine intake of 300mg, but not 600mg, for five days. The elevation of cortisol in the 300mg group lasted only for about 6 hours on the test day. Based on their finding, Lovallo et al. (2005) concluded that daily caffeine intake caused a partial but not complete tolerance to caffeine's effects on cortisol secretion and that elevated cortisol levels may occur in the afternoon hours in those consuming repeated doses throughout the day.

To summarise, due to insufficient information available, the influence of caffeine intake on long-term levels of cortisol cannot yet be determined.

4.3.5 Mood correlates of cortisol

A total of 23 observational studies reported the relationship between chronic levels of cortisol and mood factors, including stress (Braig et al., 2016; Byun, 2013; Chen et al., 2015; Geng et al., 2016; Gerber et al., 2013; Grassi-Oliveira et al., 2012; Heinze et al., 2016; Hoffman et al., 2016; Kramer et al., 2009; O'Brien et al., 2013; Qi et al., 2014; Sumra and Schillaci, 2015; Ullmann et al., 2017), anxiety (Braig et al., 2016; Faresjö et al., 2013; Heinze et al., 2016), depression or psychological distress (Booij et al., 2015; Braig et al., 2016; Byun, 2013; Caparros-Gonzalez et al., 2017; Faresjö et al., 2013; Giesbrecht et al., 2013; Heinze et al., 2016; Kramer et al., 2009; Schalinski et al., 2015; Wang et al., 2016; Wikenius et al., 2016), PTSD (Cieslak et al., 2011; Schalinski et al., 2015; Stoppelbein and Greening, 2015), coping efficacy (Cieslak et al., 2011), psychological functioning (Heinze et al., 2016), mindfulness (Jacobs et al., 2013), and fatigue (Wang et al., 2016). A summary of the studies investigating long-term levels of cortisol and mood factors is presented in Supplementary Appendix A (A.1.3).

Perceived Stress

A total of ten observational studies investigated the association between perceived stress and long-term levels of cortisol, with ten using the Perceived Stress Scale (PSS, Cohen et al. (1983)) (Byun, 2013; Chen et al., 2015; Faresjö et al., 2013; Geng et al., 2016; Gerber et al., 2013; Heinze et al., 2016; Hoffman et al., 2016; Kramer et al., 2009; O'Brien et al., 2013; Sumra and Schillaci, 2015) and one also using a computer task (Implicit Association Test, IAT).(Geng et al., 2016) Of the ten studies using the PSS, nine studies measured perceived stress over the past month (Byun, 2013; Chen et al., 2015; Faresjö et al., 2013; Geng et al., 2016; Gerber et al., 2013; Heinze et al., 2016; Hoffman et al., 2016; Kramer et al., 2009; Sumra and Schillaci, 2015) while one other study tailored the time frame into 3 months, in accordance with their HCC measure.(O'Brien et al., 2013) Nine

studies measured cortisol in hair (Chen et al., 2015; Faresjö et al., 2013; Geng et al., 2016; Gerber et al., 2013; Heinze et al., 2016; Hoffman et al., 2016; Kramer et al., 2009; O'Brien et al., 2013; Sumra and Schillaci, 2015) and one measured cortisol in saliva.(Byun, 2013)

Five studies measured HCC over three months (Chen et al., 2015; Faresjö et al., 2013; Gerber et al., 2013; O'Brien et al., 2013; Sumra and Schillaci, 2015). Four of these did not find a significant association between perceived stress and HCC regardless of the sample size (Chen et al., 2015; Faresjö et al., 2013; O'Brien et al., 2013; Sumra and Schillaci, 2015) and one reported a weak negative correlation (Gerber et al., 2013). One study measured HCC over six months but also did not find it to be significantly correlated with perceived stress which was measured over the past month (Heinze et al., 2016). Only one of the ten studies measured both HCC and perceived stress over one month (Geng et al., 2016). They found that HCC was not significantly correlated with self-reported perceived stress (measured by PSS). However, when they measured implicit perceived stress using computer tasks (IAT), HCC became strongly correlated with perceived stress. The directions of these correlations were positive (Geng et al., 2016).

Two of the 10 studies measured HCC over longer periods of time in pregnant women. Hoffman et al. (2016) measured perceived stress using the PSS at 16, 22, 28, 34, and 40 weeks of gestation and HCC at 16, 28, and 40 weeks of gestation. They found that higher HCC in the first trimester was potentially correlated with higher levels of perceived stress at 40 weeks of gestation, and HCC in the second trimester was positively correlated with perceived stress at 16 weeks of gestation (Hoffman et al., 2016). Kramer et al. (2009) took maternal hair samples to measure HCC over the preceding nine months and asked participants to answer the PSS at a clinical visit at 24–26 weeks of gestation. They did not find any association between HCC and perceived stress.

The final study measured long-term levels of cortisol over one month in saliva (Byun, 2013), using the PSS to measure perceived stress over the same time. They found that salivary cortisol on waking was not correlated with perceived stress but salivary cortisol in the evening had a positive correlation with higher levels of perceived stress.

Mixed findings were reported for the association between self-reported perceived stress and long-term cortisol. Of the 10 studies included, two reported a significant positive association between perceived stress and cortisol levels (Geng et al., 2016; Hoffman et al., 2016) and one reported a significant negative association (Gerber et al., 2013). Despite the mixed observations, a meta-analysis, which included data from eight studies (Chen et al., 2015; Faresjö et al., 2013; Geng et al., 2016; Gerber et al., 2013; Heinze et al., 2016; Hoffman et al., 2016; O'Brien et al., 2013; Sumra and Schillaci, 2015), showed a significant and small pooled correlation between long-term levels of cortisol and perceived stress (ESr = 0.11, p=0.0015, 95% CI: 0.023–0.203).

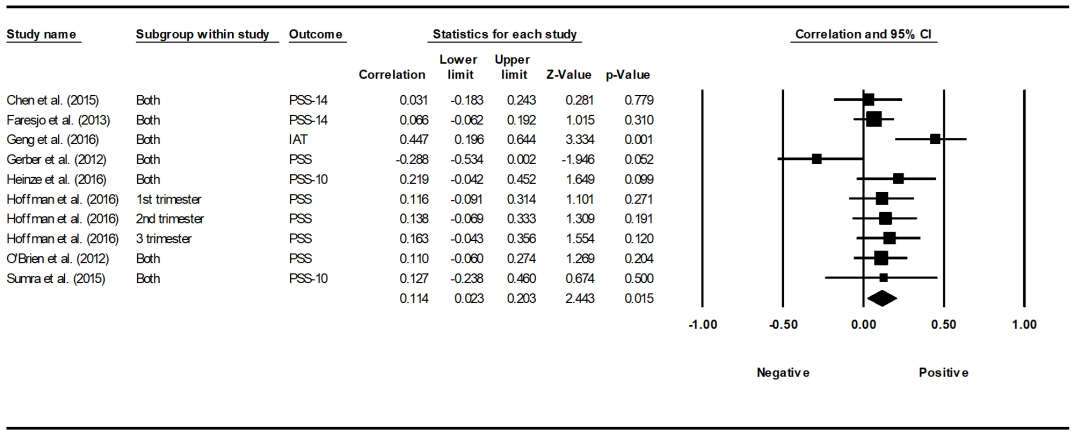


Figure 4.5 Random effect meta-analysis for the correlation between long-term levels of cortisol and perceived stress

Stressors

A total of three observational studies measured the impact of stressors on long-term levels of cortisol. The assessed stressors included negative life events in

the preceding 90 days (interview using the Psykel Life Events Scale, de Oliveira et al. (2003)) (Grassi-Oliveira et al., 2012), chaos in the neighbourhood over the preceding year (City Stress Index, Ewart and Suchday (2002)), chaos/order in the home environment without a specified time frame (Chaos, Hubbub, and Order Scale, Matheny Jr et al. (1995)) (O'Brien et al., 2013), and mentally burdening events (the Recalled Parental Rearing Behaviour Questionnaire, Penelo et al. (2012)) (Ullmann et al., 2017). All three studies measured HCC over three months (Grassi-Oliveira et al., 2012; O'Brien et al., 2013; Ullmann et al., 2017). Two of the three studies reported that HCC was positively correlated with self-reported stressful life events or mentally burdening events (Grassi-Oliveira et al., 2012; Ullmann et al., 2017). One study did not find significant associations between HCC and chaos, either in the neighbourhood or in the home (O'Brien et al., 2013).

A meta-analysis based on available data from two studies (Grant et al., 2013; O'Brien et al., 2013) showed that the association between stressors and HCC was small and not significant (ESr = 0.06, p= 0.288, 95% CI: 0.053–0.177). To summarise, two out of three studies found a significant positive association between HCC and stressors. However, this observation was not supported by the results of the meta-analysis.

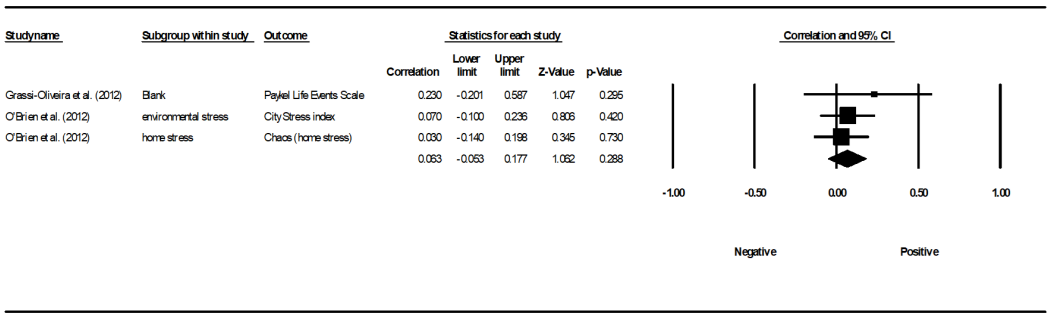


Figure 4.6 Random effect meta-analysis for the correlation between long-term levels of cortisol and mood impact of stressors

Other indices of stress

Three observational studies examined the association between HCC and other indices of stress, including chronic stress (Trier Inventory of Chronic Stress, Schlotz (2008)) (Braig et al., 2016), distress (Profile of Mood States, McNair and Heuchert (2005)) (Giesbrecht et al., 2013), and work stress (Effort–Reward Imbalance Scale, Li et al. (2006)) (Qi et al., 2014).

Two of the three studies examined the influence of stress in pregnant women by measuring self-reported chronic stress and distress (mood). However, the two studies measured cortisol in different ways. Braig et al. (2016) measured HCC over the three months before delivery in pregnant women and did not find a significant association between HCC and chronic stress over those three months. Giesbrecht et al. (2013) collected saliva samples from pregnant participants five times per day for two consecutive days at approximately 14, 20, and 32 weeks of gestation. Distress was measured at each sampling moment over the same period of time. They found that cortisol levels in saliva were positively correlated with distress among pregnant women (Giesbrecht et al., 2013). Qi et al. (2014) measured HCC using 1 cm hair samples and self-reported work stress over the preceding month and reported that neither the effort nor the reward scores were correlated with HCC. However, the effort–reward imbalance (1.83 $\times$ effort score $\div$ reward score) at work was positively correlated with HCC.

To summarise, two of the three studies reported a significant positive association between HCC and specific features of stress, namely distress and effort-reward imbalance at work. However, this observation was not supported by the meta-analysis of the three studies, which showed a non-significant pooled correlation between specific features of stress and HCC (ESr = 0.25, p= 0.097, 95% CI: 0.046–0.508).

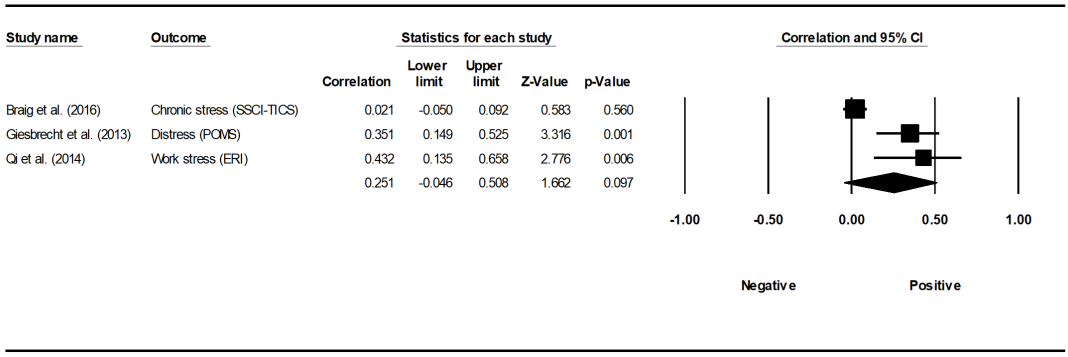


Figure 4.7 Random effect meta-analysis for the correlation between long-term levels of cortisol and other indices of stress

Anxiety

Five observational studies examined the association between self-reported anxiety and long-term levels of cortisol. All of the studies measured cortisol using hair samples of 3 cm, despite different overall lengths of studies. Five different scales were used to measure anxiety across the five studies including the Pregnancy-Related Anxiety Questionnaire (PRAQ, Braig et al. (2016)), the Hospital Anxiety and Depression scale (HADS, Braig et al. (2016); Faresjö et al. (2013)), a 4-item scale developed by Dunkel-Schetter (1998, Kramer et al. (2009)), the Overall Anxiety Severity and Impairment Scale (OASIS, Heinze et al. (2016)), the State-Trait Anxiety Inventory – State version (STAI-S, Hoffman et al. (2016)). Three studies with large sample sizes (n = 236-768) measured HCC over three months and did not find a significant association between HCC and pregnancy-related anxiety (over 3 months, Braig et al. (2016); not specified time frame, Kramer et al. (2009)), or between HCC and anxiety in general (not specified time frame, Faresjö et al. (2013)). One other study measured anxiety (at 16, 22, 28, 34, 40 weeks of gestation) and HCC (at 16, 28, 40 weeks of gestation) for the three trimesters (Hoffman et al., 2016). They found that self-reported state-trait anxiety measured at 16 and 28 weeks of pregnancy was correlated with HCC in both the second and the third trimester (Hoffman et al., 2016). One study measured HCC over 3-6 months and anxiety which was measured for a shorter time-frame

(approximately 3 months) within the forementioned 3-6 months. This study did not find a significant association between HCC and anxiety (Heinze et al., 2016).

A meta-analysis based on available data from three studies showed that the pooled correlation between long-term levels of cortisol and anxiety was small and not significant (ESr=0.08, p=0.137, 95%CI: -0.026, 0.187). In summary, the association between anxiety and long-term levels of cortisol seemed very mixed with most of the studies not supporting an association between self-reported anxiety and cortisol measured in hair. Results from the meta-analysis also pointed to the same conclusion.

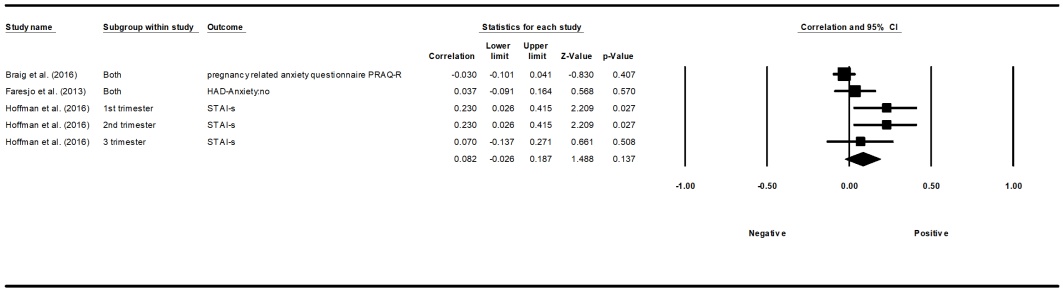


Figure 4.8 Random effect meta-analysis for the correlation between long-term levels of cortisol and anxiety

Depression and depressive symptoms

A total of 12 observational studies examined the association between long-term levels of cortisol and self-reported depression or depressive symptoms, of which ten studies measured cortisol in hair (Braig et al., 2016; Caparros-Gonzalez et al., 2017; Faresjö et al., 2013; Heinze et al., 2016; Hoffman et al., 2016; Kramer et al., 2009; Schalinski et al., 2015; Wang et al., 2016; Wikenius et al., 2016) and two measured cortisol in saliva (Booij et al., 2015; Byun, 2013). Seven tools were used to measure depression or depressive symptoms, including the Beck Depression Inventory (Booij et al., 2015; Gerber et al., 2013), the the Hospital Anxiety and Depression scale (Braig et al., 2016; Faresjö et al., 2013), the Patient Health Questionnaire (Byun, 2013), the Edinburgh Postnatal Depression Scale

(Caparros-Gonzalez et al., 2017; Wikenius et al., 2016), the Kessler Psychological Distress Scale (Heinze et al., 2016), the Quick Inventory of Depression Symptoms (Heinze et al., 2016), the Center for Epidemiologic Studies Depression Scale (Hoffman et al., 2016; Kramer et al., 2009; Wang et al., 2016), and the Clinician-administered Hamilton Rating Scale (Schalinski et al., 2015).

Five studies examined the association between HCC and depression either during pregnancy or postpartum (Braig et al., 2016; Caparros-Gonzalez et al., 2017; Hoffman et al., 2016; Kramer et al., 2009; Wikenius et al., 2016). Two studies reported that the depression scales covered the preceding week (Hoffman et al., 2016) and three months (Braig et al., 2016). The time frames were not specified in other studies (Caparros-Gonzalez et al., 2017; Kramer et al., 2009; Wikenius et al., 2016). Three studies measured HCC using hair samples of 3–9 cm from participants and did not find a significant association between HCC and depression, regardless of when the relationship was examined, i.e., in the second trimester (Wikenius et al., 2016), in the three months before delivery (Braig et al., 2016), or throughout the pregnancy (Kramer et al., 2009). Two other studies (Caparros-Gonzalez et al., 2017; Hoffman et al., 2016) took 3 cm hair samples from pregnant women at the end of the first, second, and third trimesters (i.e., in effect measuring cortisol from the three months prior to conception to the third trimester). In one study, depression was measured repeatedly (in each trimester), and in the other study, postpartum depression was measured shortly after delivery. Caparros-Gonzalez et al. (2017) found that HCC during the first and third trimesters was prospectively positively correlated with postpartum depression. Hoffman et al. (2016) found that HCC during the first trimester was prospectively positively associated with depression at 40 weeks of gestation, HCC in the second trimester was positively associated with depression at 16 and 28 weeks of gestation, and HCC in the third trimester was associated with depression at 16 and 28 weeks of gestation. In these two studies, the first trimester referred to approximately weeks 4–12 of pregnancy, the second trimester to weeks 13–26,

and the third trimester to weeks 26–40 (Cappuccio et al., 2010; Hoffman et al., 2016).

Three studies examined the association between HCC and self-reported depression in the general population (Faresjö et al., 2013; Gerber et al., 2013; Heinze et al., 2016). Two of the three studies measured HCC over three months using a 3 cm hair sample and found a significant correlation between HCC and depressive symptoms over an unspecified time frame. However, the correlation was negative in the one study (Gerber et al., 2013) and positive in the other (Faresjö et al., 2013). Neither study fully reported the time frame over which self-reported depression was measured. Heinze et al. (2016) measured HCC over six months using two 3 cm samples and measured depression over the same six months but did not find a significant association between HCC and depression over the preceding 7–30 days.

Two studies examined the association between HCC over three months and depression over and unspecified time frame in people with PTSD symptoms. Neither study found a significant association between HCC and depression (Schalinski et al., 2015; Wang et al., 2016).

Two studies examined the association between long-term levels of cortisol (measured in saliva over one month) and depression (Booij et al., 2015; Byun, 2013). Booij et al. (2015) took saliva samples from participants with depression and healthy controls three times a day (in the morning, six hours later, and another six hours later) for 30 days. However, Byun (2013) took saliva samples from participants twice a day (on waking and at 9 pm) at the beginning of the study and four weeks later. Booij et al. (2015) found that cortisol in saliva was higher and the daily cortisol slope steeper in the participants with depression when compared with the healthy controls. Byun (2013) found that although not significantly correlated with depression at the beginning of the study, elevated cortisol levels in saliva in the evening were positively correlated with self-reported depressive

symptoms four weeks later.

In summary, six out of 12 studies found a significant association between long-term levels of cortisol, either measured in hair or saliva, and depression. A meta-analysis was performed with the available data from seven studies and revealed a small but significant pooled correlation coefficient of 0.13 ($p = 0.022$, 95% CI: 0.019–0.243), suggesting that long-term levels of cortisol, either measured in hair or saliva, are correlated with depression.

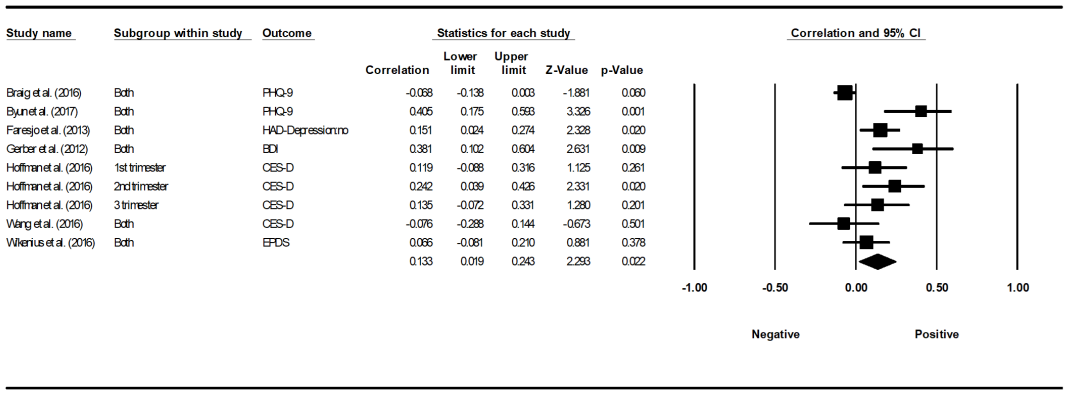


Figure 4.9 Random effect meta-analysis for the correlation between long-term levels of cortisol and depression

PTSD

Three observational studies examined the relationship between PTSD and long-term levels of cortisol (Cieslak et al., 2011; Schalinski et al., 2015; Stoppelbein and Greening, 2015). They all measured cortisol using different protocols and PTSD using different self-report instruments. Two studies measured cortisol in saliva, using diverse methodologies (Cieslak et al., 2011; Stoppelbein and Greening, 2015). Cieslak et al. (2011) collected saliva samples from participants four times a day (upon awakening and one hour, four hours, and 12 hours after waking) one week, one month, and three months after a traffic accident (i.e., trauma). The impact of trauma was measured using the Impact of Event Scale - Revised (Marmer and Weiss, 1997) one week, one month, and three months after the

accident. The impact of trauma one month after the accident was found to be negatively associated with AUCi (reflecting changes in daytime cortisol secretions) at three months, and the impact of trauma three months after the accident was found to be positively associated with AUCg (reflecting total daytime cortisol release) one month after the accident (Cieslak et al., 2011). Stoppelbein and Greening (2015) asked participants to provide saliva samples at 3–5 pm on the first day and 6–8 am and 3–5 pm on the following day. This saliva collection protocol was repeated monthly for 12 months. PTSD was assessed using self-report questionnaires, including the Life Events Checklist and PTSD Checklist. Participants were asked to complete the questionnaires monthly over the same 12-month period. Stoppelbein and Greening (2015) found that higher levels of cortisol were positively correlated with one type of self-reported PTSD symptoms, i.e., numbing symptoms, across the 12 months. Schalinski et al. (2015) measured HCC using 3 cm hair samples and clinically assessed the severity of PTSD symptoms. The time frame covered by the PTSD measure was not fully described in the study. They did not find a significant association between HCC and the severity of PTSD symptoms (Schalinski et al., 2015).

In summary, two out of three studies reported a significant positive association between PTSD and long-term levels of cortisol measured in saliva. One study did not find an association between the clinically assessed severity of PTSD symptoms and HCC. However, the data were insufficient for conducting a meta-analysis; therefore, it was not possible to interrogate the observations further.

4.3.6 Effects of mood interventions on cortisol

Mindfulness

Six studies conducted mindfulness interventions targeting psychological factors, including stress, anxiety, depression, and PTSD (Bergen-Cico et al., 2014; Brown

et al., 2016; Daubenmier et al., 2011; Jacobs et al., 2013; Lynch et al., 2011; Marcus et al., 2003). The studies measured acute levels of cortisol in saliva, and one study also measured cortisol in blood. The length of the mindfulness interventions varied from four weeks to four months. However, over half of the studies conducted mindfulness interventions over eight weeks, with mixed findings (Brown et al., 2016; Jacobs et al., 2013; Lynch et al., 2011; Marcus et al., 2003). The sample sizes of these four studies were n = 21–60.

Brown et al. (2016) conducted an eight-week mindfulness-based stress reduction intervention and made comparisons with a social support control group. The purposes of this intervention were to reduce psychological and neuroendocrine markers of stress, improve psychological morbidity, and improve the quality of caregiver–recipient relationships in family caregivers of persons in the early stages of Alzheimer's disease and other dementias. Participants attended weekly classes of 1.5–2 hours in length. Saliva samples were collected six times a day pre-intervention, post-intervention, and at the three-month follow-up. The mindfulness intervention was found to significantly improve participants' perceived stress levels and mood. However, there was no significant difference in the improvements in mood, perceived stress, and burden between the experimental group and the control group. Neither group exhibited significant changes in the diurnal cortisol response curve. Therefore, Brown et al. (2016) concluded that both mindfulness-based and social support interventions cause a reduction in stress, but mindfulness had no sustained psychosocial advantages over social support, and neither mindfulness nor social support had an intervention effect on cortisol levels.

Jacobs et al. (2013) conducted an eight-week mindfulness meditation intervention. The main purpose was to investigate whether self-reported mindfulness and changes in mindfulness following the intervention were related to resting cortisol. The hypothesised intervention effects on cortisol were not described. The

intervention group met twice per day for one-hour sessions of meditation practice. Saliva samples were collected twice a day for three consecutive days pre- and post-intervention. The resting cortisol was calculated as the average of the afternoon and bedtime cortisol levels. Jacobs et al. (2013) found that self-reported mindfulness significantly increased after the intervention, but there were no significant changes in cortisol levels. However, they found that levels of mindfulness were significantly and inversely associated with resting cortisol levels: larger increases in mindfulness were associated with an average decrese in levels of pre- to post-intervention resting cortisol, smaller increases or slight decreases in mindfulness were associated with an average increase in levels of pre- to post-intervention resting cortisol.

Lynch et al. (2011) conducted an eight-week mindfulness-based coping intervention with a waiting-list control group. The aim of this study was to explore the feasibility of the intervention, especially the sensitivity of salivary measures of biomarkers of stress, including cortisol. No specific hypothesis was given regarding the direction of the effects on cortisol. Participants attended one 1.5-hour meeting per week for eight consecutive weeks. Saliva samples were collected eight times throughout the day for two consecutive days pre- and post-intervention. They found that the mindfulness intervention significantly improved perceived stress, anxiety, and depression and increased the self-reported mindfulness of participants. However, they did not find any significant changes in cortisol levels or in alpha-amylase levels after the intervention (Lynch et al., 2011).

Marcus et al. (2003) conducted an eight-week mindfulness-based stress reduction intervention. The purpose of the study was to identify changes in the psychological and physiological stress response of intervention participants who were undergoing treatment for substance abuse in a therapeutic community. However, they did not describe the direction of intervention effects on cortisol levels. Participants attended a 2.5-hour mindfulness meditation session per week for eight

consecutive weeks. They were asked to practise mindfulness outside the sessions for 45–60 minutes six days per week. Saliva samples were collected four times on the first morning pre- and post-intervention. They found a non-significant decrease in self-reported perceived stress after the intervention. However, awakening salivary cortisol levels were significantly lower following the intervention (Marcus et al., 2003).

Bergen-Cico et al. (2014) conducted a four-week mindfulness-based stress reduction intervention for veterans with PTSD. The objective was to investigate whether there was a dose–response relationship between participation in the intervention and diurnal cortisol. They hypothesised that engagement in the intervention would reduce diurnal cortisol output and CAR among veterans with PTSD. A total of 19 participants were asked to attend a 90-minute mindfulness-based meditation group session per week for four consecutive weeks. An active control group was included consisting of veterans with similar PTSD and subthreshold PTSD receiving the typical primary care treatment. Saliva samples were collected five times a day for two consecutive days pre- and post-intervention. Significant changes in cortisol levels were found, which were associated with engagement and number of sessions of the intervention that were completed. In addition, participants who completed all four sessions had a significant decrease in CAR and an increase in AUC when compared with the control group. Changes in either PTSD symptoms or self-reported stress were not described (Bergen-Cico et al., 2014).

Daubenmier et al. (2011) conducted a four-month mindfulness intervention for stress eating with a waitlist control group and hypothesised that mindfulness training would enhance awareness of and responsiveness to bodily sensations and reduce psychological stress, emotional eating, and cortisol secretion, which would lead to reductions in abdominal fat. Treatment outcomes were three-fold: improvement in stress eating, reduction in cortisol, and reduction in abdominal

fat. A total of 24 participants were asked to attend a weekly 2.5-hour class for nine weeks and a seven-hour silent day of guided meditation practice after the sixth class. Cortisol was measured in both saliva and blood. Saliva samples were collected three times a day on four different days pre- and post-intervention. Fasting morning blood samples were taken pre- and post-intervention. The study found that participants in the intervention group showed significant improvements in mindfulness, anxiety, and externally based eating after the intervention when compared with control participants. However, differences between the intervention and control groups were not significant for emotional eating, average CAR, overall weight, abdominal fat, or fat distribution over time. Neither group exhibited significant changes in the cortisol slope or morning serum cortisol concentrations. However, in subgroup analyses, Daubenmier et al. (2011) found that obese participants in the intervention group had significant reductions in CAR and maintained body weight while obese participants in the control group had stable CAR and gained weight. They also reported associations between reduced abdominal fat and improvements in self-reported mindfulness, chronic stress, and CAR in the treatment group. Based on these findings, Daubenmier et al. (2011) concluded that the intervention was successful in increasing mindfulness and responsiveness to bodily sensations and reducing anxiety and eating in response to external food cues. However, this treatment did not reduce abdominal fat, fat distribution, or overall weight and only tended to reduce eating in response to emotions and cortisol levels.

In summary, half the mindfulness interventions resulted in reductions in cortisol levels following the interventions. Two of these studies reported a positive relationship between participation in the intervention and reductions in cortisol levels. However, half of the interventions did not find significant treatment effects on cortisol. Based on these findings, it is unclear whether mindfulness-based stress interventions reduce cortisol levels.

Social and emotional support

Holt-Lunstad et al. (2008) conducted a four-week support enhancement intervention among healthy married couples with the primary purpose of examining whether warm touch enhancement could improve "physiological indices of mood and the stress system," including decreasing cortisol levels. A total of 20 couples (n = 40) in the intervention group were asked to attend two one-hour group sessions for couple contact enhancement techniques. The couples were instructed to practise the techniques for 30 minutes three times per week for four weeks. A behavioural monitoring control group was designed where subjects were instructed not to change anything about their normal behaviours with their spouse and to simply keep a diary of their physical affection and mood. Saliva samples were collected five times throughout a day pre- and post-intervention. Stress, mood, and social support were not measured psychologically. Holt-Lunstad et al. (2008) did not find intervention effects on salivary cortisol. However, they found improvements in other physiological indices, such as salivary oxytocin and alpha-amylase levels, in the intervention group. They also found significantly lower blood pressure in the husbands in the intervention group post-treatment. In summary, findings from this trial suggested that the social and emotional support intervention was unlikely to decrease salivary cortisol over time. However, the significant improvements in other physiological indices suggest pathways through which psychological interventions influence physical health other than the HPA axis.

Other stress management interventions

Three studies investigated the effects of other stress management methods, such as stress prevention seminars (Limm et al., 2011), cognitive behavioural sessions (Sannes et al., 2015), and relaxation and diaphragmatic breathing (Tsiouli et al., 2014) on chronic levels of cortisol. Although the lengths of the studies varied from eight weeks to one year, all the studies measured cortisol in saliva.

Limm et al. (2011) conducted a work-site stress management intervention over one year to examine the long-term effects of the intervention on psychological and biological reactions to work stress. The hypothesised intervention effects on cortisol were not described. The intervention consisted of two full-day group sessions followed by two refresher sessions within the next eight months. At each group session, participants were asked to remember and share individual stressful situations at work. Several tools for coping with difficult and stressful situations were introduced by the trainer at the group session. A waitlist (n = 20) was used as the control group. Saliva samples to measure cortisol were collected seven times throughout a day both pre-intervention and at the one-year follow-up. They found that perceived stress reactivity, which was the primary outcome, decreased in both groups, but a significantly greater reduction was seen in the intervention group at the one-year follow-up. A greater reduction in the activation of the sympathetic nervous system, measured by α-amylase, was also seen in the intervention group. However, there was no intervention effect on cortisol. For other self-reported psychological factors (i.e., depression, anxiety, and effort–reward imbalance), improvements were seen, which were greater in the intervention group than the control group, but the difference was not significant (Limm et al., 2011).

Sannes et al. (2015) conducted a 10-week cognitive behavioural stress management intervention with a one-month follow-up to examine whether the intervention could induce changes in the cortisol rhythm of healthy, stressed women who were at risk of breast cancer due to family history. The hypothesis was that women in the cognitive behavioural stress management group would demonstrate lower intra-individual cortisol variability (i.e., the degree to which an individual's cortisol output may be erratic on a given day). There were no other hypothesised cortisol outcomes. The intervention consisted of 10 weekly two-hour sessions covering topics of awareness of the effects of stress, cognitive reframing, cognitive coping skills training, etc. Saliva samples to measure cortisol were collected

five times a day for two consecutive days pre- and post-intervention and at the one-month follow-up. They found lower levels of intra-individual cortisol variability among intervention participants when compared with the waitlist control group. They also observed improvements towards a steeper cortisol slope and lower overall cortisol output (AUCg) in the intervention group. However, they did not find significant changes in other cortisol indices (i.e., awakening cortisol levels, evening cortisol levels, and CAR) following the intervention. Changes in self-reported stress were not fully described. Based on these results, Sannes et al. (2015) concluded that the 10-week cognitive behavioural stress management intervention contributed to a less erratic cortisol output throughout the day in stressed women at risk of breast cancer.

Tsiouli et al. (2014) conducted an eight-week stress management intervention with a healthy lifestyle education control group to evaluate the effectiveness of the intervention in reducing perceived and parenting stress, increasing internal locus of control, promoting a healthy lifestyle, and normalising cortisol levels in parents of children and adolescents with type 1 diabetes. However, neither the definition of normal levels of cortisol nor the hypothesised changes in cortisol post-intervention were well described. Participants in the intervention group were involved in a relaxation training technique for 37 minutes and were asked to perform these techniques twice a day for eight weeks. They were instructed to adopt a healthy lifestyle (instructions included recommendations on nutrition and physical activity). Saliva samples were collected five times throughout a single day pre- and post-intervention. They found a significant reduction in stress in the intervention group that was significantly different from the control group. They also found significant improvements in lifestyle factors, such as sleep, stress-related symptoms, and eating habits in both groups. However, no significant changes in the diurnal rhythm of cortisol were observed in either the intervention group or the control group (Tsiouli et al., 2014).

In summary, all of the stress management interventions reported significant treatment effects on participants. Two studies that assessed psychological or lifestyle outcomes reported reduced stress (Limm et al., 2011; Tsiouli et al., 2014), reduced alpha-amylase (Limm et al., 2011), and improved lifestyle (Tsiouli et al., 2014). However, neither study observed significant changes in cortisol levels post-intervention (Limm et al., 2011; Tsiouli et al., 2014). The study that only measured cortisol indices as outcomes reported improvements in cortisol indices, including cortisol slope, overall cortisol output, and intra-individual cortisol variability (Sannes et al., 2015). Therefore, at present the evidence for the effects of these stress-management interventions on cortisol is weak.

Alternative therapies

Two studies examined the effects of alternative therapies, including acupuncture (Huang et al., 2012) and massage (Leivadi et al., 1999). Both studies conducted the intervention over five weeks, measured cortisol in saliva, and reported changes in cortisol levels after the intervention.

Huang et al. (2012) conducted a five-week stress reduction acupuncture intervention with a relaxation-only group and a waitlist control group to examine whether traditional Chinese acupuncture to treat chronic stress could change diurnal salivary cortisol profiles in healthy individuals. The hypothesised intervention effects on cortisol levels were not described. The intervention group received weekly acupuncture sessions for five weeks. The relaxation-only group received weekly relaxation sessions without acupuncture for five weeks. The control group did not receive any treatment for five weeks. Saliva samples were collected four times a day on nine days: pre- and post-intervention and at follow-up. A decrease in perceived stress was observed in all groups; however, the differences pre- and post-intervention and between and within groups were not significant. Increased CAR was found in both the intervention and relaxation-only groups,

but the increases were not significant. No significant changes in cortisol levels or daily average cortisol concentrations were found between the groups following the intervention (Huang et al., 2012).

Leivadi et al. (1999) conducted a five-week massage intervention with a relaxation control group to examine whether massage therapy could reduce stress hormone levels and other potential stressors, such as anxiety and neck, shoulder, and back pain, among dance students. The massage intervention involved two 30-minute sessions per week for five weeks, and the frequency of the relaxation sessions was the same. Saliva samples were collected twice a day pre- and post-intervention. Both groups reported improvements in mood and anxiety after the sessions, but a decrease in salivary cortisol levels was found only in the intervention group. However, these changes were only observed immediately after the intervention. Changes in mood, anxiety, and cortisol levels were not observed over time. All participants reported a decrease in neck, shoulder, and back pain after the treatment sessions (Leivadi et al., 1999).

Two studies involving alternative therapies were reviewed. Only the massage therapy showed significant improvements in mood and anxiety among participants, but neither showed significant effects on cortisol levels.

4.4 Discussion

The main aim of this review was to systematically identify and review the available literature on the associations between lifestyle and mood factors and long-term levels of cortisol in healthy adults. The quality and scientific rigour of research conducted over the last two decades in this area were assessed. Overall, the evidence on associations between lifestyle and long-term cortisol levels was mixed, with considerable heterogeneity between studies in approaches to assessing cortisol, lifestyle, and psychological factors, with observational studies meeting

criteria for moderate quality, and interventional studies meeting criteria for poor quality.

Nonetheless, from this review, we can glean some evidence in support of a positive association between physical activity and long-term levels of cortisol, with a meta-analysis of observational studies showing a significant pooled correlation. Although the evidence from trials also seemed to suggest that physical activity is likely to induce changes in cortisol levels over time, the direction of the changes remained unclear due to mixed findings. The meta-analysis of observational studies also showed a significant pooled positive correlation between BMI/bodyweight and long-term cortisol levels. However, the only trial that aimed to reduce weight through diet with and without exercise did not find changes in cortisol levels. Other dietary interventions, regardless of whether they induced changes in body weight or body composition or not, were unlikely to induce changes in cortisol levels over the long term. Observational studies suggested that perceived stress and depression were likely to be positively associated with long-term cortisol levels. However, the observational evidence did not support a significant association between stressors, specific features of stress and anxiety, and cortisol levels. Posttraumatic stress disorder was shown to be positively associated with cortisol levels, although this conclusion was tentative due to insufficient data. Overall, the trial evidence examining the effects of psychological interventions on long-term cortisol levels was insufficient and mixed. Finally, evidence from both observational and interventional studies was insufficient to draw a clear conclusion on the associations between long-term cortisol levels and smoking, sleep, alcohol consumption, or caffeine consumption.

4.4.1 Heterogeneity in how lifestyle and mood were measured

One observation arising from this review relates to the heterogeneity in the measurement of lifestyle and mood across studies, which may have contributed to the mixed findings. For example, of the eight studies examining the association between BMI/bodyweight and cortisol, two studies collected self-reported BMI (Garcia-Leon et al., 2018; Qi et al., 2014), four studies measured BMI/bodyweight during research visits (Chen et al., 2015; Larsen et al., 2016; Manenschijn et al., 2011; Wang et al., 2016), and two studies did not describe how BMI was measured (Geng et al., 2016; Heinze et al., 2016). The three observational studies examining the association between physical activity and HCC measured physical activity differently (activity levels recorded for seven consecutive days by accelerometers, Gerber et al. (2013); self-reported frequency of exercise per week, Sumra and Schillaci (2015); and self-reported regular/non-regular participants of physical exercise, Garcia-Leon et al. (2018)). Therefore, it was difficult to make comparisons between studies. However, meta-analyses supported the correlation between BMI and cortisol and physical activity and cortisol, suggesting that these associations are likely to be robust and not likely to be due to chance.

The heterogeneity in lifestyle interventions was somewhat reflected by the variations in monitoring compliance to the interventions. For example, in the two yoga interventions (Batista et al., 2015; Chen et al., 2017), the definition and measurement of compliance to the intervention were not clearly described. There was heterogeneity in the measurement of compliance among the studies that reported compliance. For example, dietary assessments were completely different across the three dietary interventions. Brown et al. (2002) provided the participants with food and obtained food records relating to the weekly meal that they were allowed to choose freely during the intervention period. (Diment et al., 2012) collected all food packaging provided to the participants and weighed the un-

eaten food. (Torres et al., 2008) assessed dietary compliance using 24-hour urine collections from participants every two weeks. The uncertainty of the definition of adherence to lifestyle interventions and the variability in the measurement of compliance created difficulties in comparing results across different interventions and influenced the accuracy of the results within the studies. Without accurately tracking the lifestyle changes throughout the interventions, the degree to which changes in cortisol were induced by the lifestyle changes is hard to evaluate.

The same issue was observed in studies measuring cortisol and mood. For instance, the five studies that examined the relationship between cortisol levels and anxiety measured anxiety using four different scales (i.e., Pregnancy Related Anxiety Questionnaire, Hospital Anxiety and Depression Scale, Overall Anxiety Severity and Impairment Scale, and State–Trait Anxiety Inventory). Of the three stress management interventions, one measured stress using the Stress Reactivity Scale and the Effort–Reward Imbalance Scale without specifying the time frame for which it was measured (Limm et al., 2011); one measured stress in the preceding month, at baseline, and, after intervention using the PSS (Cohen et al., 1983), the Parenting Stress Index (Abidin et al., 2006), and the Health Locus of Control Scale (Wallston et al., 1978; Tsiouli et al., 2014); and the third only measured stress in the preceding month, using the PSS only at baseline as a screening tool for participation (Sannes et al., 2015). Similar to the lifestyle measures, specific mood measures assess different aspects of mood. For example, the PSS measures general distress and the inability to cope (Cohen et al., 1983), and the Effort–Reward Imbalance Scale measures work stress, with a focus on efforts invested in job performance and rewards received in turn (Siegrist, 1996). As shown in this review, the associations between cortisol and various psychological outcomes may not be consistent (e.g., the pooled correlation was significant for cortisol and perceived stress and depression but not for cortisol and anxiety), indicating that the nature of the mood under investigation and the source of the negative mood experienced may be important.

4.4.2 Heterogeneity in approaches to measuring cortisol

Another observation of this review was the considerable variation in the measurement of long-term cortisol levels across the studies. Variability was evident at four levels: the type of cortisol measured, time frame captured, indices examined, and sampling protocol adopted.

Type of cortisol measured

Heterogeneity was observed across the studies in the types of cortisol measured. A broad overview of the studies showed that 80% of the observational studies measured cortisol in hair, and only one out of 27 trials (4%) measured cortisol in hair. None of the observational studies measured cortisol in blood, and six out of 27 trials measured cortisol in blood. Of the studies that examined the same lifestyle factors, different types of cortisol were measured. For example, out of six physical activity interventions, three measured cortisol in saliva, two in blood, and one in urine.

Cortisol measured in different samples reflects different fractions of cortisol. Blood samples can capture total cortisol circulating in the blood, which includes both free and bound cortisol. As explained in Chapter 2, only free cortisol is bioactive and related to effects on target cells (Lightman and Conway-Campbell, 2010). Free cortisol can also be calculated from the total cortisol measured in blood (Coolens et al., 1987). However, whether this calculation accurately reflects the amount of free cortisol circulating in blood is influenced by the level of CBG in blood (Lewis et al., 2005). It is generally assumed that cortisol measured in saliva (i.e., salivary cortisol) is free cortisol that has entered the salivary glands primarily by passive diffusion (Kirschbaum and Hellhammer, 1994). The concentration of cortisol in saliva is less than 10% of that in blood (Levine et al., 2007). Although salivary cortisol is often used as an alternative to free cortisol

measured in blood, previous evidence has suggested that the correlation between salivary and free serum cortisol varies between individuals in daily paired samples (Levine et al., 2007). Cortisol measured in both blood and saliva are affected by short-term fluctuations. Urine samples are often collected to measure urinary free cortisol (Turpeinen and Hämäläinen, 2013) over a longer period of time (e.g., 12–24 hours). Therefore, urinary cortisol captures the overall excretion of cortisol throughout the measured period. Although urinary free cortisol also reflects unbound or free cortisol, the correlation between urinary, salivary, and serum cortisol is still undetermined. Studies have suggested that salivary and urinary cortisol measured over the same 24-hour period are likely to provide different information about cortisol secretion; therefore, they should not be treated interchangeably (Yehuda et al., 2003). It is believed that cortisol measured in hair also reflects free cortisol in plasma (Kapoor et al., 2018). However, this general assumption is still under debate (Kalliokoski et al., 2019); therefore, the comparability of hair cortisol to saliva/urine/blood cortisol remains unclear. Results from this systematic review echo previous findings that cortisol measured in different samples is hardly comparable and should not be used interchangeably (Yehuda et al., 2003; Hellhammer et al., 2009; van Holland et al., 2012).

Indices examined

The second facet of the heterogeneity in measuring cortisol is the variety of cortisol indices measured even in studies measuring cortisol in the same type of human tissues. As discussed in Chapter 2, different indices of cortisol capture different characteristics of cortisol activity and should not be used interchangeably (Vedhara et al., 2006), which made the already complex results from different studies even more difficult to compare. For example, although two yoga interventions reviewed in this study both measured cortisol in saliva, one of them found significantly higher morning values and lower bedtime values of salivary cortisol

among participants post-intervention (Batista et al., 2015), and the other yoga intervention did not specify the cortisol index measured in saliva but did not find intervention effects on salivary cortisol concentrations (Chen et al., 2017). Another example is the mindfulness intervention. All six studies measured cortisol in saliva but two measured awakening salivary cortisol and reported significant intervention effects on cortisol (Marcus et al., 2003; Bergen-Cico et al., 2014), one measured diurnal cortisol and did not find significant intervention effects (Brown et al., 2016), one measured resting afternoon/evening cortisol and did not find significant intervention effects (Jacobs et al., 2013), and the two other did not specify which cortisol indices they measured and did not find significant intervention effects on cortisol (Daubenmier et al., 2011; Lynch et al., 2011).

With a limited number of studies measuring the effect of one lifestyle intervention on the same index of cortisol, it is therefore difficult to conclude if a certain type of lifestyle intervention affects cortisol levels or if the effect, if any, will only be revealed when measured using certain indices of cortisol. This issue has highlighted an important field requiring urgent attention: identifying which cortisol indices should be focused on when examining the influence of lifestyle and mood on cortisol levels. Further research can only demonstrate the impact of changes in lifestyle or mood on cortisol levels once this issue has been clarified.

Sampling procedure adopted

Studies included in this review not only measured different types and indices of cortisol but also adopted a variety of sampling procedures, which may further contribute to inconsistencies in the evidence. For example, although the six stress reduction mindfulness interventions all measured cortisol in saliva, the sample collection protocols differed considerably (e.g., the time samples were collected, number of samples collected per day, and number of consecutive days the samples were collected). As explained in Chapter 2, the accuracy and stabil-

ity of cortisol outcomes depend highly on the sampling protocol when sampling acute measures, such as blood and saliva. The sampling protocol includes the exact time the sample is taken, the frequency of sampling, the length of the sample collection period, and the adherence to the sampling protocol (Daubenmier et al., 2011; Saxbe, 2008; Turpeinen and Hämäläinen, 2013). Deviations from the sampling protocols can result in significant variations in the results of different studies (Clow et al., 2004). In addition, many studies in this review that used these acute measures failed to report the details of the sampling procedures adopted. Considering the significant differences in relation to the sampling procedure, lack of clarity about the rigour in timing and sufficiency of samples could have contributed to the inconsistency in the evidence. A solution to this problem may be to compare findings from studies that used consistent methods. However, as shown in this review, there are few studies that have adopted acute measures of cortisol and used the same sampling procedures to examine the same cortisol outcomes (O'Leary et al., 2016). With such widespread variability, it is not surprising that findings across studies were mixed or even contradictory. The variability across studies also makes it difficult to understand which cortisol indices are more sensitive to change. This observation further indicated that the ability of these acute measures to reflect the chronic relationship between lifestyle, mood, and cortisol or the enduring treatment effects on cortisol levels is limited.

The sampling procedure for hair cortisol is relatively straightforward and was consistent between studies: hair samples of a certain length were taken close to the scalp from the posterior vertex area of the head. However, the lengths of the hair samples being taken and analysed in the studies were different (1–9 cm). For example, the eight observational studies that examined the association between BMI/bodyweight and cortisol all measured cortisol in hair, but three studies took hair samples of 1–2 cm, four took samples of 3 cm, and one took samples of 6 cm. It is commonly assumed that a hair sample of 1 cm captures the systematic cortisol levels over the preceding month (Meyer and Novak, 2012).

Although this assumption has not yet been confirmed in human participants, and whether a specific length of hair sample reflects cortisol levels over a specific time frame is still inconclusive (Kapoor et al., 2018), the variations in hair sampling protocols undoubtedly add another variable when we synthesise the current literature. Therefore, although several studies included in this review examined the same lifestyle or mood factors and hair cortisol levels, the findings were still inconclusive.

4.4.3 The temporal issue of the association between cortisol and lifestyle or mood

Investigating the association concurrently

It was observed in this review that although a number of studies measured cortisol and lifestyle or mood concurrently, the time frames captured by the measures did not necessarily correspond. For example, three observational studies examined physical activity and HCC. However, two of the studies measured physical activity over the preceding week (Gerber et al., 2013; Sumra and Schillaci, 2015) but measured HCC over the preceding three months. One study asked participants whether or not they engaged in regular physical activity without specifying a time frame and measured HCC over the preceding three months (Garcia-Leon et al., 2018). Of the nine studies that investigated the association between perceived stress and cortisol, only three studies measured perceived stress and cortisol levels over the same time period (Byun, 2013; O'Brien et al., 2013; Geng et al., 2016) while the other six studies measured self-perceived stress in the preceding month and cortisol over the preceding 3–9 months (Chen et al., 2015; Faresjö et al., 2013; Gerber et al., 2013; Heinze et al., 2016; Kramer et al., 2009; Sumra and Schillaci, 2015). One of these studies reported a significant positive association between perceived stress and cortisol (Geng et al., 2016) and one reported a

negative association.(Gerber et al., 2013). In addition, several studies did not fully report the time frame across which lifestyle/mood factors were measured. For example, the study that examined the association between HCC over the preceding three months and sleep quality did not specify the time frame over which sleep quality was measured, although it was one of the lifestyle factors to be influenced by cortisol in their conceptual framework (Wang et al., 2016). Although these studies aimed to measure cortisol and lifestyle or mood at the same time, the actual time frames over which cortisol and lifestyle or mood were measured were not always the same. As a result, this raised the issue of whether the examined associations between cortisol and lifestyle or mood reflected the actual association over the time frame the cortisol levels were captured or the time frame over which lifestyle/mood was measured. This may further explain the mixed findings in this review. As a result of this limitation, future studies should clearly report the time frames over which these variables are measured and cautiously interpret the results.

However, when the studies that did measure lifestyle/mood and cortisol over the same period of time were examined, the findings were still confusing. For instance, of the aforementioned studies that measured perceived stress and cortisol, one study reported a positive association between evening salivary cortisol and perceived stress over the same month (Byun, 2013), but the other two studies reported no association between HCC and perceived stress over the preceding one month (Geng et al., 2016) or three months (O'Brien et al., 2013). Only one study included in this review measured lifestyle (alcohol consumption and smoking) and HCC over the same three-month period but did not find significant associations (Heinze et al., 2016). For these studies, a major issue is the accuracy of recalled mood over the measured time frame, as this can potentially be biased by their current mood, appraisal of past experiences, coping strategies, and personalities (Diener et al., 1984; Bradley and Mogg, 1994). The further in the past that the participants were asked to recall their mood, the more biased the results may be;

therefore, such results should be treated with caution.

Investigating the association between lifestyle or mood and subsequent cortisol levels

Most of the trials, by their nature, provided evidence of the association between lifestyle or mood and subsequent cortisol levels. For example, four out of six trials involving physical activity reported significant changes in cortisol levels after interventions (Batista et al., 2015; Mangine et al., 2016; Papacosta et al., 2013; Tanskanen et al., 2011). Goldberg et al. (2014) reported lower HCC among participants who stopped smoking for a month when compared with the month before the smoking cessation intervention. Two interventions found blunted cortisol responses to tasks in individuals undergoing alcohol abstinence interventions when compared with non-alcoholic individuals (Coiro et al., 2007; Sinha, 2011). Three out of six mindfulness interventions reported reduced cortisol levels following the intervention (Marcus et al., 2003; Bergen-Cico et al., 2014; Daubenmier et al., 2011). These findings suggest that trial evidence is helpful for understanding the effect of lifestyle or mood on cortisol (whether the effect exists and what the effect is) and the concurrent relationships between lifestyle or mood and cortisol, which are usually reported in observational studies.

Investigating the association between cortisol levels and subsequent mood

Two studies included in this review described findings on the association between cortisol and subsequent mood. Hoffman et al. (2016) reported that higher HCC in the first trimester of pregnancy was prospectively correlated with higher levels of perceived stress and depression at 40 weeks of gestation. (Cappuccio et al., 2010) reported that HCC in the first and third trimester of pregnancy positively

predicted postpartum depressive symptoms.

Similar associations between cortisol and major depressive disorder have been reported previously. For example, higher baseline CAR has been identified as a significant risk factor for developing major depressive disorder in adolescence (Adam et al., 2010; Ellenbogen et al., 2011; Vrshek-Schallhorn et al., 2013). There is also evidence suggesting that individual differences in morning salivary cortisol levels may be a risk factor for subsequent major depressive disorder in women (Harris et al., 2000). Based on these findings, it is suggested that our understanding of the relationship between cortisol and lifestyle or mood would benefit from prospective studies conducted over long time frames in which it is possible to explore not only the concurrent relationship between these factors but also conduct time-series analyses to help disentangle the cause and effect or, more likely, the circular relationship between cortisol, lifestyle, and mood.

The temporal inconsistency of when lifestyle or mood and cortisol were measured may contribute to the difficulties in (1) determining whether there is a correlation between lifestyle or mood and cortisol and (2) determining the direction of such a correlation, if any exists. For lifestyle, this time gap may potentially overlook the possible influence of changes in lifestyle on cortisol over the time frame of interest. This is particularly problematic for HCC because without evidence reflecting both lifestyle and HCC over the same time frame, it is difficult to uncover the effect of lifestyle on HCC at the time cortisol was deposited in the hair (Stalder et al., 2017). For mood, the conceptual framework of many studies included in this review described cortisol as a physiological indicator of exposure to stress (Chen et al., 2015; Faresjö et al., 2013; O'Brien et al., 2013). However, without enough studies measuring this association concurrently, it is unclear how representative cortisol is of mood over a certain time period. Indeed, this is further complicated by evidence that the relationship between cortisol and mood may be circular, with cortisol both affecting subsequent mood and mood affecting subsequent cortisol

(Grassi-Oliveira et al., 2012; Ullmann et al., 2017).

4.4.4 Issues from the interventions

A further observation from this review is that the effects of lifestyle or psychological interventions on long-term levels of cortisol remain unclear due to insufficient or mixed evidence and methodological limitations of the studies. The uncertainty has four facets. First, the overall picture was mixed, with 18 out of 27 trials showing significant effects on long-term cortisol levels.

Second, the persistence of the intervention effects on cortisol was unclear. For example, although both the yoga interventions reported significant reductions in cortisol immediately following the practice, only one observed the intervention effect over the longer term (Batista et al., 2015). In addition, 21 out of 27 trials did not include or report long-term follow-ups. Only five studies measured cortisol levels at follow-ups (Brown et al., 2016; Goldberg et al., 2014; Huang et al., 2012; Limm et al., 2011; Sannes et al., 2015), with different lengths of follow-ups ranging from two weeks (Huang et al., 2012) to one year (Limm et al., 2011). Only two out of these five studies reported significant long-term effects on cortisol one-month post-intervention (lower AUCg in the stress management intervention group, Sannes et al. (2015) and reductions in hair cortisol after a smoking cessation intervention, Goldberg et al. (2014)). Thus, while the evidence on the long-term persistence of intervention effects is unclear, these two studies suggest that it may be reasonable that effects persist for up to one month for some interventions.

Third, most trials described cortisol as a biomarker of mood and hypothesised that lifestyle or mood interventions would influence mood, which may be captured by changes in cortisol levels. However, this hypothesised mechanism was not always observed in these studies. Of the interventions that reported signifi-

cant effects on cortisol, several of them did not find significant treatment effects on the targeted mechanisms. For instance, one stress reduction intervention, in which lower awakening salivary cortisol levels were found, did not result in a significant improvement in self-reported perceived stress (Marcus et al., 2003). If an intervention did not result in any change in the hypothesised mechanism, even when the effect on cortisol was significant, it was impossible to distinguish whether the change in cortisol was attributable to the effectiveness of the intervention or to chance. A few studies did not measure or fully report changes in the targeted mechanisms. For example, one of the two yoga studies only measured stress using cortisol but not self-reported stress levels (Chen et al., 2017). Similarly, one intervention targeted at stress in veterans with PTSD did not report changes in self-reported or clinically assessed stress or PTSD (Bergen-Cico et al., 2014). A CBT-based stress management intervention did not measure self-reported psychological outcomes (Sannes et al., 2015). Previous literature has shown that cortisol levels (both short- and long-term levels) may not correlate well with self-reported mood measures, and these different measures should not be treated interchangeably (van Holland et al., 2012). Therefore, if a study did not measure or report the hypothesised mechanisms (i.e., the change in both self-reported mood and cortisol), regardless of whether the cortisol levels changed significantly after the intervention, it was impossible to conclude whether the intervention effectively changed participants' moods.

Several studies measured and reported mood mediators, with only the lifestyle interventions reporting significant changes in both mood and cortisol levels. For example, two physical activity interventions found both enhanced mood and decreased evening cortisol levels in yoga (Batista et al., 2015) and taper training (Papacosta et al., 2013). The mindfulness-based smoking cessation intervention reported both improvements in mindfulness and reductions in cortisol among participants, although changes in cortisol were not different from those induced by the CBT intervention (Goldberg et al., 2014). None of the other mindfulness-

based stress reduction interventions that measured and reported improvements in mood post-intervention found significant effects on cortisol (Brown et al., 2016; Jacobs et al., 2013; Lynch et al., 2011). However, due to different mood and cortisol measures being used in these studies (e.g., PSS and salivary cortisol, Batista et al. (2015) or Depression, Anxiety, and Stress Scale and hair cortisol, Goldberg et al. (2014)), it is difficult to conclude whether lifestyle interventions were more likely to influence long-term cortisol levels through enhancing mood than psychological interventions.

Finally, it is worth noting that another key limitation of this literature is the relative failure to hypothesise the direction of intervention effects on cortisol. Out of the 27 trials, less than half ($k = 12$) hypothesised the direction of the effect on cortisol, and only half of these ($k = 6$) reported findings that were in line with their hypotheses (Bergen-Cico et al., 2014; Daubenmier et al., 2011; Goldberg et al., 2014; Miller et al., 2015; Papacosta et al., 2013; Sannes et al., 2015). The other six studies that hypothesised an effect either did not find a significant intervention effect on cortisol (Brown et al., 2016; Chen et al., 2017; Diment et al., 2012; Holt-Lunstad et al., 2008; Leivadi et al., 1999) or found intervention effects in the opposite direction to what they hypothesised (Tam et al., 2014). Looking at the evidence, it is difficult to conclude whether a specific type of intervention is effective in changing chronic cortisol levels given the limited number of studies reporting significant effects in the expected direction. However, the six interventions that confirmed their hypotheses had clear descriptions of which cortisol indices they measured. It is likely that a lifestyle or mood intervention that aims to change cortisol levels may benefit from a specific hypothesis that outlines the indices for cortisol and direction of the possible effects.

4.4.5 Strengths and limitations

This is the first systematic review to synthesise the available literature on the associations between lifestyle and psychological factors and long-term cortisol levels in healthy adults. It is also the first systematic review to synthesise relevant evidence from both observational and interventional studies. By doing so, it has contributed to understanding the role of cortisol in lifestyle, stress, and mood when assessed over the long term. This review pointed to gaps between and within observational and interventional studies in cortisol research, which revealed several directions for future research. Another strength of this review is the robust methods that were used throughout the review process and the quality and scientific rigour with which the studies were assessed, taking into consideration both the study design and reporting. However, studies included in this review were highly heterogeneous, and thus a meta-analysis was only conducted for observational studies on BMI, physical activity, perceived stress, stressors, other indices of stress, anxiety, and depression. Findings of the meta-analyses led to the conclusion that BMI, physical activity, perceived stress, and depression are significant correlates of chronic levels of cortisol. However, due to the heterogeneity and mixed findings from relevant trials, this conclusion should be interpreted with caution. Another limitation should be noted that the overall quality of the available studies on the relationship between lifestyle and psychological factors and cortisol was only moderate for observational studies and poor for interventional studies. The lack of quality was salient in the external validity and bias in both the design and reporting of the studies, which further emphasises the need for robust systematic reviews in this research area. As was pointed out by a previous review, the lack of consensus within the literature about what cortisol outcomes are most meaningful to health is still a challenge in cortisol research (Saxbe, 2008). An urgent step for future research is to address the limitations and concerns discussed in this review and to rethink our knowledge of cortisol

as a biomarker of stress and behaviours. Duplicate title and abstract screening were not conducted due to the size of the search. However, a second reviewer was involved in full-text screening, which increased the certainty that relevant papers were included.

4.4.6 Recommendations for future research

Variability in approaches to measuring lifestyle, mood, and cortisol; the time intervals between when lifestyle/mood and cortisol were measured; and the issues surrounding the effects of the interventions on cortisol all contributed to uncertainty in the literature and the conclusions that could be drawn.

First, with the small amount of evidence available, the relationships between chronic cortisol levels and sleep, smoking, alcohol consumption, and caffeine consumption are unclear and require more research attention from both observational studies and trials. However, there is emerging evidence of a positive association between BMI/bodyweight, physical activity, perceived stress, and depression and cortisol, suggesting that interventional trials in these areas may be fruitful areas of enquiry. The findings from this review suggest that such trials would benefit from (1) using measures of cortisol that are able to capture levels of the hormone over the longer term, such as HCC, and (2) adding longer-term follow-ups post-intervention to investigate the persistence of the intervention effects on cortisol. To fully understand the potential intervention effects on cortisol, future interventions should clearly describe the hypotheses and aims of the study in relation to cortisol and obtain the cortisol levels at baseline, or over longer periods pre-intervention, as a reference point. The hypothesised mechanisms (e.g., the mediating mood factors) necessary for the effectiveness of the interventions should also be clearly described, measured, and reported. For the association between cortisol and mood, it is possible that the long-term association between these two factors is circular. There is evidence supporting the prospective effect

of mood on cortisol but insufficient evidence supporting the effect of cortisol on mood, which provides a useful direction for future research.

Chapter 5

Cohort Study: Aims and Methods

Figure 5.1 illustrates the conceptual model used to guide the hypotheses that were explored in the cohort study as part of this thesis. Evidence was reviewed in Chapter Two to Four and revealed that the role of stress and the hormone cortisol in influencing IVF treatment outcomes and the role of lifestyle in influencing this relationship (i.e., as a mediator or direct influencer) remain unclear. More specifically, the association between chronic cortisol levels in the month and three months before commencing IVF treatment and several IVF treatment outcomes (i.e., clinical pregnancy, biochemical pregnancy, or discontinued treatment) remains unclear. As shown in Figure 5.1, it has not yet been demonstrated in this thesis whether stress and lifestyle are correlated, and if so, how they are correlated in relation to IVF treatment. Therefore, the second purpose of the present study is to explore the interactions between lifestyle (i.e., BMI, physical activity, smoking, sleep, and alcohol and caffeine consumption), mood (i.e., anxiety, depression, and negative and positive mood), and cortisol levels on IVF treatment outcomes (i.e., clinical pregnancy, biochemical pregnancy, or discontinued treatment).

This chapter introduces the wider theoretical context underpinning the present

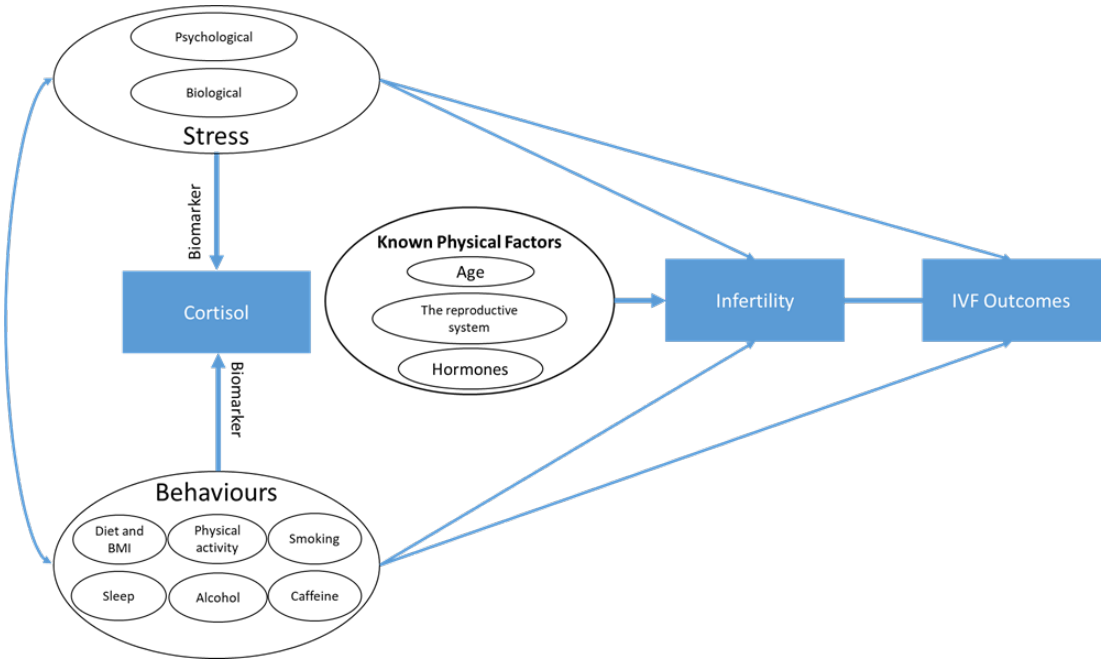


Figure 5.1 Thesis conceptual model

study, presents the aims and hypotheses, and describes the methods used. The first aim of the cohort study is to explore the association between HCC and IVF treatment outcomes, and the second is to explore the role of lifestyle and mood factors in this association.

5.1 Study aims and hypotheses

5.1.1 Aim 1

To examine the relationship between HCC before commencing IVF treatment and IVF outcomes, i.e., clinical pregnancy, biochemical pregnancy, or discontinuation of treatment.

Research question 1.1, 1.2 & 1.3

Do HCCs in the three months prior to commencing IVF differ between women who subsequently became clinically pregnant (primary outcome), biochemically pregnant, or discontinue treatment and those who do not? If these associations exist, do they exist both before and after adjusting for demographic and clinical factors?

It is hypothesised that HCCs in the three months prior to commencing IVF differ significantly between women who subsequently became clinically pregnant and those who do not. However, in view of the equivocal nature of the evidence regarding cortisol and IVF outcomes (Massey et al., 2016), we do not predict the direction of the difference between these groups. It is hypothesised that HCCs in the three months prior to commencing IVF differ between women who subsequently became biochemically pregnant or discontinue treatment and those who do not. It is hypothesised that these associations exist in both unadjusted and adjusted models accounting for demographic and clinical factors. The directions of the differences between the groups are not hypothesised.

Research question 1.4

Do HCCs in the one month prior to commencing IVF differ between women who

subsequently became clinically pregnant, biochemically pregnant, or discontinue treatment and those who do not? If these associations exist, do they exist both before and after adjusting for demographic and clinical factors?

First, it is hypothesised that HCCs in the three months and in the month prior to commencing IVF will be significantly correlated. Second, HCC in the month preceding IVF treatment will differ between women who subsequently become biochemically pregnant or discontinue treatment at any point during the treatment and those who do not. It is hypothesised that these associations exist in both the unadjusted and adjusted models accounting for demographic and clinical factors. The directions of the differences between the groups are not hypothesised.

5.1.2 Aim 2

To explore whether lifestyle (i.e., BMI, physical activity, sleep, smoking, and alcohol and caffeine consumption) and mood (i.e., anxiety, depression, and negative and positive mood) prior to commencing IVF treatment are related to treatment outcomes, (i.e., clinical pregnancy, biochemical pregnancy, or discontinuation of treatment).

Research question 2.1, 2.2 and 2.3

Does lifestyle (i.e., BMI, physical activity, sleep, smoking, and alcohol and caffeine consumption) in the month prior to commencing IVF differ significantly between women who subsequently become clinically or biochemically pregnant or discontinue treatment and those who do not?

It is hypothesised that lifestyle factors (i.e., BMI, physical activity, sleep, smoking, and alcohol and caffeine consumption) in the month prior to commencing IVF treatment will differ significantly between women who subsequently become

clinically or biochemically pregnant or discontinue treatment and those who do not. However, in view of the complexity of lifestyle factors (Homan et al., 2007), the directions of the differences between the groups are not hypothesised.

Research question 2.4, 2.5, and 2.6

Is mood (i.e., anxiety, depression, and negative and positive mood) in the month prior to commencing IVF treatment significantly associated with the likelihood of women subsequently becoming clinically or biochemically pregnant or discontinuing treatment?

It is hypothesised that greater levels of anxiety, depression, and negative mood in the month prior to commencing IVF treatment are significantly associated with a reduced likelihood of becoming clinically or biochemically pregnant and an increased likelihood of discontinuing treatment. It is hypothesised that greater levels of positive mood are significantly associated with an increased likelihood of becoming clinically or biochemically pregnant and a reduced likelihood of discontinuing treatment.

5.1.3 Aim 3

To explore whether HCC levels mediate the association between lifestyle (i.e., physical activity, sleep, smoking, and alcohol and caffeine consumption) and mood (i.e., anxiety, impact of events, depression, and negative and positive mood) prior to commencing IVF treatment and IVF outcomes (i.e., clinical pregnancy, biochemical pregnancy, or discontinuation of treatment).

Research question 3.1

Do HCC levels mediate the association between lifestyle factors (i.e., physical

activity, sleep, smoking, and alcohol and caffeine consumption) in the month prior to commencing IVF and IVF treatment outcomes (i.e., clinical pregnancy, biochemical pregnancy, or discontinuation of treatment)?

It is hypothesised that HCC levels in the month prior to commencing treatment mediate the association between lifestyle factors (i.e., BMI, physical activity, sleep, smoking, and alcohol and caffeine consumption) in the month prior to commencing IVF treatment and whether women subsequently became clinically or biochemically pregnant or discontinue treatment. However, in view of the complexity of lifestyle factors (Homan et al., 2007), we neither hypothesise which lifestyle factors are associated with HCC, clinical pregnancy, biochemical pregnancy, and discontinuation of treatment nor the direction of the associations that exist.

5.2 Methods

5.2.1 Participants

Power calculation

The study primarily seeks to document the associations between HCCs and the likelihood of achieving clinical pregnancy in female patients undergoing IVF treatment. Therefore, a sample size of at least n = 200 participants is needed to predict a number of feasible effect sizes with either 80% or 90% confidence in a two-tailed paired samples t-test with alpha set to 0.05.

Patient group eligibility criteria

(1) Women undergoing short or long protocol IVF (with or without ICSI)

(2) No fertility treatment in the previous 3 months

(3) Ability to speak and understand English

(4) Ability to give informed consent

(5) Able to provide a minimum of 3cm of hair taken from the vertex posterior of the head

(6) Aged between 18 years and 45 years

(7) No current or recent involvement in research involving medications prior to recruitment

5.2.2 Ethical approval and considerations

Ethical approval was granted by the NHS Health Research Authority, North East Newcastle & North Tyneside 1 Research Ethics Committee (ref: 17/NE/0278). The main ethical consideration in relation to this study was the possibility of inducing further discomfort in patients already taking part in a burdensome treatment program and the confidentiality and security issues relating to lifestyle data collection via mobile phone. These issues were circumvented by minimising the burden of the research on the participants. For example, it was decided that a limited number of lifestyle and mood measures would be included in the survey, and these data were only collected twice prior to commencing treatment. The collection of hair samples was non-invasive and completed during patients' visits to the clinic. All patient information was kept strictly confidential using encrypted passwords, and a secure server was used to store all the data. Each patient was given a unique study ID (i.e., a special code given to the patient to track their data), which contained no identifiable information. Mobile phone numbers were only used to send lifestyle survey links to participants and were confidentially

destroyed immediately afterwards.

5.2.3 Procedure

Recruitment (see Figure 5.2) of women undergoing IVF/ICSI treatment was conducted at CARE Fertility Nottingham (UK). The recruitment procedure was designed to fit with the typical practice at the clinic. The initial approach to eligible patients was made by providing the patient with a patient information sheet (PIS) with the patient's initial consultation letter sent from the clinic. Patients who were considering undergoing IVF first visited the clinic and had their initial consultation with one of the IVF doctors before consenting to the treatment. During their initial consultation, doctors informed the patients about the current research and gave them the PIS to read. Information about the study, in the form of a poster, was also displayed in the relevant clinical areas. The PIS was also available for patients to read freely in the patient lounge at the clinic. This ensured that eligible patients were exposed to the research and had enough time to consider participation.

Patients who decided to have IVF/ICSI treatment at CARE Fertility Nottingham visited the clinic to provide consent for the treatment and receive injection training. At that time, eligible patients were immediately approached by the researcher. Further detailed information about the research, including the aims, methods, benefits, and possible concerns, was given to the patients. The PIS was also given to patients at that time. Patients who decided to take part in the study were then given the consent forms to read and sign. Hair samples were taken by the researcher (RJ) immediately after receiving informed consent. The link to the lifestyle questionnaire was sent to participants via text message shortly thereafter. Patients who declined to participate were asked for reasons and were not contacted further. Patients who indicated that they wished to have more time to consider taking part were given the contact details of the research team.

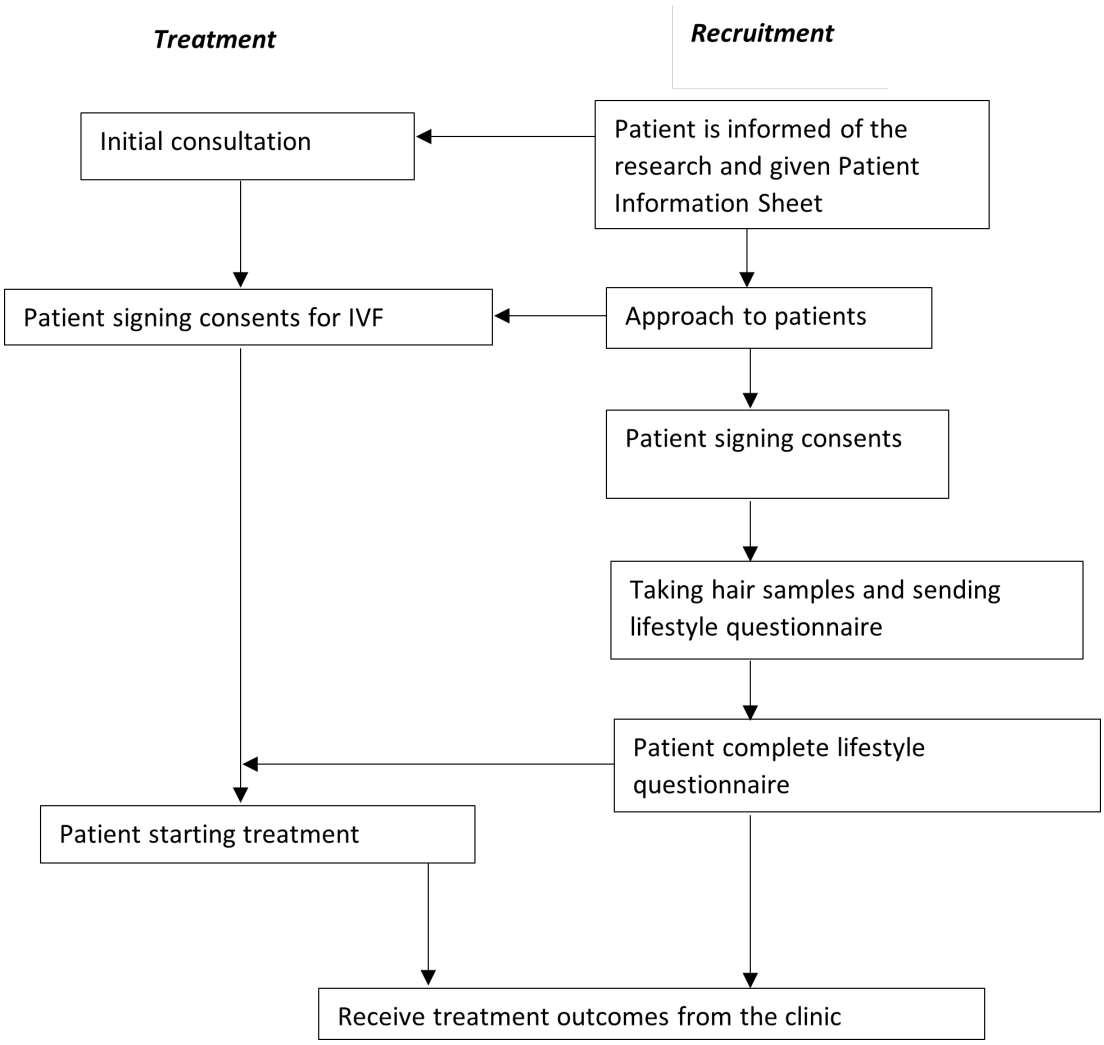


Figure 5.2 Flow of the treatment and recruitment procedure

Patients who indicated they wish to have further time to consider taking part or not were given the contact details of the research team. The recruitment took place from January 15th 2018 to August 20th 2019. A total of n = 371 patients were approached initially and subsequently n = 290 patients agreed to take part during this period. The analyses presented in this thesis relate to these patients.

Participants were carefully tracked using an anonymous study ID by stuff at the clinic for clinical information (age, BMI, previous miscarriage, months trying to conceive, previous IVF treatment, antral follicle count, and fertilisation rate). In this study the treatment phase refers to the period of time from the first day commencing treatment medications to the day they received clinical pregnancy test results. Clinical data were collected from January 25th 2019 to October 4th

2019 upon availability by the staff at the clinic, who were blind to the cortisol and lifestyle data.

5.2.4 Data management

A Microsoft Access database with a password was created to store patient information, including study ID, hair sample number, and important time points (e.g., date approached, date of completion of the questionnaire, and date the sample was taken). A special code was also given to patients who were approached but declined to take part. Reasons for not taking part were documented in the Access database. The database was stored on the university server and was updated on daily basis. Hair samples were labelled and kept at the Division of Primary Care, University of Nottingham.

5.2.5 Measures

Hair sampling

Human hair grows 1 cm per month on average; therefore, the 1 cm of hair closest to the scalp roughly represents the previous month's cortisol exposure (Sauvé et al., 2007). This study measured cortisol levels for approximately three months prior to commencing treatment from each participant using a 3 cm segment of hair nearest to the scalp, which is the same methodology used in the previous study (Massey et al., 2016). Hair samples were taken with scissors by cutting as close as possible to the scalp from the posterior vertex of the participant's head. This area has the least variation in cortisol (16%) when compared with other areas of the scalp (31%) (Sauvé et al., 2007). Before cutting the hair, participants were given the opportunity to look at the amount of hair to be cut. Hair samples were wrapped in aluminium foil and later transferred into 15 ml

Falcon Tubes for storage in a dry and dark cupboard.

In the current study, a 1 cm segment of hair was also analysed to determine the cortisol levels over approximately the previous one month for a subsample of participants (n = 190). Specifically, each hair sample from a non-random subsample (i.e., the last 100 patients recruited) was divided into 2 subsamples, including a 1 cm segment and a 3 cm segment of hair nearest to the scalp for cortisol measurement. The two samples were examined separately to provide data on cortisol levels from each participant over the preceding one month and three months accordingly to explore (a) whether cortisol levels in the month and the three months prior to commencing IVF treatment differed in their associations with clinical pregnancy and (b) the relationship between participants' lifestyles and cortisol levels in the month prior to commencing the treatment. HCC was determined using valid and reliable enzyme immunoassays (enzyme-linked immunosorbent assay [ELISA]). The hair cortisol assay procedure is described in detail in Section 5.2.7.

Questionnaires

Lifestyle and mood factors that (a) may directly influence fertility (e.g., frequency of unprotected sexual intercourse) and (b) may be indirectly related to fertility through their effects on cortisol (e.g., sleep quality, caffeine consumption, physical exercise, anxiety, depression, positive and negative mood, and infertility-related negative mood) were also measured via an online survey. To reduce the time to complete the survey for patients, perceive stress was not measured in this study while infertility-related negative mood (which captures infertility-related stress) was measured. Participants accessed the survey via a link sent to their phone or directly on a portable tablet available at the clinic. The survey was delivered to participants on the same day they signed the informed consent and the hair sample was collected. The online survey was designed using a university-

supported online survey tool: Jisc Online Survey. Participants were asked to complete the survey as soon as possible and no later than the start of their treatment cycle. The survey contained several sub-questionnaires (described below) to examine different lifestyle factors that may affect fertility and cortisol levels. All sum scores for the sub-questionnaires were computed in accordance with the authors' guidelines. Questionnaires were reverse coded were appropriate. All questionnaire measures used in this study were subject to reliability analysis. A reasonable level of internal consistency is generally considered to be values $\geq$0.7 (DeVellis, 2003; Kline, 2005).

Anxiety: State–Trait Anxiety Inventory, State Version Short Form

Anxiety was measured using the six-item short version of STAI-S Short Form, which is a shortened version of the original 40-item STAI developed by Spielberger et al. (1983) (Marteau and Bekker, 1992). The STAI-S produces scores similar to the original scale and has been widely used in general and clinical populations to assess the levels of state anxiety (Marteau and Bekker, 1992). It has demonstrated good reliability and validity among pregnant women (Marteau and Bekker, 1992; Meijer et al., 2014). Each item was scored on a four-point Likert scale ranging from not at all = 1 to very much = 4. Scores on the six-item STAI are in the range 20–80, with higher scores indicating greater levels of anxiety. This scale was used previously in patients undergoing IVF and was chosen in the current study due to its brevity (Massey et al., 2016). Instructions for the scale were modified to assess participants' anxiety in relation to their infertility. The six-item STAI had good reliability in this study (Cronbach's $\alpha = 0.84$).

Depression: Center for Epidemiologic Studies Depression Scale

Depression was measured using the 20-item CES-D (Radloff, 1977), which is a widely used depression measure and has been demonstrated to have high reliabil-

ity and validity in the infertile population and patients undergoing IVF (Domar et al., 1992; Gourounti et al., 2012). It assesses the frequency of depressive symptoms experienced by respondents over the preceding seven days. Each item is scored on a four-point Likert scale ranging from rarely or none of the time = 0 to most or all of the time = 3. Total scores on this measure are in the range 0–60, with higher scores indicating greater levels of depression, and a cut-off score $\geq$16 indicating clinical depression (Radloff, 1977; Roberts and Vernon, 1983). This cut-off criterion has been used previously in the infertile population and patients undergoing IVF (Pasch et al., 2012; Massey et al., 2016). The CED-S demonstrated acceptable reliability in the current study (Cronbach's $\alpha = 0.73$).

Positive and negative mood: The Positive and Negative Affect Schedule

The overall positive and negative mood were measured using the Positive and Negative Affect Schedule (PANAS, Watson et al. (1988)). The PANAS consists of two subscales measuring the extent to which respondents feel enthusiastic, active, or alert (positive affect), and the extent to which they feel sad, anxious, angry, or guilty (negative affect). Each item is scored on a five-point scale ranging from very slightly or not at all = 1 to extremely = 5. Total scores on the PANAS are in the range 10–50 for each subscale. It has been demonstrated to have high reliability and validity in previous studies among infertile women and patients undergoing IVF (Kiesswetter et al., 2020; Gonen and Bokek-Cohen, 2020). The PANAS demonstrated good reliability in this study (Cronbach's $\alpha = 0.76$).

Infertility-related negative mood: Impact of Event Scale

Infertility-related negative mood was measured using the IES (Horowitz et al., 1979). The IES was developed to measure the distress resulting from a specific life event or trauma (Horowitz et al., 1979). The IES was adapted to infertility

in the current study for relevance. It has previously been demonstrated to have high reliability and validity in people with fertility difficulties (De et al., 2017). Each item is scored on a four-point scale ranging from not at all = 0 to often = 5. Total scores on the IES are in the range 0–75, with higher scores indicating greater distress associated with the stressful event (i.e., infertility). Scores on the IES can further be divided into four categories according to the severity of the trauma impact: non-meaningful impact, mild trauma impact, moderate trauma impact, and severe trauma impact. The reliability of the IES in this study was high (Cronbach's $\alpha = 0.89$).

Sleep: The Pittsburgh Sleep Quality Index

Participants' sleep quality was measured using the PSQI (Buysse et al., 1989)). The PSQI is a self-report scale with seven components assessing seven aspects of sleep including subjective sleep quality, sleep latency, sleep duration, sleep disturbance, use of sleep medication, and daytime dysfunction. Each question is scored on a scale ranging from absence of sleep complaints = 0 to most severe sleep complaints = 3. The global score of the PSQI is calculated by adding scores from the seven components (range: 0–21), with a score >5 indicating poor sleep quality and a higher score indicating poorer sleep quality. The PSQI has demonstrated good reliability and validity in previous studies involving the infertile population (Eryilmaz et al., 2011). The reliability of the PSQI in this study was acceptable (Cronbach's $\alpha = 0.70$).

Physical activity: The Global Physical Activity Questionnaire

Physical activity was measured using the Global Physical Activity Questionnaire (GPAQ, Armstrong and Bull (2006)). The GPAQ is a self-report scale assessing physical activity, including context, such as work, travel to and from places, recreational activities, and time spent in inactivity. Physical activities are classified

into three intensity levels: moderate (e.g., cycling and swimming), vigorous (e.g., carrying or lifting heavy loads), and inactivity (e.g., sitting). The ratio of a person's working metabolic rate relative to their resting metabolic rate (metabolic equivalent [MET]) is used to express the intensity of physical activities: one MET for inactivity, four METs for moderate activities, and eight METs for vigorous activities. Respondents were asked to report whether they were involved in these activities, and if they were, the frequencies and typical lengths in a typical week. A respondent's overall energy expenditure (i.e., total METs) over a typical week was calculated based on these answers. The total physical activity scores indicate three different levels of total physical activity: high (total physical activity score $\geq$50 h/week), moderate (total physical activity score $\geq$10 h/week), and low (total physical activity score $<$10 h/week). The GPAQ has not previously been used among patients undergoing IVF. Because the GPAQ does not consist of multiple Likert-type scales and items that measure the same latent trait (Tavakol and Dennick, 2011), Cronbach's α was not suitable for measuring its reliability and hence is not reported here.

Caffeine consumption: The Caffeine Assessment Tool

Caffeine intake was measured with the Caffeine Assessment Tool (Boylan et al., 2008), which is a validated food frequency questionnaire self-report assessment of habitual consumption of caffeine (Boylan et al., 2008). Respondents were asked to report their habitual consumption of food and beverages containing caffeine (i.e., coffee, tea, soft drinks, energy drinks, chocolate, and chocolate coated or flavoured biscuits or dessert) at daily, weekly, and monthly levels, with clear indications of the brands of the food/beverages. The amount of caffeine contained in one portion of each type of food/beverage was acquired from the Caffeine Informer website: https://www.caffeineinformer.com. Each individual's habitual consumption of caffeine was then calculated based on their reported portions of food/beverages and the amount of caffeine per portion. This calculation was

done at three levels: daily consumption, weekly consumption, and monthly consumption. If a participant did not indicate the brand of the food or beverage, an estimated amount was assigned according to the average caffeine contents found in the same type of food/beverage on the Caffeine Informer website. To obtain an accurate record, participants were instructed not to report their weekly and monthly consumption by multiplying their daily caffeine consumption by the number of days in a week or month. Although the Caffeine Assessment Tool has not been used in IVF patients, it has been reported with good validity among pregnant women (Boylan et al., 2008; CARE Study Group and others, 2008). Because the Caffeine Assessment Tool is a food diary–like tool instead of a Likert scale, Cronbach's α is not reported here for the same reason described above.

Alcohol consumption: The Alcohol Use Disorders Identification Test

Alcohol consumption was measured using the Alcohol Use Disorders Identification Test (AUDIT, Saunders et al. (1993)). The AUDIT consists of 10 questions assessing three components: alcohol consumption, drinking behaviour and dependence, and alcohol-related problems. Participants were asked to report the frequency with which they experienced the described occasions on a Likert scale. Example items of the questions are as follows: "How often do you have a drink containing alcohol? (Answer from selection: Never/Monthly or less/2–4 times a month/2–3 times a week/4 or more times a week)" and "How often do you have six or more drinks on one occasion? (Answer from selection: Never/Less than monthly/Monthly/Weekly/Daily or almost daily)." We applied the time frame of this scale to the past month to be consistent with other scales in the survey. For question one to eight, scores ranged from 0 (the least frequent) to 4 (the most frequent). Questions 9 and 10 were scored 0 (the least frequent), 2, or 4 (the most frequent). The total scores of the AUDIT are in the range 0–40, with a score $\geq$8 indicating risky or hazardous levels of drinking behaviours. The AUDIT has not previously been used among patients undergoing IVF. In the current study, the

reliability of the AUDIT was low (Cronbach's $\alpha = 0.69$).

Smoking

Participants were asked whether or not they smoke (yes or no). Participants who answered "yes" were then asked how many cigarettes they smoked on average per day and the brand of cigarettes they usually smoked.

Hair characteristics

Several hair characteristics were considered as possible confounders of hair cortisol levels in the current study, as suggested in previous literature (Stalder and Kirschbaum, 2012; Wosu et al., 2013). Therefore, a brief questionnaire was used to collect participants' hair characteristics. Participants were asked how many times they washed their hair per week: answers of 1–3 times per week were coded as "low hair wash frequency," and answers of $\geq$4 times per week were coded as "high hair wash frequency" in the analysis. Participants were also asked whether they had ever dyed their hair (yes or no) and whether they had dyed their hair in the preceding three months (yes or no). A free text question was used to asked if participants used any hair products (e.g., hair spray or wax) in the preceding three months. Any answers to this question other than "no" were coded as "yes."

Clinical characteristics and treatment endpoints

A list of relevant clinical and demographic information was extracted from patients' clinical notes. These factors included age, BMI, length of time trying to conceive (months), previous fertility treatments (yes or no), previous live birth (yes, no, or unknown), previous miscarriage (yes or no), and antral follicle count. All the information was extracted upon availability by staff at the clinic who were blind to participants' cortisol levels and lifestyle data.

A list of treatment endpoints were also collected by staff at the clinic, including the number of oocytes collected, the number of oocytes fertilised, and the fertilisation rate (%). Fertilisation rate was calculated as the number of oocytes fertilised as a percentage of the number of oocytes collected.

5.2.6 Clinical outcomes

Clinical pregnancy

Clinical pregnancy, defined as a pregnancy diagnosed by ultrasonographic visualisation of foetal heart activity at six weeks post embryo transfer, was the primary outcome of this study. Analyses focusing on clinical pregnancy excluded women who discontinued treatment at some stage prior to the ultrasonographic pregnancy scan. Clinical pregnancy outcomes were extracted from patients' medical notes within two months of patients commencing treatment.

Biochemical pregnancy

Biochemical pregnancy was defined as positive result on a urine or blood hCG pregnancy test. Analyses focusing on biochemical pregnancy excluded women who had withdrawn from or failed treatment at some stage prior to the biochemical pregnancy test. Biochemical pregnancy outcomes were extracted from patients' medical notes within two months of patients commencing treatment.

Discontinued treatment

Discontinued treatment was defined as discontinuation of the treatment cycle due to complications occurring at any stage of treatment (e.g., OHSS, no eggs collected, or failed fertilisation) and loss of pregnancy after hCG pregnancy test.

Women who did not complete treatment by the time of data collection due to planned frozen cycles or other conditions (e.g., being ill) or who withdrew from treatment were excluded from the discontinued treatment outcome.

5.2.7 Hair cortisol assay

After collection, hair samples were stored at Division of Primary Care and were assayed in batches throughout the study phase in the Reproduction Technology Teaching Laboratory, University of Nottingham, Queen's Medical Centre. All samples were assayed for cortisol between 2018, August 7th and 2019, October 15th.

Hair cortisol extraction and assay were conducted following the established protocol provided by Meyer et al. (2014), which included washing, weighing, grinding, extracting, and assaying. To wash the hair samples, 5 mL of high-performance liquid chromatography-grade isopropanol was added to each Falcon tube containing a hair sample followed by repeated inversion for 5 min using a rotator. This procedure was repeated before each sample was removed from the tube, placed on a paper towel with the sample number and the root end carefully labelled, and left to dry completely for two days. From the root end of the dried samples, 10–20 mg of the first 3 cm of the dried samples were cut into pre-weighted 2 mL Eppendorf® Safe-Lock tubes with a 5 mm stainless steel ball inside. Each tube was weighed again to obtain the weight of the hair samples. The tubes were carefully loaded on a TissueLyser and ground for 15 min at 30 MHz. High-performance liquid chromatography-grade methanol (1 mL) was then added, and the samples were shaken on a Rotating Mixer overnight at 35 rpm. The next day, the tubes were centrifuged for 3 min at 10,000 rpm in a micro centrifuge, and the supernatant was transferred to clean Eppendorf® tubes and centrifuged for a second time. In the next step, 700 μl of the methanol extract was pipetted into 75*12 mm glass tubes, which were then used to dry the solvent under air

using the sample concentrator. The dried extract was then re-dissolved in 250 μl of assay buffer and vortexed well before assaying.

The hair cortisol assay followed the assay protocol on the instructions of the ELISA kits. The HCC of the first 141 samples were determined using a high sensitivity commercially available enzyme immunoassay (ELISA, IBL International, Hamburg, Germany). The ELISA kit had reduced sensitivity for hair cortisol and failed our quality control parameter for assays for the following n = 22 samples. Therefore, HCCs from patient hair samples numbered 161–290 were determined using a different ELISA kit (DRG International, Inc., Springfield, USA). The ELISA kits were designed for the determination of free cortisol in human saliva. For each ELISA kit, the intra- and inter-assay coefficients of variance deemed acceptable for assay were <10%. Hair cortisol concentrations were expressed in pg/mg. The instructed assay protocols for the two kits were different (see Appendix), which resulted in different ranges of expected results for HCC. The range of readings from IBL kits (median = 76.2, range = 28.4–435.4 pg/mg) were significantly higher than from DRG kits (median = 14.6, range = 5.4–366.1 pg/mg). This difference in ELISA kits was considered a confounder, and it was controlled for in the statistical analyses and reported.

5.2.8 Approaches to data analysis

All data analysis was performed using SPSS (version 26.0 for Windows, IBM Corp, USA) and STATA (version 16.1 for Windows, StataCorp LLC, USA).

Data cleaning

One participant was identified as not being eligible for the study and was removed from analysis due to being older than 45 years. The number of cases with missing value(s) were counted. For demographic and clinical characteristics (i.e., age,

ethnicity, previous birth history, previous miscarriage, length of time trying to conceive, previous treatment at CARE Fertility Nottingham, BMI, and antral follicle count), missing values were found in 35.2% of cases. Of the cases with missing values, 69 had one missing value, 21 had two missing values, two had three missing values, and two had eight missing values. For treatment outcomes, two cases had one missing value, and three cases had four missing values. For clinical outcomes, 34 cases (12.7%) were missing all three clinical outcomes and 39 (14.6%) were missing two outcomes either due to discontinuation of treatment (n = 71) or no data provided by the clinic (n = 2). For HCC values, a total number of nine cases (3.4%) were recorded as missing due to readings being too high to read (n = 1), lab error (n = 1), or being extreme values (n = 7); therefore, they were excluded from the analyses. The number of cases with missing data for demographic and clinical characteristics were as follows: age, n = 2; BMI, n = 2; miscarriage, n = 48; ethnicity, n = 13; ovarian reserve, n = 3; number of eggs collected, n = 3; number of sperm inseminated, n = 3; number of eggs fertilised, n = 5; and fertilisation rate, n = 5. The number clinical outcomes were as follows: clinical pregnancy, n = 73; biochemical pregnancy, n = 73; and discontinued treatment, n = 34. The missing data were largely due to events during treatment. To account for the influence of missing data on regression models, sensitivity analyses were conducted by using multiple imputation (MI).

5.2.9 Hair cortisol data

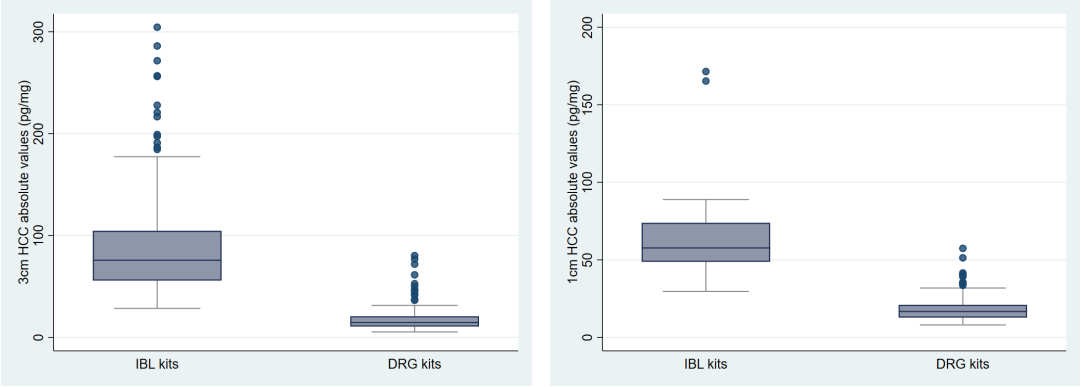
True HCC values from both ELISA kits were transformed into Z-scores to obtain standardised values. Visual examination of the histograms showed that the Z-scores were not normally distributed. Therefore, the Z-scores were subjected to log transformation for normal distribution. Unless specified, all comparisons involving HCCs were conducted using the transformed standardised values.

Table 5.1 presents 1 cm HCCs and 3 cm HCCs as measured by the ELISA kits.

Mann–Whitney U tests were conducted to examine the difference in the absolute values of the 3 cm and 1 cm HCCs by the different cortisol assays (IBL *vs* DRG). The results showed that the 3 cm HCCs (absolute values) from the IBL assays were significantly higher (mean rank = 490.9) than those from the DRG assays (mean rank = 65.4, z = -13.4, p< 0.001). Similarly, the absolute values of the 1 cm HCCs from the IBL assays (mean rank = 143.9) were significantly higher than those from the DRG assays (mean rank = 63.0, z = -9.4, p< 0.001). Due to significant differences in the absolute values of HCCs from the different assays, the 3 cm and 1 cm HCC absolute values were transformed into Z-scores to allow analyses using HCCs from both types of assay. Log transformations were further conducted on the Z-scores transformed HCC values due to the Z-scores not being normally distributed. Pearson's correlation coefficient revealed that the three-month HCCs and the one-month HCCs of patients undergoing IVF were significantly correlated (r = 0.67, p< 0.001). Paired samples t-tests showed that the one-month HCCs were significantly lower than the three-month HCCs (log transformed Z-scores, t = -11.53, p< 0.001, CI = -0.24—0.17). Figure 5.3 shows the distributions of the absolute values of the 3 cm HCCs by assay. Figure 5.4 shows the distributions of HCC values by assay after Z-score transformation and log transformation.

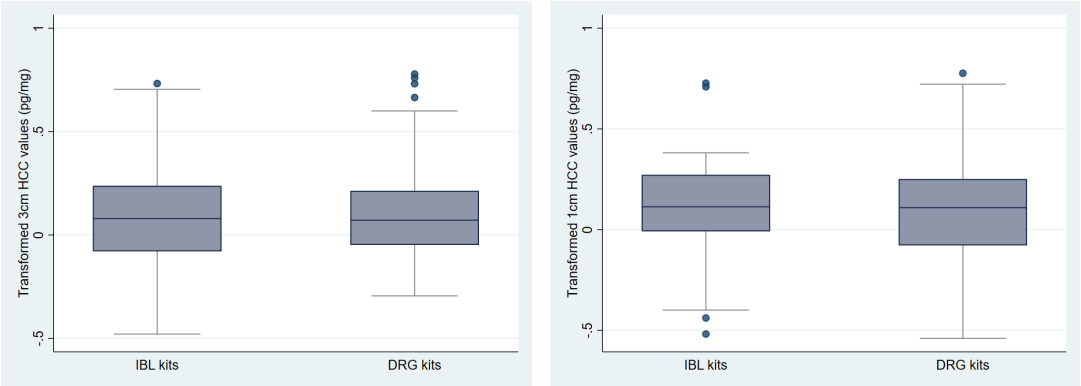
Table 5.1 HCC absolute values before transformation

	3cm HCC	3cm HCC	1cm HCC	1cm HCC
	IBL assay kit	DGR assay kit	IBL assay kit	DGR assay kit
Median	75.73	14.50	57.66	16.66
25% quartiles	55.73	10.63	48.70	12.74
75% quartiles	104.59	20.72	73.88	20.99
Minimum	28.43	5.35	29.62	8.00
Maximum	304.48	80.23	171.43	57.38



(a) 3cm HCC absolute values (b) 1cm HCC absolute values

Figure 5.3 HCC absolute values (pg/mg) by ELISA. (a) shows the 3cm HCC absolute values by ELISA and (b) shows the 1cm HCC absolute values by ELISA. Boxes in this figure represent 25% quartiles to 75% quartiles. The solid line inside the boxes represent medians. The whisks represent the lower and upper adjacent values. The dots represent outside values after removal of extreme values from the dataset.



(a) Transformed 3cm HCC values (b) Transformed 1cm HCC values

Figure 5.4 Log-transformed and Z-score transformed HCC values (pg/mg) by ELISA. (a) shows the transformed 3cm HCC values (pg/mg) by ELISA and (b) shows the transformed 1cm HCC values (pg/mg) by ELISA. Boxes in this figure represent 25% quartiles to 75% quartiles. The solid line inside the boxes represent medians. The whisks represent the lower and upper adjacent values. The dots represent outside values after removal of extreme values from the dataset.

Patient and treatment characteristics

Descriptive statistics were computed prior to analysis for all major patient and treatment variables that were collected from the clinical database, including female age (years), BMI, ethnicity, previous birth history (yes or no), length of time trying to conceive (months), previous miscarriage (yes or no), number of treatment cycles, antral follicle count, and fertilisation rate (%).

Between-group differences

Independent sample t-tests were conducted to examine the differences in continuous demographic and clinical factors (i.e., age, BMI, antral follicle count, months trying to conceive, number of oocytes matured, and number of embryos fertilised) between women who subsequently achieved clinical or biochemical pregnancy or discontinued treatment and those who did not. Chi-squared tests were conducted to examine the differences in dichotomous demographic factors (i.e., ethnicity, previous birth, or previous miscarriage) between women who subsequently achieved clinical or biochemical pregnancy or discontinued treatment and those who did not.

Correlations

Histogram examinations showed that the HCC absolute values were not normally distributed. Therefore, differences in the 3 cm HCCs and 1 cm HCCs between the IBL and DGR assays were examined using the Mann–Whitney U test. To combine the 3 cm HCCs from the IBL and DGR assays, absolute HCC values were transformed into Z-scores before further log transformation to achieve a normal distribution. This approach was also applied to the 1 cm HCC absolute values. Differences between the combined 3 cm HCCs and combined 1 cm HCCs were examined using paired sample t-tests. Correlations between lifestyle (i.e.,

BMI, physical activity, sleep, smoking, and alcohol and caffeine consumption) and mood variables (i.e., anxiety, depression, and negative and positive mood) were examined using Pearson's r.

Multicollinearity

Multicollinearity can occur when predictors within a regression model are too highly correlated, which reduces the validity of the model and increases the chances of making a type II error, i.e., incorrectly concluding that a predictor is non-significant. Treatment outcomes (i.e., the number of oocytes collected, the number of sperm inseminated, the number of oocytes fertilised, and the fertilisation rate) were expected to be correlated, as they reflect overlapping aspects of the treatment. Therefore, only the fertilisation rate was included in the regression analyses as a collective variable for all treatment outcomes. The scales used for this study were also expected to correlate, as they measured overlapping constructs (i.e., stress, anxiety, and depression). Therefore, the extent of multicollinearity was examined prior to analysis by assessing the variance inflation factor and tolerance statistics. A variance inflation factor threshold statistic of >10.0 was used to indicate acceptability, corresponding to a tolerance value of 0.1.

Linear regression

The independent associations between hair characteristics (i.e., hair washing frequency, hair dye in the preceding three months, hair treatment in the preceding three months, and the use of hair products) and 1 cm and 3 cm HCCs were examined using univariate linear regression. Hair wash frequency (i.e., low: 1–3 times per week and high: 4–7 times per week), hair dye in the preceding three months (yes or no), hair treatment in the preceding three months (yes or no),

and the use of hair products (yes or no) were treated as binary variables in the regression models.

Logistic regression

In this study, logistic regressions were conducted to explore the associations between HCCs and IVF outcomes, including clinical pregnancy (1 = pregnant and 0 = not pregnant), biochemical pregnancy (1 = pregnant and 0 = not pregnant), and discontinued treatment (1 = discontinued treatment and 0 = continued treatment). The statistical assumptions pertaining to the logistic regression were explored prior to analysis. Logistic regression was used because it is suitable for dichotomous, categorical outcomes and was therefore used to examine whether self-reported mood and HCCs were capable of predicting whether patients obtained a clinical pregnancy or not. The logistic regression models calculated the odds of the various events occurring (i.e., clinical pregnancy, biochemical pregnancy, or discontinued treatment), with the associated 95% CI, as a function of the predictor variables. An OR > 1 indicates that the event (i.e., clinical pregnancy, biochemical pregnancy, or discontinued treatment) is at increased likelihood to occur.

The independent associations between demographic and clinical factors (i.e., age, BMI, ethnicity, previous miscarriage, previous IVF treatment, months trying to conceive, antral follicle count, and fertilisation rate) and treatment outcomes (i.e., clinical pregnancy, biochemical pregnancy, or discontinued treatment) were examined using univariate logistic regression. Age, BMI, months trying to conceive, antral follicle count, and fertilisation rate were treated as continuous variables in the regression models predicting clinical pregnancy (yes or no), biochemical pregnancy (yes or no), and discontinued treatment (yes or no). Ethnicity (Black, Asian and minority ethnic [BAME] or white British [non-BAME]), previous miscarriage (yes or no), and previous IVF treatment (yes or no) were treated as

binary variables in these models.

The univariate regression to predict each of the three clinical outcomes used log transformed cortisol values for the individual assays. When the data from the IBL and DGR assays were combined, the log transformed Z-scores were used to aid in combining the data.

Multivariate logistic regression analyses explored the relationships between hair cortisol and treatment outcomes (i.e., clinical pregnancy, biochemical pregnancy, and discontinued treatment) after controlling for demographic and clinical covariates. As age is the most influential factor in determining IVF outcomes, it was included as a covariate in all multivariate analyses. Previous literature has shown that psychological stress and miscarriage are related (Qu et al., 2017). Women with a history of miscarriage are at risk of experiencing anxiety during subsequent pregnancies (Geller et al., 2004; Woods-Giscombé et al., 2010), and miscarriage history has been shown to influence IVF outcomes (Bates Jr and Ginsburg, 2002; Kupka et al., 2004). Therefore, miscarriage history was considered a clinical characteristic covariate of all three clinical outcomes. Fertilisation rate (number of oocytes fertilised as a percentage of the number of oocytes collected, %) was identified as a covariate in a previous study exploring the association between hair cortisol and clinical pregnancy in IVF (Massey et al., 2016), and therefore, it was included in the multivariate analyses as a covariate of clinical pregnancy and biochemical pregnancy. However, because no oocytes collected and failed fertilisation could both be reasons for terminating treatment, the fertilisation rate was not included in the model predicting discontinued treatment.

Sensitivity analysis

To account for the influence of missing data in the relevant variables in the multivariate regression models, sensitivity analysis using MI datasets was conducted.

The percentages of missing values for each variable included in the model were in the range 1–53% out of n = 192 (see Results: Sensitivity analysis with MI datasets). Therefore, 20 datasets were generated. Multivariate regression models predicting clinical pregnancy, biochemical pregnancy, and discontinued treatment outcomes using MI datasets and estimates were combined using Rubin's rules.

Chapter 6

Cohort Study: Results

6.1 Patient flow

A total of 290 patients (out of 370 that were approached) consented to take part in the study after being approached by the researcher, yielding a recruitment rate of 78.4%. Of these, one participant withdrew from the study, and 26 participants were lost to follow-up. A total of 258 patients (89%) started ovarian stimulation as planned and 225 patients (78%) subsequently underwent egg collection. A total of 24 patients froze eggs or embryos as planned, and 33 patients cancelled treatment due to a variety of conditions, including infection, risk of OHSS, no oocytes collected, immature oocytes, unsuitable embryo for transfer, failed cleavage, etc. Therefore, 198 patients (68%) proceeded to embryo transfer, and their data were available to analyse the primary outcome (i.e., clinical pregnancy). Six patients cancelled treatment after embryo transfer due to failed fertilisation or other failures in treatment. A total of 93 patients (32%) conceived, as evidenced by a positive urine pregnancy test (biochemical pregnancy). Subsequent transvaginal scan confirmed a viable intrauterine pregnancy in 82 patients (28% of the whole cohort). Of the 192 patients who had pregnancy tests, HCC data were available for 190 patients, and these patients were included in the data analysis. Figure 6.1 illustrates the flow diagram of patient recruitment and the number of patients with full data available in the present study.

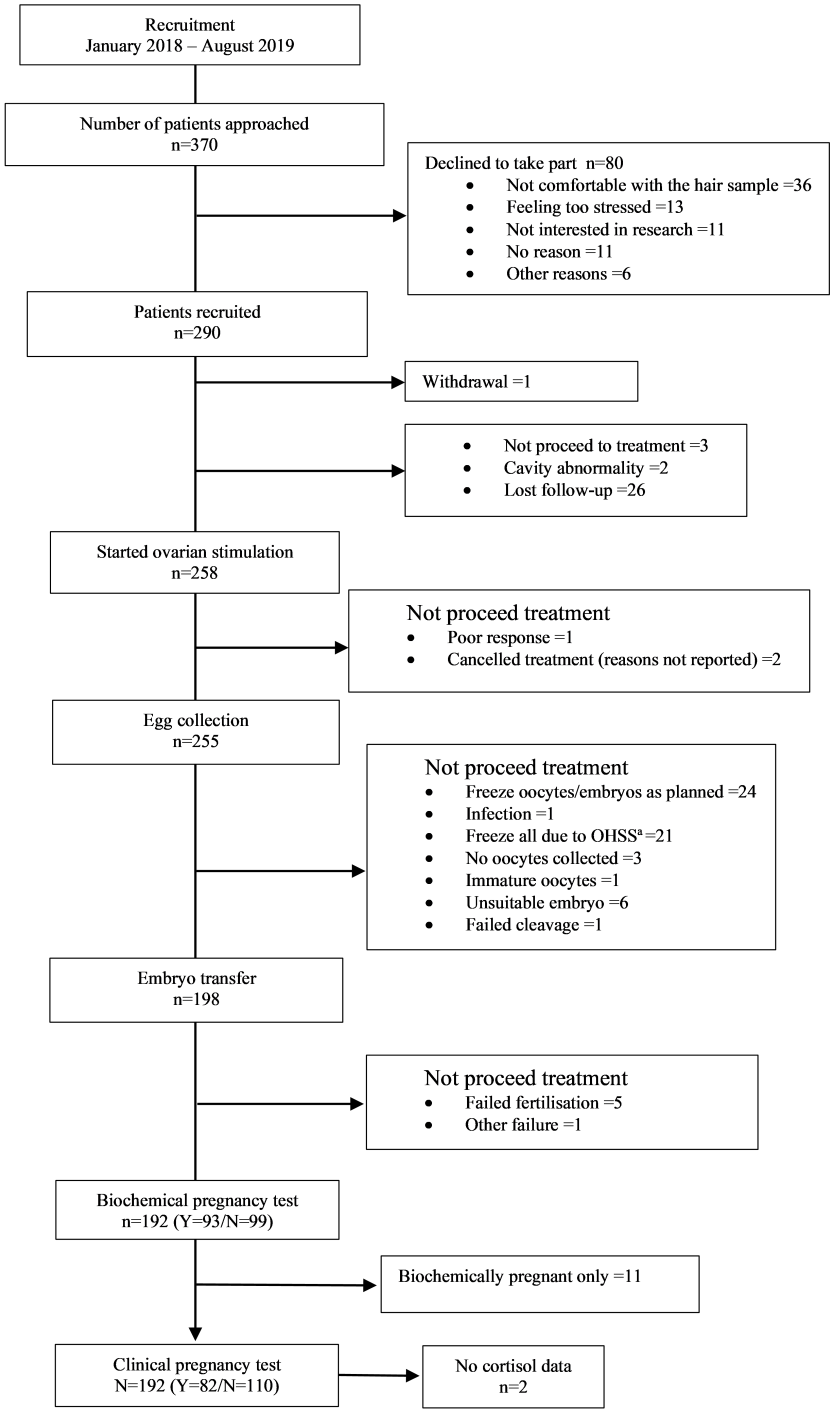


Figure 6.1 Flow diagram of patient recruitment and data collection

6.2 Cohort characteristics

The cohort was predominately white British (84.3%) with an average age of 34.4 years (SD = 4.6 years). Approximately one-third of the patients (n = 81) had a previous live birth, and 39 patients reported a previous miscarriage. The average duration for which patients were trying to conceive was 38.7 months (SD = 25.7 months). Table 6.1 summarises the main demographic and clinical characteristics of the participants. Lifestyle and mood of the participants are presented in later sections. However, it should be noted that a high proportion of participants (n = 186, 69.7%) reported high levels of depression that reached the cut-off for clinical depression (scores $\geq$16) (Radloff, 1977; Roberts and Vernon, 1983). Independent sample t-tests showed that women who subsequently became pregnant and those who did not only differed significantly in age ($p = 0.036$), number of oocytes matured ($p = 0.045$), and number of embryos fertilised ($p = 0.041$). Independent t-tests were also conducted to examine the differences between women who achieved biochemical pregnancy and those who did not, and the results were similar in that they differed significantly by age ($p = 0.0076$), number of oocytes collected ($p = 0.0070$), number of oocytes matured ($p = 0.0004$), and number of embryos fertilised ($p = 0.0021$). Women who did not achieve biochemical pregnancy were also more likely to have had a prior miscarriage ($\chi^2 = 4.45$, $p = 0.035$).

Table 6.1 Cohort characteristics

	Whole sample (n=234)	**Subsequently pregnant (n=82)**	**Subsequently not pregnant (n=110)**
Age (mean, *SD*)	34.2 (4.5)	33.5 (4.1)	34.8 (4.6)
Ethnicity (n, %)			
White British	188 (85.1%)	71 (89.9%)	87 (85.3%)
Asian	17 (7.7%)	4 (5.1%)	8 (7.8%)
Black African	1 (.5%)	0	0
Other white background	10 (5.0%)	2 (2.5%)	7 (6.9%)
Other mixed background	4 (1.8%)	2 (2.5%)	0
Previous birth (n, %)			
Yes	81 (31.2%)	24 (30%)	35 (31.8%)
No	156 (60.0%)	48 (60%)	72 (65.5%)
Not reported	23 (8.9%)	8 (10%)	3 (2.7)
Duration trying to conceive (n, %)			
Less than 1 year	17 (8.7%)	4 (4.9%)	11 (10%)
1–2 years	62 (31.6%)	25 (30.5%)	27 (24.6%)
2–3 years	50 (25.5%)	18 (22.0%)	23 (20.9%)
3–4 years	30 (15.3%)	13 (15.9%)	15 (13.6%)
More than 4 years	37 (18.9%)	10 (12.2%)	23 (20.9%)

continued on the next page

Miscarriage history (n, %)			
Yes	39 (19.3%)	7 (10%)	20 (20.8%)
No	163 (80.7%)	63 (90%)	76 (79.2%)
BMI[b] (mean, SD)	25.5	25.4	25.6
Antral follicle count (mean, SD)	23.8	25.6	21.8
Number of oocytes collected (mean, SD)	11.5	11.0	10.1
Number of mature oocytes* (mean, SD)	10.0	9.8	8.6
Number of embryos fertilised* (mean, SD)	6.8	7.0	6.0
Fertilisation rate (mean, SD)	59.5	64.8	60.6
Hair wash frequency			
1–3 hair washes per week (n, %)	131 (69.7%)	36 (70.6%)	61 (74.4%)
4 or more hair washes per week (n, %)	57 (30.3%)	15 (29.4%)	21 (25.6%)
Hair dye (n, %)	184 (81.8%)	53 (77.9%)	75 (81.5%)
Hair dye in three months (n, %)	156 (69.6%)	42 (62.7%)	67 (72.8%)
Use of hair product (n, %)	172 (83.9%)	54 (85.7%)	70 (83.3%)

[a]Data in this table are presented by n (%) or mean (*SD*).

[b]BMI: body mass index

*Significant difference between women who were subsequently pregnant and those who were not.

6.3 Part 1: cortisol levels and *in vitro* fertilisation treatment outcomes

6.3.1 Hair cortisol concentration among patients undergoing *in vitro* fertilisation treatment

Univariate linear regressions were conducted to explore the associations between hair characteristics (i.e., hair wash frequency, hair dye in the preceding three months, hair treatment in the preceding three months, and use of hair products) and HCC over the preceding three months and one month. This was done to identify potential confounders of HCCs. The results (see Table 6.2) showed that frequency of hair washing (low: 1–3 times per week *vs* high: 4–7 times per week, B = -0.07, 95% CI = -0.16–0.02), hair dye in the preceding three months (yes or no, B = 0.01, 95% CI = -0.07–0.10), hair treatment in the preceding three months (yes or no, B = -0.04, 95% CI = -0.32–0.23), and the use of hair products (yes or no, B = 0.003, 95% CI = -0.11–0.11) were not associated with 3 cm HCCs. Frequency of hair washing (low: 1–3 times per week vs high: 4–7 times per week, B= -0.06, 95% CI = -0.16–0.05), hair dye in the preceding three months (yes or no, B = -0.08, 95% CI = -0.18–0.01), hair treatment in the preceding three months (yes or no, B = -0.08, 95% CI = -0.32–0.15), and the use of hair products (yes or no, B = 0.002, 95% CI = -0.13–0.13) were also not associated with 1 cm HCCs. Together these analyses show that none of the measured hair characteristics were associated with either 3 cm or 1 cm HCCs. Therefore, they were not considered in later analyses.

Table 6.2 Univariate logistic regression models showing hair characteristic predictors of hair cortisol levels

	3cm HCC from combined assay results	1cm HCC from combined assay results
	B (95%CI), *p*	B (95%CI), *p*
Hair wash (low/high)	-.07 (-.16, .02), .15	-.06 (-.16, .05), .28
Adjusted R^2	.001	.001
n	185	117
Hair dye in the previous 3 months (yes/no)	.01 (-.07, .10), .73	-.08 (-.18, .01), .07
Adjusted R^2	-.004	.02
n	219	139
Hair treatment in the previous 3 months (yes/no)	-.04 (-.32, .23), .75	-.08 (-.32, .15), .49
Adjusted R^2	-.004	-.003
n	220	140
Use of hair product (yes/no)	.003 (-.11, .11), .96	.002 (-.13, .13), .97
Adjusted R^2	-.005	-.01
n	200	126

6.3.2 Predicting three treatment outcomes: examining demographic and clinical covariates

Univariate logistic regressions were conducted to examine the association between demographic and clinical variables and clinical pregnancy (primary outcome): age, BMI, ethnicity (Black Asian and Minority Ethnic, BAME/non-Black Asian and Minority Ethnic, non-BAME), previous miscarriage (yes or no), previous live birth (yes or no), previous IVF treatment (yes or no), months trying to conceive, antral follicle count, and fertilisation rate. The results showed that decreasing age was significantly associated with increasing likelihood of becoming clinically pregnant (OR = 0.93, 95% CI = 0.87–1.00, p = 0.04). However, BMI (OR = 1.01, 95% CI = 0.95–1.08, p = 0.73), BAME background (OR = 0.69, 95% CI = 0.29–1.65, p = 0.40), previous miscarriage (OR = 0.42, 95% CI = 0.17–1.06, p = 0.07), previous live birth (OR = 1.03, 95% CI = 0.55–1.94, p = 0.93), previous IVF treatment (OR = 0.55, 95% CI = 0.23–1.34, p = 0.19), months trying to

conceive ($OR = 0.99$, 95% CI = 0.98–1.00, $p = 0.22$), antral follicle count ($OR = 1.01$, 95% CI = 1.00–1.03, $p = 0.17$), and fertilisation rate ($OR = 1.01$, 95% CI = 1.00–1.02, $p = 0.17$) were not independently associated with clinical pregnancy (Table 6.3).

Table 6.3 Univariate logistic regression models showing individual predictors of treatment outcomes

	Clinical pregnancy	**Biochemical pregnancy**	**Discontinued treatment**
	odds ratio (95%CI), *p*	*odds ratio* (95%CI), *p*	*odds ratio* (95%CI), *p*
Age	**-.07 (.87, .996), .04***	**.91 (.86, .98), .009****	.99 (.92, 1.06), .75
Pseudo R^2	.02	.03	.0005
n	192	192	232
BMI	1.01 (.95, 1.05), .73	1.03 (.96, 1.10), .42	1.03 (.95, 1.12), .52
Pseudo R^2	.0005	.003	.002
n	192	192	232
BAME[a]	.69 (.29, 1.65), .40	.95 (.41, 2.20), .90	1.96 (.83, 4.63), .12
Pseudo R^2	.003	.0001	.01
n	183	183	221
Previous miscarriage	.42 (.17, 1.06), .07	**.39 (.16, .95), .04***	**2.81 (1.24, 6.35), .013***
Pseudo R^2	.02	.02	.03
n	166	166	200
Previous live birth	1.03 (.55, 1.94), .93	.79 (.42, 1.47), .45	.85 (.38, 1.89), .69
Pseudo R^2	.00	.002	.001
n	179	179	213
Previous IVF treatment	.55 (.23, 1.34), .19	.42 (.17, 1.03), .06	.94 (.34, 2.62), .90
Pseudo R^2	.007	.01	.0001
n	192	192	231
Months trying to conceive	.99 (.98, 1.00), .22	.99 (.98, 1.00), .18	.99 (.97, 1.01), .39
Pseudo R^2	.007	.008	.006
n	169	169	194
Antral follicle count	1.01 (.995, 1.03), .17	1.02 (.999, 1.04), .07	1.01 (.99, 1.02), .41
Pseudo R^2	.008	.01	.003
n	191	191	231
Fertilisation rate	1.01 (.996, 1.02), .17	1.01 (.995, 1.02), .20	NA
Pseudo R^2	.007	.006	NA
n	192	192	NA

[a]BAME: Black Asian and Minority Ethnic.

Similar univariate logistic regressions were conducted to examine the association between demographic and clinical variables and biochemical pregnancy: age, BMI, ethnicity (BAME/non-BAME), previous miscarriage (yes or no), previous IVF treatment (yes or no), months trying to conceive, antral follicle count, and fertilisation rate. The results showed that being older (OR = 0.92, 95% CI = 0.86–0.98, p = 0.009) and having a previous miscarriage (OR = 0.39, 95% CI = 0.16–0.95, p = 0.04) were significantly and independently associated with a reduced likelihood of becoming biochemically pregnant. However, BMI (OR = 1.03, 95% CI = 0.96–1.10, p = 0.42), ethnicity (OR = 0.95, 95% CI = 0.41–2.20, p = 0.90), previous live birth (OR = 0.79, 95% CI = 0.42–1.47, p = 0.45), previous IVF treatment (OR = 0.42, 95% CI = 0.17–1.03, p = 0.06), months trying to conceive (OR = 1.00, 95% CI = 1.00–1.00, p = 0.48), antral follicle count (OR = 1.02, 95% CI = 1.00–1.04, p = 0.07), and fertilisation rate (OR = 1.01, 95% CI = 0.995–1.02, p = 0.20) were not independently associated with clinical pregnancy 6.3.

Univariate logistic regressions to predict the discontinued treatment outcome showed that age (OR = 0.99, 95% CI = 0.92–1.06, p = 0.75), BMI (OR = 1.03, 95% CI = 0.95–1.12, p = .52), ethnicity (OR = 1.96, 95% CI = 0.83–4.63, p = 0.12), previous live birth (OR = 0.85, 95% CI = 0.38–1.89, p = 0.69), previous IVF treatment (OR = 0.94, 95% CI = 0.34–2.62, p = 0.90), months trying to conceive (OR = 0.99, 95% CI = 0.97–1.01, p = 0.39), and antral follicle count (OR = 1.01, 95% CI = 0.99–1.02, p = 0.41) were not significantly associated with discontinued treatment. However, previous miscarriage significantly predicted discontinued treatment (OR = 2.81, 95% CI = 1.24–6.35, p = 0.013). These results are presented in Table 6.3.

In summary, the univariate regression analyses showed that age was significantly associated with both the clinical and biochemical pregnancy treatment outcomes, and previous miscarriage was significantly positively correlated with not achieving

biochemical pregnancy and discontinuing treatment. Therefore, these two factors were considered covariates in the subsequent models of HCCs in predicting clinical pregnancy, biochemical pregnancy, and discontinued treatment.

6.3.3 Relationship between hair cortisol concentration in the three months before treatment and outcomes of IVF treatment: univariate analysis

Table 6.4 presents 1 cm HCCs and 3 cm HCCs by outcome groups and ELISA kits.

Table 6.4 Hair cortisol concentration (HCC) values in patients by treatment outcome and reference group

	3cm HCC	3cm HCC	1cm HCC	1cm HCC
	IBL assay kit	DGR assay kit	IBL assay kit	DGR assay kit
Clinical pregnancy outcome				
Yes				
(median, IQR, n)	81.28 (27.4), 39	14.43 (8.6), 41	59.95 (30.6), 13	18.46 (10.4), 41
No				
(median, IQR, n)	67.38 (22.4), 55	14.28 (11.4), 52	53.69 (17.2), 15	16.47 (10.2), 53
Biochemical pregnancy outcome				
Yes				
(median, IQR, n)	78.48 (28.4), 48	13.84 (8.9), 43	58.52 (31.1), 16	15.80 (10.0), 43
No				
(median, IQR, n)	61.56 (32.9), 46	14.52 (12.0), 50	56.06 (17.2), 12	16.54 (10.4), 51
Discontinued treatment				
Yes				
(median, IQR, n)	60.01 (75.6), 22	13.88 (8.7), 18	62.32 (26.9), 8	17.11 (6.5), 17
No				
(median, IQR, n)	73.99 (37.8), 94	14.42 (10.8), 93	57.66 (20.0), 28	16.09 (10.2), 94

When looking at 3 cm HCCs from the IBL assay only, the univariate logistic regressions showed that 3 cm HCCs were not significantly associated with clinical pregnancy (primary outcome), biochemical pregnancy, or discontinued treatment (see Table 6.5). Similarly, the 3 cm HCCs from the DGR assay were not significantly associated with any of the treatment outcomes (see Table 6.5). When

looking at 3 cm HCCs (log transformed Z-scores) from both the IBL and DGR assays, the univariate logistic regressions showed that 3 cm HCCs did not predict clinical pregnancy ($OR = 2.09$, 95% CI = 0.61–7.11, $p = 0.24$), biochemical pregnancy ($OR = 2.00$, 95% CI = 0.59–6.78, $p = 0.26$), or discontinued treatment ($OR = 0.44$, 95% CI = 0.09–2.07, $p = 0.30$).

Table 6.5 Univariate logistic regression models showing 3 cm hair cortisol concentrations (HCC) predicting treatment outcomes

	Clinical pregnancy	**Biochemical pregnancy**	**Discontinued treatment**
	odds ratio (95%CI), *p*	*odds ratio* (95%CI), *p*	*odds ratio* (95%CI), *p*
3cm HCC from IBL assay only	1.66 (.73, 3.77), .23	1.75 (.76, 4.01), .19	.58 (.21, 1.60), .29
Pseudo R^2, n	.01, 94	.01, 94	.01, 116
3cm HCC from DGR assay only	1.17 (.57, 2.42), .67	1.06 (.51, 2.18), .88	.86 (.34, 2.18), .75
Pseudo R^2, n	.001, 93	.0002, 93	.001, 111
3cm HCC from both assays[a]	2.09 (.61, 7.11), .24	2.00 (.59, 6.78), .26	.44 (.09, 2.07), .30
Pseudo R^2, n	.01, 187	.01, 187	.01, 227

[a] Log transformed *Z*-scores of 3cm HCC absolute values

6.3.4 Relationship between hair cortisol concentrations in the three months before treatment and outcomes of IVF treatment: multivariate analysis with results from the IBL assay

Multivariate logistic regression was conducted to predict clinical pregnancy (primary outcome), with age, previous miscarriage (yes or no), and 3 cm HCCs from the IBL assay as independent variables. Results showed that 3 cm HCCs from the IBL assay did not predict clinical pregnancy ($OR = 2.53$, 95% CI = 0.85–7.58, $p = 0.10$) after controlling for other factors. Age ($OR = 0.92$, 95% CI = 0.82–1.05, $p = 0.23$) and previous miscarriage ($OR = 0.24$, 95% CI = 0.04–1.29, $p = 0.10$) were also not significantly associated with clinical pregnancy. These results are shown in Table 6.6.

Table 6.6 Multivariate logistic regression models showing 3 cm hair cortisol concentrations (HCCs) from the IBL assay predicting treatment outcomes while controlling for age and miscarriage history

	Clinical pregnancy	**Biochemical pregnancy**	**Discontinued treatment**
	odds ratio (95%CI), *p*	*odds ratio* (95%CI), *p*	*odds ratio* (95%CI), *p*
Age (per unit)	.92 (.82, 1.05), .23	**.86 (.75, .98), .03***	.88 (.77, 1.01), .06
Previous miscarriage (Yes/No)	.24 (.04, 1.29), .10	.28 (.06, 1.30), .10	**4.25 (1.16, 15.64), .03***
3cm HCC from IBL assay (per unit)	2.53 (.85, 7.58), .10	**3.55 (1.05, 12.04), .04***	.49 (.14, 1.70), .26
Pseudo R^2	.08	.12	.09
n	75	75	93

*p<.05, **p<.01

A second multivariate logistic regression was conducted to predict biochemical pregnancy with 3 cm HCCs from the IBL assay while controlling for age and previous miscarriage (see Table 6.6). The results showed that 3 cm HCCs from the IBL assay were significantly associated with biochemical pregnancy, with a per unit increase in 3 cm HCCs associated with 3.6 times greater likelihood of achieving biochemical pregnancy ($OR = 3.55$, 95% CI = 1.05–12.04, $p = 0.04$). A per unit increase in age was also associated with a 14% reduction in the likelihood of achieving biochemical pregnancy ($OR = 0.86$, 95% CI = 0.75–0.98, $p = 0.03$). Previous miscarriage was not significantly associated with biochemical pregnancy in this model ($OR = 0.28$, 95% CI = 0.06–1.30, $p = 0.10$).

A third multivariate logistic regression was conducted to predict discontinued treatment using 3 cm HCCs from the IBL assay while controlling for age and previous miscarriage (see Table 6.6). The 3 cm HCCs from the IBL assay did not predict discontinued treatment after controlling for the other covariates ($OR = 0.49$, 95% CI = 0.14–1.70 $p = 0.26$). Age ($OR = 0.88$, 95% CI = 0.77–1.01, $p = 0.06$) was also not significantly associated with discontinued treatment. However, women with a previous miscarriage were four times more likely to discontinue treatment ($OR = 4.25$, 95% CI = 1.16–15.64, $p = 0.03$) than those without.

In summary, when looking at the associations between the 3 cm HCCs from the IBL assay and treatment outcomes while controlling for age and previous miscarriage, the 3 cm HCCs were not significantly associated with clinical pregnancy and discontinued treatment. However, increased 3 cm HCCs were significantly associated with an increased likelihood of achieving biochemical pregnancy. Increased age was significantly associated with a reduced likelihood of achieving biochemical pregnancy, and previous miscarriage was significantly associated with an increased likelihood of discontinuing treatment.

6.3.5 Relationship between hair cortisol concentration in the three months before treatment and outcomes of IVF treatment: multivariate analysis with results from the DRG assay

Multivariate logistic regression was conducted to predict clinical pregnancy using age, previous miscarriage (yes or no), and 3 cm HCCs from DRG assay results as independent variables. The results showed that 3 cm HCCs from the DRG assay did not predict either clinical pregnancy ($OR = 1.13$, 95% CI = 0.53–2.41, $p = 0.75$), biochemical pregnancy ($OR = 1.00$, 95% CI = 0.47–2.13, $p = 0.99$), or discontinued treatment ($OR = 0.88$, 95% CI = 0.31–2.49, $p = 0.81$) after controlling for age and previous miscarriage. Age was not associated with clinical pregnancy ($OR = 0.97$, 95% CI = 0.88–1.07, $p = 0.52$), biochemical pregnancy ($OR = 0.96$, 95% CI = 0.87–1.05, $p = 0.37$), or discontinued treatment ($OR = 1.07$, 95% CI = 0.95–1.21, $p = 0.25$) in these models. Previous miscarriage was not associated with clinical pregnancy ($OR = 0.62$, 95% CI = 0.17–2.25, $p = 0.46$) or biochemical pregnancy ($OR = 0.56$, 95% CI = 0.15–2.06, $p = 0.38$). However, women with a previous miscarriage were 3.5 times more likely to discontinue treatment ($OR = 3.45$, 95% CI = 1.02–11.66, $p = 0.046$) in these models than

those without. The results are shown in Table 6.7.

Table 6.7 Multivariate logistic regression models showing combined 3 cm hair cortisol concentrations (HCCs) from the DRG assay predicting treatment outcomes while controlling for age and miscarriage history

	Clinical pregnancy	**Biochemical pregnancy**	**Discontinued treatment**
	odds ratio (95%CI), *p*	*odds ratio* (95%CI), *p*	*odds ratio* (95%CI), *p*
Age (per unit)	.97 (.88, 1.07), .52	.96 (.87, 1.05), .37	1.07 (.95, 1.21), .25
Previous miscarriage (Yes/No)	.62 (.17, 2.25), .46	.56 (.15, 2.06), .38	**3.45 (1.02, 11.66), .046***
3cm HCC from DRG assay (per unit)	1.13 (.53, 2.41), .75	1.00 (.47, 2.13), .99	.88 (.31, 2.49), .81
Pseudo R^2	.01	.01	.06
n	86	86	102

*p<.05, **p<.01

In summary, when looking at the associations between the 3 cm HCCs from the DRG assay and treatment outcomes while controlling for age and previous miscarriage, the 3 cm HCCs were not significantly associated with clinical pregnancy, biochemical pregnancy, or discontinued treatment. Age was not significantly associated with the three treatment outcomes in these models, but previous miscarriage was significantly associated with an increased risk of discontinuing treatment.

6.3.6 Relationship between hair cortisol concentration in the three months before treatment and outcomes of IVF treatment: multivariate analysis with combined results from two assays

Multivariate logistic regression models were also conducted with 3 cm HCCs from the combined assay results to predict clinical pregnancy, biochemical pregnancy, and discontinued treatment while controlling for age and previous miscarriage

(see Table 6.8). The results showed that the 3 cm HCCs from the combined assay results did not predict clinical pregnancy ($OR = 2.56$, 95% CI = 0.63–10.37, $p = 0.19$), biochemical pregnancy ($OR = 2.55$, 95% CI = 0.62–10.52, $p = 0.20$), or discontinued treatment ($OR = 0.31$, 95% CI = 0.05–1.81, $p = 0.19$) after controlling for age and previous miscarriage. Age was not significantly associated with clinical pregnancy ($OR = 0.95$, 95% CI = 0.88–1.03, $p = 0.22$), biochemical pregnancy ($OR = 0.93$, 95% CI = 0.86–1.00, $p = 0.051$), or discontinued treatment ($OR = 0.98$, 95% CI = 0.90–1.06, $p = 0.58$) in these models. Previous miscarriage was not significantly associated with clinical pregnancy ($OR = 0.40$, 95% CI = 0.15–1.10, $p = 0.08$) or biochemical pregnancy ($OR = 0.40$, 95% CI = 0.15–1.05, $p = 0.06$). However, women with a previous miscarriage were 3.5 times more likely to discontinue treatment ($OR = 3.51$, 95% CI = 1.48–8.29, $p = 0.004$) than those without.

Table 6.8 Multivariate logistic regression models showing combined 3 cm hair cortisol concentrations (HCCs) from both assays predicting treatment outcomes while controlling for age and miscarriage history

	Clinical pregnancy	**Biochemical pregnancy**	**Discontinued treatment**
	odds ratio (95%CI), *p*	*odds ratio* (95%CI), *p*	*odds ratio* (95%CI), *p*
Age (per unit)	.95 (.88, 1.03), .22	.93 (.86, 1.00), .051	.98 (.90, 1.06), .58
Previous miscarriage (Yes/No)	.40 (.15, 1.10), .08	.40 (.15, 1.05), .06	**3.51 (1.48, 8.29), .004****
3cm HCC from both assays[a] (per unit)	2.56 (.63, 10.37), .09	2.55 (.62, 10.52), .20	.31 (.05, 1.81), .19
Pseudo R^2	.03	.04	.05
n	131	161	195

*p<.05, **p<.01

[a] Log-transformed z scores of 3cm HCC absolute values

In summary, the 3 cm HCCs from the combined assay results were not significantly associated with clinical pregnancy, biochemical pregnancy, or discontinued treatment outcomes after controlling for age and miscarriage history. Age was not associated with the three treatment outcomes in these models, but previous miscarriage was significantly associated with an increased risk of discontinuing treatment.

6.3.7 Relationship between hair cortisol concentrations in the month before treatment and outcomes of IVF treatment: univariate analysis

The same analyses were repeated for 1 cm HCCs (see Table 6.9 to Table 6.12). When looking at 1 cm HCCs from the IBL assay only, the univariate logistic regression showed that 1 cm HCCs were weakly but significantly associated with clinical pregnancy ($OR = 30.41$, 95% CI = 1.02–908.59, $p = 0.049$). However, the associations between 1 cm HCCs (IBL assay) and either biochemical pregnancy

or discontinued treatment were non-significant (see Table 6.9). The 1 cm HCCs from the DGR assay were not significantly associated with any of the treatment outcomes (see Table 6.9). When considering 1 cm HCCs (log transformed *Z*-scores) from both the IBL and DGR assays, the univariate logistic regression showed that 1 cm HCCs did not predict clinical pregnancy ($OR = 1.37$, 95% CI = 0.36–5.17, $p = 0.65$), biochemical pregnancy ($OR = 0.79$, 95% CI = 0.21–2.96, $p = 0.73$), or discontinued treatment ($OR = 0.62$, 95% CI = 0.12–3.27, $p = 0.57$). The results from these univariate logistic regressions are presented in Table 6.9.

Table 6.9 Univariate logistic regression models showing 1 cm hair cortisol concentrations (HCCs) predicting treatment outcomes

	Clinical pregnancy	**Biochemical pregnancy**	**Discontinued treatment**
	odds ratio (95%CI), *p*	*odds ratio* (95%CI), *p*	*odds ratio* (95%CI), *p*
Univariable			
1cm HCC from IBL assay	**30.41 (1.02, 908.59), .049***	3.44 (.29, 41.44), .330	1.51 (.15, 15.61), .73
Pseudo R^2, n	.15, 28	.03, 28	.003, 36
1cm HCC from DGR assay	.75 (.29, 1.95), .56	.67 (.26, 1.75), .41	.50 (.13, 1.97), .33
Pseudo R^2, n	.003, 94	.01, 94	.01, 111
1cm HCC from both assays[a]	1.37 (.36, 5.17), .65	.79 (.21, 2.96), .73	.62 (.12, 3.27), .57
Pseudo R^2, n	.001, 122	.001, 122	.002, 147

[a] Log-transformed z scores of 1cm HCC absolute values

6.3.8 Relationship between hair cortisol concentrations in the month before treatment and outcomes of IVF treatment: multivariate analysis with results from the IBL assay

Multivariate logistic regression was conducted to predict clinical pregnancy using age, previous miscarriage (yes or no), and 1 cm HCCs from the IBL assay as independent variables (see Table 6.10). The results showed that 1 cm HCCs from the IBL assay were significantly associated with clinical pregnancy ($OR = 74.97$,

95% CI = 1.11–5078.59, $p = 0.045$) after controlling for the other factors. Age ($OR = 0.89$, 95% CI = 0.68–1.18, $p = 0.42$) was not significantly associated with clinical pregnancy. Due to the small sample size (n = 22) in this model, the previous miscarriage variable was omitted.

Table 6.10 Multivariable logistic regression models showing 1 cm hair cortisol concentrations (HCCs) from the IBL assay predicting treatment outcomes while controlling for age and miscarriage history

	Clinical pregnancy	**Biochemical pregnancy**	**Discontinued treatment**
	odds ratio (95%CI), *p*	*odds ratio* (95%CI), *p*	*odds ratio* (95%CI), *p*
Age (per unit)	.89 (.68, 1.18), .42	.83 (.64, 1.08), .17	.92 (.72, 1.17), .47
Previous miscarriage (Yes/No)	Omitted variable	.66 (.02, 17.69), .81	2.79 (.16, 47.95), .48
1cm HCC from IBL assay (per unit)	**74.97 (1.11 5078.59), .045***	4.98 (.24, 104.02), .30	1.18 (.15, 21.74), .65
Pseudo R^2	.23	.12	.03
n	22	24	31

*p<.05, **p<.01

A second multivariate logistic regression was conducted to predict biochemical pregnancy with 1 cm HCCs from the IBL assay while controlling for age and previous miscarriage (see Table 6.10). The results showed that 1 cm HCCs from the IBL assay ($OR = 4.98$, 95% CI = 0.24–104.02, $p = 0.30$), age ($OR = 0.83$, 95% CI = 0.64–1.08, $p = 0.17$), and previous miscarriage ($OR = 0.66$, 95% CI = 0.02–17.69, $p = 0.81$) were not significantly associated with biochemical pregnancy in this model.

Similarly, 1 cm HCCs from the IBL assay ($OR = 1.18$, 95% CI = 0.15–21.74 $p = 0.65$), age ($OR = 0.92$, 95% CI = 0.72–1.17, $p = 0.47$), and previous miscarriage ($OR = 2.79$, 95% CI = 0.16–47.95, $p = 0.48$) were not significantly associated with discontinued treatment (see Table 6.10).

In summary, when looking at the associations between 1 cm HCCs from the IBL assay and treatment outcomes while controlling for age and previous miscarriage,

1 cm HCCs were significantly associated with clinical pregnancy but not biochemical pregnancy or discontinued treatment. Age and previous miscarriage were not significantly associated with the three treatment outcomes in these models.

6.3.9 Relationship between hair cortisol concentrations in the month before treatment and outcomes of IVF treatment: multivariate analysis with results from the DRG assay

A multivariate logistic regression was conducted to predict clinical pregnancy using age, previous miscarriage (yes or no), and 1 cm HCCs from the DRG assay as independent variables. The results showed that 1 cm HCCs from the DRG assay did not predict clinical pregnancy (OR = 0.76, 95% CI = 0.28–2.03, p = 0.58), biochemical pregnancy (OR = 0.66, 95% CI = 0.25–1.78, p = 0.42), or discontinued treatment (OR = 0.48, 95% CI = 0.10–2.22, p = 0.25) after controlling for age and previous miscarriage. Age was not associated with clinical pregnancy (OR = 0.97, 95% CI = 0.88–1.06, p = 0.48), biochemical pregnancy (OR = 0.95, 95% CI = 0.87–1.05, p = 0.35), or discontinued treatment (OR = 1.10, 95% CI = 0.96–1.25, p = 0.16) in these models. Previous miscarriage was not associated with clinical pregnancy (OR = 0.57, 95% CI = 0.16–2.03, p = 0.38), biochemical pregnancy (OR = 0.51, 95% CI = 0.14–1.84, p = 0.30), or discontinued treatment (OR = 2.58, 95% CI = 0.72–9.19, p = 0.14). The results are shown in Table 6.11.

Table 6.11 Multivariate logistic regression models showing combined 1 cm hair cortisol concentrations (HCCs) from the DRG assay predicting treatment outcomes while controlling for age and miscarriage history

	Clinical pregnancy	**Biochemical pregnancy**	**Discontinued treatment**
	odds ratio (95%CI), *p*	*odds ratio* (95%CI), *p*	*odds ratio* (95%CI), *p*
Age (per unit)	.97 (.88, 1.06), .48	.95 (.87, 1.05), .35	1.10 (.96, 1.25), .16
Previous miscarriage (Yes/No)	.57 (.16, 2.03), .38	.51 (.14, 1.84), .30	2.58 (.72, 9.19), .14
3cm HCC from IBL assay (per unit)	.76 (.28, 2.03), .58	.66 (.25, 1.78), .42	.48 (.10, 2.22), .25
Pseudo R^2	.01	.02	.06
n	87	87	102

*p<.05, **p<.01

In summary, when looking at the associations between 1 cm HCCs from the DRG assay and treatment outcomes while controlling for age and previous miscarriage, 1 cm HCCs were not significantly associated with clinical pregnancy, biochemical pregnancy, or discontinued treatment. Age and previous miscarriage were also not significant predictors in these models.

6.3.10 Relationship between hair cortisol concentrations in the month before treatment and outcomes of IVF treatment: multivariate analysis with combined results from two assays

Multivariate logistic regression models were conducted with 1 cm HCCs from the combined assay results to predict treatment outcomes while controlling for age and previous miscarriage (see Table 6.12). The results showed that 1 cm HCCs did not predict clinical pregnancy (OR = 1.49, 95% CI = 0.37–5.94, p = 0.58), biochemical pregnancy (OR = 0.79, 95% CI = 0.20–3.14, p = 0.74), or discontinued treatment (OR = 0.78, 95% CI = 0.14–4.46, p = 0.78) after

controlling for the other factors. Age was not significantly associated with clinical pregnancy (OR = 0.96, 95% CI = 0.88–1.05, p = 0.35), biochemical pregnancy (OR = 0.93, 95% CI = 0.85–1.02, p = 0.14), or discontinued treatment (OR = 1.04, 95% CI = 0.94–1.15, p = 0.48) in these models. Similarly, previous miscarriage was not significantly associated with clinical pregnancy (OR = 0.41, 95% CI = 0.12–1.39, p = 0.15), biochemical pregnancy (OR = 0.48, 95% CI = 0.15–1.54, p = 0.22), or discontinued treatment (OR = 2.24, 95% CI = 0.74–6.75, p = 0.15) in these models.

Table 6.12 Multivariate logistic regression models showing combined 1 cm hair cortisol concentrations (HCCs) from both assays predicting treatment outcomes while controlling for age and miscarriage history

	Clinical pregnancy	**Biochemical pregnancy**	**Discontinued treatment**
	odds ratio (95%CI), *p*	*odds ratio* (95%CI), *p*	*odds ratio* (95%CI), *p*
Age (per unit)	.96 (.88, 1.05), .35	.93 (.85, 1.02), .14	1.04 (.94, 1.15), .48
Previous miscarriage (Yes/No)	.41 (.12, 1.39), .15	.48 (.15, 1.54), .22	2.24 (.74, 6.75), .15
1cm HCC from both assays[a] (per unit)	1.49 (.37, 5.94), .58	.79 (.20, 3.14), .74	.78 (.14, 4.46), .78
Pseudo R^2	.02	.03	.02
n	111	111	133

*p<.05, **p<.01

[a] Log-transformed z scores of 1cm HCC absolute values

In summary, the 1 cm HCCs from the combined assay results were not significantly associated with clinical pregnancy, biochemical pregnancy, or discontinued treatment outcomes after controlling for age and miscarriage history. Age and previous miscarriage were also not significantly associated with the three outcomes in these models.

6.3.11 Sensitivity analysis using multiple imputation

To account for the influence of missing data in relevant variables in the multivariate regression models, sensitivity analysis was conducted with MI datasets (n = 20). The results were largely unchanged, i.e., neither 3 cm HCCs nor 1 cm HCCs from the combined assay results significantly predicted clinical pregnancy, biochemical pregnancy, or discontinued treatment. However, increasing age was significantly associated with a decreasing likelihood of achieving clinical pregnancy and biochemical pregnancy, and previous miscarriage was significantly associated with an increased likelihood of discontinuing treatment (see Table 6.13 and 6.14).

Table 6.13 Missing value table

Variables	Missing (n)	% of 192
Age	2	1%
Miscarriage	48	25%
Combined 3cm HCCs from both assays	5	3%
Combined 1cm HCCs from both assays	101	53%

Table 6.14 Sensitivity analysis multivariate logistic regression models using multiple imputation datasets to show combined 3 cm hair cortisol concentrations (HCCs) from both assays predicting treatment outcomes while controlling for age and previous miscarriage

	Clinical pregnancy	**Biochemical pregnancy**	**Discontinued treatment**
	odds ratio (95%CI), *p*	*odds ratio* (95%CI), *p*	*odds ratio* (95%CI), *p*
Age (per unit)	**-.07 (-.14, -.003), .04***	**-.09 (-.16, -.02), .010****	-.02 (-.10, .05), .59
Previous miscarriage (Yes/No)	-.68 (-1.62, .26), .15	-.76 (-1.69, .17), .11	**1.00 (.17, 1.83), .02***
3cm HCC from both assays[a] (per unit)	.92 (-.34, 2.18), .15	.95 (-.33, 2.23), .14	-.97 (-2.56, .62), .23
N	192	192	232

*p<.05, **p<.01

Table 6.15 Sensitivity analysis multivariate logistic regression models using multiple imputation datasets to show combined 1 cm hair cortisol concentrations (HCCs) from both assays predicting treatment outcomes while controlling for age and previous miscarriage

	Clinical pregnancy	**Biochemical pregnancy**	**Discontinued treatment**
	odds ratio (95%CI), *p*	*odds ratio* (95%CI), *p*	*odds ratio* (95%CI), *p*
Age (per unit)	**-.07 (-.14, -.00), .050***	**-.09 (-.15, -.02), .014***	-.02 (-.10, .05), .53
Previous miscarriage (Yes/No)	-.62 (-1.54, .31), .19	-.70 (-1.61, .22), .14	**.94 (.12, 1.76), .02***
1cm HCC from both assays[a] (per unit)	.33 (-.94, 1.60), .61	.13 (-1.14, 1.40), .84	-.25 (-1.77, 1.27), .74
N	192	192	232

*p<.05, **p<.01

6.4 Part 2: lifestyle of women undergoing IVF treatment

6.4.1 Lifestyle in patients undergoing IVF treatment

A summary of lifestyle factors of patients undergoing IVF treatment who completed the survey and the subsequent treatment is presented in Table 6.16. The results by treatment outcomes (i.e., clinical pregnancy, biochemical pregnancy, and discontinued treatment) are also presented in Table 6.16. Approximately 60% of the whole cohort reported poor sleep quality (scores >5), according to the cut-off of the Pittsburgh Sleep Quality Index (PSQI). Other literature reported that the mean scores for sleep quality measured by the PSQI among healthy young women (N=261, mean age = 26) was 3.07 (Muzni et al., 2021). Independent sample t-test showed that the mean score for sleep quality in the current cohort was significantly higher than reported in the literature (p<0.001). Approximately 65% of the current cohort engaged in moderate to high levels of physical activity, according to the cut-offs for low, moderate, and high levels on the Global

Physical Activity Questionnaire (GPAQ). Previous literature also observed a high proportion of women who engaged in regular exercise (46%–62%, Domar et al. (2015)). The average caffeine consumption was approximately 2,329.7 mg/month (approximately 78 mg/day) in this cohort, which is lower than the 300 mg/day recommended by the UK government's Food Standards Agency (CARE Study Group and others, 2008). Results showed that the levels of alcohol consumption and smoking behaviour were low in this cohort (see Table 6.16).

Table 6.16 Lifestyle factors of patients undergoing *in vitro* fertilisation/intracytoplasmic sperm injection treatment by treatment outcomes

	Whole cohort (n=225)	Completed treatment (n=192)	Clinically pregnant (n=110)	Not clinically pregnant (n=82)	Biochemically pregnant (n=93)	Not biochemically pregnant (n=99)	Discontinued treatment (n=40)
Sleep	6.9 (3.3)	6.6 (3.1)	6.7 (3.2)	6.3 (2.9)	6.4 (2.9)	6.7 (3.2)	7.8 (3.8)
Poor sleep quality (score > 5, n/%)	133 (60.2)	88 (56.4)	41 (61.2)	47 (52.8)	46 (61.3)	42 (51.9)	24 (68.6)
Physical activity							
Total hours/week	43.9 (69.2)	45.9 (71.2)	51.8 (81.4)	37.9 (54.3)	39.3 (59.3)	51.9 (80.5)	49.5 (73.7)
Low (n/%)	78 (25.3%)	55 (34.6%)	24 (35.3%)	31 (34.1%)	27 (35.5%)	28 (33.7%)	10 (28.6%)
Medium (n/%)	87 (39.4%)	60 (37.7%)	27 (39.7%)	33 (26.3%)	30 (39.5%)	30 (36.1%)	16 (45.7%)
High (n/%)	56 (25.3%)	44 (27.7%)	17 (25.0%)	27 (29.7%)	19 (25.0%)	25 (30.1%)	9 (25.7%)
Caffeine consumption (mg/-month)	2329.7 (3009.3)	2442.5 (3078.5)	2957.4 (3487.0)	2062.0 (2694.9)	2784.2 (3378.9)	2133.4 (2763.2)	2005.7 (3009.4)
Alcohol consumption	3.1 (1.7)	3.1 (1.7)	3.0 (1.7)	3.2 (1.7)	3.2 (1.7)	3.0 (1.7)	2.6 (1.6)
Smoking							
Yes	5	5	2	3	2	3	0
No	220	155	66	89	74	81	35

6.4.2 Lifestyle and treatment outcomes in patients undergoing IVF treatment: univariate analyses

Univariate logistic regressions showed that alcohol consumption, sleep, physical activity, caffeine consumption, and smoking behaviours were not significantly associated with the clinical pregnancy outcome (see Table 6.17).

Table 6.17 Univariate logistic regressions showing the association between lifestyle factors and *in vitro* fertilisation/intracytoplasmic sperm injection treatment outcomes

	Clinical pregnancy	**Biochemical pregnancy**	**Discontinued treatment**
	odds ratio (95%CI), *p*	*odds ratio* (95%CI), *p*	*odds ratio* (95%CI), *p*
Alcohol consumption	1.10 (.89, 1.35), .38	1.07 (.87, 1.31), .54	.83 (.62, 1.11), .22
Pseudo R^2	.004	.002	.01
n	124	124	145
Sleep (per unit)	.76 (.41, 1.40), .37	.73 (.40, 1.35), .32	2.12 (.97, 4.66), .06
Pseudo R^2	.004	.005	.02
n	156	156	191
Physical activity (minutes/month)	.78 (.59, 1.03), .08	.78 (.59, 1.03), .08	.92 (.68, 1.26), .61
Pseudo R^2	.02	.02	.002
n	129	139	161
Caffeine consumption (mg/month)	1.06 (.86, 1.31), .57	.97 (.79, 1.19), .78	.88 (.69, 1.11), .28
Pseudo R^2	.002	.001	.01
n	129	129	157
Smoking	1.11 (.18, 6.85), .91	1.37 (.22, 8.43), .73	N/A
Pseudo R^2	.00	.001	N/A
n	160	160	N/A

Similar univariate logistic regressions were conducted to examine the association between the same lifestyle factors and biochemical pregnancy and discontinued treatment (see Table 6.17). The results showed that alcohol consumption, sleep, physical activity, and caffeine consumption were not significantly associated with the biochemical pregnancy and discontinued treatment outcomes (see Table 6.17). Smoking was not associated with the biochemical pregnancy outcome, but none

of the patients who discontinued treatment reported smoking; therefore, no regression analysis was conducted.

6.4.3 The association between lifestyle factors and 3 cm hair cortisol concentrations: univariate analyses

Univariate linear regression models were conducted to examine the associations between lifestyle factors and 3 cm HCCs from the IBL assay, 3 cm HCCs from the DRG assay, and 3 cm HCCs from the combined assay results (see Table 6.18). The results showed that alcohol consumption, sleep, physical activity (minutes/-month), caffeine consumption (mg/month) over the past month, and smoking (yes or no) were not significantly associated with 3 cm HCCs from the IBL assay, 3 cm HCCs from the DRG assay, or 3 cm HCCs from the combined assay results.

6.4.4 The association between lifestyle and 1 cm hair cortisol concentrations: univariate analyses

Univariate linear regressions were conducted to examine the associations between lifestyle factors and 1 cm HCCs (Table 6.19). Alcohol consumption, sleep, physical activity, and caffeine consumption over the past month were not significantly associated with 1 cm HCCs from the IBL assay, 1 cm HCCs from the DRG assay, or 1 cm HCCs from the combined assay results. Smoking was not significantly associated with 1 cm HCCs from the DRG assay or 1 cm HCCs from the combined assay results. The number of participants who smoked (n = 2) was lower than the required number for linear regression thus such analysis for the association between smoking and 1 cm HCCs from the IBL assay was not conducted.

Table 6.18 Univariate linear regression models showing lifestyle predictors of 3 cm hair cortisol concentrations (HCCs) from the independent assays and combined assay results

	3cm HCC from IBL assay	3cm HCC from DRG assay	3cm HCC from combined assay results
	B (95%CI), *p*	B (95%CI), *p*	B (95%CI), *p*
Alcohol consumption	.01 (-.05, .08), .69	.04 (-.04, .12), .33	.01 (-.01, .03), .36
Adjusted R^2	-.01	-.001	-.00
n	85	78	163
Sleep (per unit)	.11 (-.06, .28), .22	.12 (-.10, .33), .28	.05 (-.01, .11), .10
Adjusted R^2	.004	.002	.01
n	117	102	219
Physical activity (minutes/month)	-.04 (-.11, .04), .37	-.001 (-.11, .11), .98	-.01 (-.04, .02), .46
Adjusted R^2	-.002	-.01	-.00
n	98	78	176
Caffeine consumption (mg/month)	.03 (-.03, .09), .33	.06 (-.01, .13), .10	.02 (-.00, .04), .07
Adjusted R^2	-.003	.02	.01
n	84	94	178
Smoking	.20 (-.52, .91), .59	-.10 (-.73, .54), .77	.02 (-.20, .23), .89
Adjusted R^2	-.01	-.01	-.00
n	120	103	223

Table 6.19 Univariate linear regression models showing lifestyle predictors of 1 cm hair cortisol concentrations (HCCs) from independent assays and the combined assay results

	1cm HCC from IBL assay	1cm HCC from DRG assay	1cm HCC from combined assay results
	B (95%CI), *p*	B (95%CI), *p*	B (95%CI), *p*
Alcohol consumption	.01 (-.06, .07), .79	.03 (-.03, .09), .27	.01 (-.02, .04), .35
Adjusted R^2	-.04	.003	-.00
n	27	78	105
Sleep (per unit)	.07 (-.13, .27), .49	.06 (-.11, .22), .50	.03 (-.05, .12), .41
Adjusted R^2	-.01	-.01	-.00
n	37	102	139
Physical activity (minutes/month)	-.07 (-.18, .03), .17	-.01 (-.09, .07), .82	-.02 (-.06, .02), .41
Adjusted R^2	.03	-.01	-.00
n	31	78	109
Caffeine consumption (mg/month)	.0003 (-.07, .07), .99	.11 (-.04, .06), .68	.01 (-.02,.03), .66
Adjusted R^2	-.03	-.01	-.01
n	32	94	126
Smoking	NR	-.21 (-.70, .28), .40	-.15 (-.45, .16), .35
Adjusted R^2	NR	-.003	-.00
n	NR	103	140

6.5 Part 3: mood of women undergoing IVF treatment

6.5.1 Mood and treatment outcomes in patients undergoing IVF treatment

A summary of lifestyle factors of patients undergoing IVF treatment who completed the survey and the subsequent treatment is presented in Table 6.16. The results by treatment outcomes (i.e., clinical pregnancy, biochemical pregnancy, and discontinued treatment) are also presented in Table 6.16.

Table 6.20 Mood of patients undergoing IVF by treatment outcome[a]

	Whole cohort (n=225)	Completed treatment (n=192)	Clinically pregnant (n=110)	Not clinically pregnant (n=82)	Biochemically pregnant (n=93)	Not biochemically pregnant (n=99)	Discontinued treatment (n=40)
Anxiety	48.0 (12.9)	47.8 (12.2)	48.8 (13.0)	46.5 (10.9)	46.9 (11.7)	48.7 (12.7)	51.2 (14.7)
Impact of events	27.4 (15.7)	28.5 (15.7)	30.4 (16.6)	26 (14.2)	26.7 (14.5)	30.3 (16.7)	29.4 (14.8)
Depression							
Depression total score	18.8 (7.0)	19.1 (7.0)	19.8 (7.6)	18.1 (5.9)	18.4 (6.0)	19.8 (7.8)	17.8 (7.6)
Clinical level[b] (scores $\geq$ 16) (n,%)	186 (69.7%)	137 (71.4%)	59 (72.0%)	78 (70.9%)	68 (73.1%)	69 (69.7%)	24 (60.0%)
Positive mood	30.0 (7.9)	30.0 (8.0)	29.2 (8.3)	31.0 (7.7)	30.4 (7.7)	29.6 (8.1)	30.0 (8.5)
Negative mood	22.5 (8.2)	22.9 (7.9)	23.9 (8.4)	21.6 (7.1)	22.1 (7.5)	23.6 (8.3)	22.0 (9.3)

[a] Data presented in this table are means (*SD*) unless specified otherwise.
[b] A cut off score of 16 or higher indicating clinical depression. (Radloff, 1977; Roberts and Vernon, 1983)

Univariate logistic regressions showed that anxiety, impact of events, depression, positive mood, and negative mood were not significantly associated with the clinical pregnancy outcome (see Table 6.21).

Table 6.21 Univariate linear regression models showing mood predictors of 3 cm hair cortisol concentrations (HCCs) from the independent assays and the combined assay results

	Clinical pregnancy	**Biochemical pregnancy**	**Discontinued treatment**
	odds ratio (95%CI), *p*	*odds ratio* (95%CI), *p*	*odds ratio* (95%CI), *p*
Independent variable			
Anxiety	.98 (.96, 1.01), .23	.99 (.96, 1.01), .37	1.02 (.99, 1.05), .15
Pseudo R^2	.01	.004	.01
n	160	160	195
Independent variable			
Impact of events	.98 (.96, 1.00), .10	.99 (.96, 1.01), .72	1.00 (.98, 1.03), .79
Pseudo R^2	.01	.003	.00
n	139	160	172
Independent variable			
Depression	.28 (.04, 2.30), .24	.41 (.05, 3.24), .40	.22 (.02, 2.60), .23
Pseudo R^2	.01	.003	.01
n	160	160	195
Independent variable			
Positive mood	1.03 (.99, 1.07), .20	1.01 (.97, 1.05), .56	1.00 (.95, 1.05), .99
Pseudo R^2	.01	.002	.00
n	144	144	175
Independent variable			
Negative mood	.46 (.16, .16), .13	.57 (.21, 1.54), .27	.60 (.19, 1.90), .39
Pseudo R^2	.01	.01	.005
n	144	144	175

Similar univariate logistic regressions were conducted to examine the association between mood factors and biochemical pregnancy and discontinued treatment. The results showed that anxiety, impact of events, depression, positive mood, and negative mood were not significantly associated with either the biochemical pregnancy or discontinued treatment outcome (see Table 6.21).

6.5.2 The association between mood and 3 cm hair cortisol concentrations: univariate analyses

Univariate linear regressions were conducted to examine the independent associations between mood factors and 3 cm HCCs from the IBL assay, 3 cm HCCs from the DRG assay, and 3 cm HCCs from the combined assay results (see Table 6.22). The results showed that anxiety, impact of events, positive mood, and negative mood were not significantly associated with 3 cm HCCs from the IBL assay, 3 cm HCC from the DRG assay, or 3 cm HCCs from the combined assay results. However, depression was significantly positively associated with 3 cm HCCs from the IBL assay ($B = 0.63$, 95% CI: 0.03–1.24, $p = 0.04$).

Table 6.22 Univariate linear regression models showing mood predictors of 3 cm hair cortisol concentrations (HCCs) from the independent assays and the combined assay results

	3cm HCC from the IBL assay	3cm HCC from the DRG assay	3cm HCC from combined assay results
	B (95%CI), *p*	B (95%CI), *p*	B (95%CI), *p*
Independent variable Anxiety	.002 (-.005, .01), .52	-.07 (-.01, .01), .10	-.00 (-.00, .00), .59
Adjusted R^2	-.005	.02	-.00
n	120	103	223
Independent variable Depression	**.63 (.03, 1.24), .04***	-.13 (-.81, .56), .71	.14 (-.06, .34), .17
Adjusted R^2	.03	-.01	.00
n	120	103	223
Independent variable Impact of events	.003 (-.003, .01), .30	-.004 (-.01, .002), .20	.00 (-.00, .00), .97
Adjusted R^2	.001	.07	-.01
n	117	83	200
Independent variable Positive mood	.005 (-.01, .02), .44	.008 (-.01, .02), .23	.00 (-.00, .01), .16
Adjusted R^2	-.004	.004	.00
n	98	103	201
Independent variable Negative mood	.10 (-.19, .40), .49	-.06 (-.36, .24), .69	.01 (-.08, .11), .76
Adjusted R^2	-.01	-.01	-.00
n	98	103	201

6.5.3 The association between mood and 1 cm hair cortisol concentrations: univariate analyses

Univariate linear regressions were conducted to examine the independent associations between mood factors and 1 cm HCCs from the IBL assay, 1 cm HCCs from the DRG assay, and 1 cm HCCs from the combined assay results (see Table 6.23). The results showed that anxiety, depression, impact of events, positive mood, and negative mood were not significantly associated with 1 cm HCCs from the IBL

assay, 1 cm HCCs from the DRG assay, or 1 cm HCCs from the combined assay results.

Table 6.23 Univariate linear regression models showing mood predictors of 1 cm hair cortisol concentrations (HCCs) from the independent assays and the combined assay results

	1cm HCC from the IBL assay	1cm HCC from the DRG assay	1cm HCC from combined assay results
	B (95%CI), p	B (95%CI), p	B (95%CI), p
Independent variable Anxiety	.01 (-.002, .02), .14	-.002 (-.01, .004), .51	.00 (-.00, .00), .96
Adjusted R^2	.03	-.01	-.01
n	37	103	140
Independent variable Depression	.45 (-.41, 1.31), .29	-.15 (-.68, .38), .57	-.03 (-.32, .26), .82
Adjusted R^2	.004	-.01	-.01
n	37	103	140
Independent variable Impact of events	.004 (-.004, .01), .30	-.002 (-.01, .003), .37	-.00 (-.00, .00), .68
Adjusted R^2	.003	-.002	-.01
n	37	83	120
Independent variable Positive mood	.01 (-.01, .02), .45	.003 (-.01, .01), .58	.00 (-.00, .01), .36
Adjusted R^2	-.01	-.01	-.00
n	37	103	140
Independent variable Negative mood	.31 (-.05, .68), .09	-.004 (-.23, .23), .97	.05 (-.08, .17), .47
Adjusted R^2	.05	-.01	-.00
n	37	103	140

6.5.4 Correlations between lifestyle and mood factors

Correlations between lifestyle and mood factors were examined using Person's r (see Table 6.24). The results showed that all the mood factors (i.e., anxiety, depression, impact of event, positive mood, and negative mood) were significantly correlated ($p < 0.001$). Poorer sleep quality was significantly positively corre-

lated with anxiety, depression, impact of event, and negative mood ($p < 0.001$) and significantly negatively correlated with positive mood (r = -0.25, $p < 0.001$). Caffeine consumption was significantly positively correlated with alcohol consumption ($p < 0.001$). Physical activity was significantly positively correlated with caffeine consumption ($p < 0.001$) and weakly but significantly correlated with negative mood ($r = 0.15$, $p = 0.048$).

Table 6.24 Correlations between lifestyle and mood factors

	Alcohol consumption	Sleep (per unit)	Physical activity (minutes/month)	Caffeine consumption (mg/month)	Anxiety	depression	Impact of event	Positive mood
Alcohol consumption								
Sleep (per unit)	.09							
Physical activity (minutes/month)	.12	.06						
Caffeine consumption (mg/month)	**.25****	-.03	**.24*****					
Anxiety	-.06	**.34*****	.06	-.06				
Depression	.01	**.43*****	.12	.03	**.57*****			
Impact of event	-.09	**.28*****	.03	-.03	**.57*****	**.50*****		
Positive mood	.06	**-.25*****	.05	.01	**-.32*****	**-.26*****	**-.25*****	
Negative mood	-.02	**.30*****	**.15***	.02	**.64*****	**.63*****	**.59*****	**-.30*****

*p<.05, **p<.01, ***p<.001

6.5.5 Depression predicting biochemical pregnancy through 3 cm hair cortisol concentrations: mediation analysis

In the previous sections, it was shown that 3 cm HCCs from the IBL assay were significantly associated with biochemical pregnancy following treatment, and depression scores were significantly positively associated with 3 cm HCCs from the IBL assay. These results suggested a possible mediation effect of 3 cm HCCs on the association between depression and biochemical pregnancy. Therefore, a path analysis was conducted to examine this mediation effect, with age and previous miscarriage as covariates of 3 cm HCCs and biochemical pregnancy (see Figure 6.2), based on previous analyses. The direct effect of depression on 3 cm HCCs (a), direct effect of 3 cm HCCs on biochemical pregnancy (b), direct effect of depression on biochemical pregnancy (c'), and the indirect effect of depression on biochemical pregnancy through 3 cm HCCs (c = a × b) were examined.

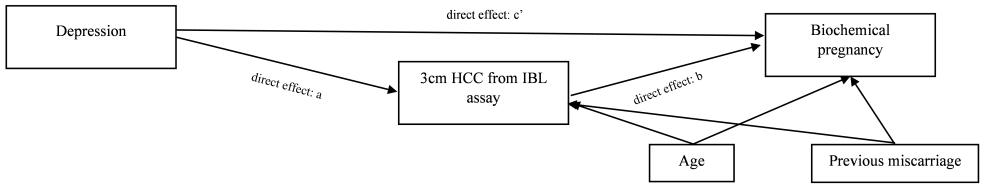


Figure 6.2 Path analysis with depression predicting biochemical pregnancy through 3 cm hair cortisol concentrations (HCCs) from the IBL assay

The results (see Figure 6.3) showed that the effect size of the path from depression to 3 cm HCCs from the IBL assays was significant (B = 0.91, 95% CI = 0.05–1.77, $p = 0.03$) while controlling for age and previous miscarriage. The effect size of the path from 3 cm HCCs to biochemical pregnancy was also significant (B = 1.58, 95% CI = 0.18–2.98, $p = 0.03$). The direct effect of depression on biochemical pregnancy (*c'*) was non-significant (effect size = -3.15, 95% CI = -7.29–0.99, $p = 0.14$). However, the total indirect effect of depression on clinical pregnancy

through 3 cm HCCs (c) was significant (effect size = 1.44, bootstrapped 95% CI = 0.04–4.52).

Based on the thesis model, a sensitivity analysis (another path analysis) was conducted to examine the effects of depression on biochemical pregnancy through 3 cm HCCs from the IBL assay, with age and previous miscarriage as covariates of only biochemical pregnancy (see Figure 5.1). Similar to the previous path analysis, the direct effect of depression on 3 cm HCCs (a), direct effect of 3 cm HCCs on biochemical pregnancy (b), direct effect of depression on biochemical pregnancy (*c'*), and indirect effect of depression on biochemical pregnancy through 3 cm HCCs (c = a × b) were examined. The results (see Figure 6.4) showed that the effect size of the path from depression to 3 cm HCCs from the IBL assays was not significant (B = 0.71, 95% CI = -0.13–1.55, $p = 0.10$). However, the effect size of the path from 3 cm HCCs to biochemical pregnancy was significant (B = 1.58, 95% CI = 0.18–2.98, $p = 0.03$). The direct effect of depression on biochemical pregnancy (c') was not significant (effect size = -3.15, 95% CI = -7.29–0.99, $p = 0.14$). The total indirect effect of depression on clinical pregnancy through 3 cm HCCs (c) was significant (effect size = 1.13, bootstrapped 95% CI = -0.14–3.87).

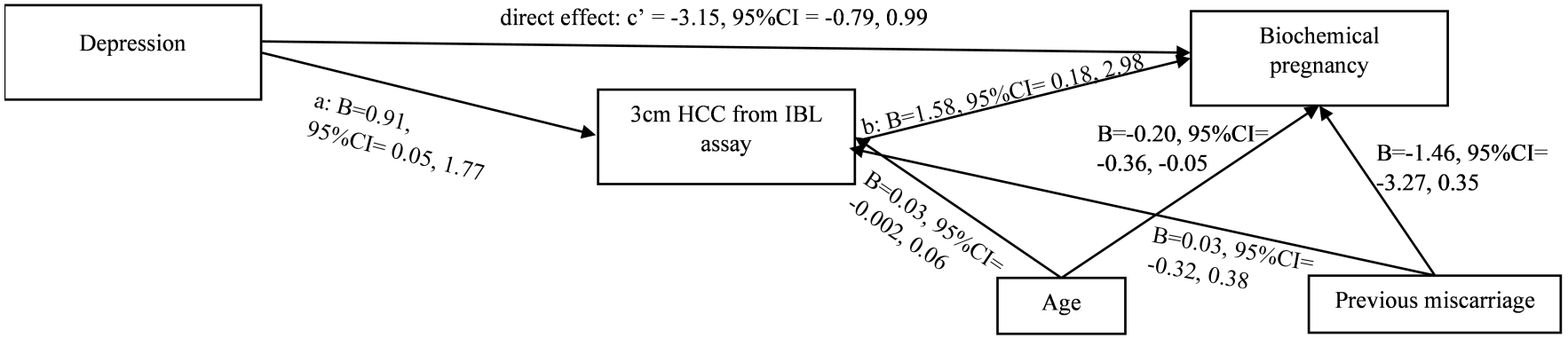


Figure 6.3 Path model 1, n = 62, showing the unstandardised regression coefficient for the association between depression and 3 cm hair cortisol concentrations (HCCs) from the IBL assay and the association between 3 cm HCCs from the IBL assay and biochemical pregnancy

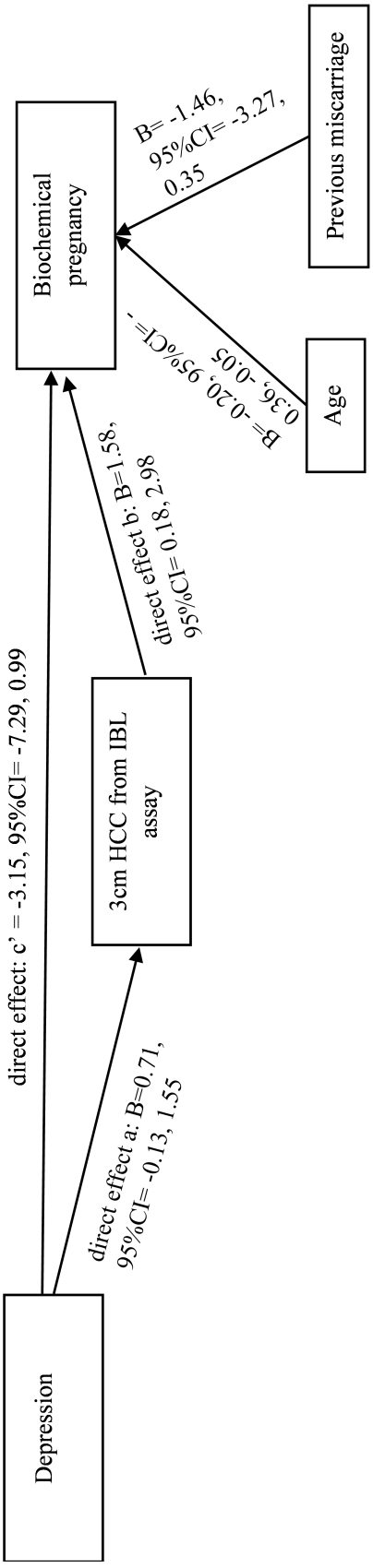


Figure 6.4 Path model 2, n = 62, showing the unstandardised regression coefficient for the association between depression and 3 cm hair cortisol concentrations (HCCs) from the IBL assay and the association between 3 cm HCCs from the IBL assay and biochemical pregnancy

Chapter 7

Discussion

The current study examined research questions relating to the associations between hair cortisol and clinical pregnancy, biochemical pregnancy, and discontinued treatment and the associations between hair cortisol and lifestyle and mood factors in women undergoing *in vitro* fertilisation (IVF) treatment. This chapter will discuss the findings reported in the previous chapter, the implications and limitations of the findings, the contributions and limitations of the current study, and recommendations for future research.

7.1 The association between hair cortisol concentrations and pregnancy outcomes

7.1.1 Hair cortisol and clinical pregnancy

The primary aim of this study was to examine the association between HCCs and clinical pregnancy following IVF treatment. The results showed that HCCs in the three months prior to commencing treatment were not significantly associated with the likelihood of achieving clinical pregnancy. The same results were found both before and after controlling for age and previous miscarriage. The association between 1 cm HCCs prior to commencing treatment and clinical pregnancy reached statistical significance for the IBL ELISA both before and after controlling for age and previous miscarriage. However, given the small sample sizes included in these analyses (n = 28 in the univariate analysis and n = 22 in the multivariate analysis), these results are potentially unreliable. In fact, because the 1 cm HCCs and 3 cm HCCs were strongly correlated, yet the 3 cm HCCs (with larger sample sizes) were not significantly correlated with clinical pregnancy results, it is likely that the significant associations between 1 cm HCCs and clinical pregnancy were due to a type 1 error. The large confidence intervals for these significant results also suggest that a type 1 error is the most likely explanation.

However, this finding contradicts earlier findings, e.g., Massey et al. (2016) reported a significant association between higher HCCs and a reduced likelihood of achieving clinical pregnancy among female patients undergoing IVF treatment. The null associations in the present work could have resulted from small sample sizes when examining the two assays separately (n = 75 for IBL and n = 86 for DRG). When examining the combined cortisol results from both assays, the considerable inter-assay variability may have reduced the accuracy of the combined

cortisol results, leading to the null associations. However, the null associations were supported by the sensitivity analyses using the MI datasets (n = 192–232). In summary, pre-treatment cortisol levels measured in hair may not be associated with the likelihood of clinical pregnancy following IVF treatment.

7.1.2 Is cortisol the pathway through which stress influences *in vitro* fertilisation treatment?

The null association between HCCs and clinical pregnancy could not simply lead to the conclusion that pre-treatment stress is not important in IVF treatment. There may be two possible ways to explain the null association. First, if pre-treatment chronic stress was associated with IVF treatment outcomes, the effect of stress on IVF treatment outcomes may happen mainly through other psychobiological pathways that were not measured in this study. This study focused on cortisol and, in so doing, on the HPA axis. But both the sympathetic and parasympathetic nervous systems also respond to emotional experiences and may also be relevant to IVF treatment outcomes through the secretion of other biomarkers that interact with female sex hormones (Smeenk et al., 2005; Zhou et al., 2019). For instance, the secretion of catecholamines (e.g., epinephrine and norepinephrine) may influence the secretion of GnRH, which regulates the secretion of other hormones involved in female reproduction (e.g., LH) (Schenker et al., 1992). Zhou et al. (2019) examined the levels of salivary alpha-amylase, the secretion of which is controlled by both the sympathetic and parasympathetic nervous systems, and clinical pregnancy outcomes in patients undergoing IVF treatment. They found that higher levels of salivary alpha-amylase in female partners, male partners, and couples were associated with greater risks of not achieving clinical pregnancy following treatment (Zhou et al., 2019). An et al. (2013) reported that women who became clinically pregnant had significantly lower levels of norepinephrine, which is heavily involved in the SAM pathway (see Chapter 2), on the

day of oocyte retrieval than those who did not. These study findings indicate that psychological stress is likely to influence IVF treatment outcomes through these other biological pathways. However, such studies mainly focused on acute measures (e.g., at one time point pre-treatment or during treatment) of the biomarker (e.g., salivary alpha-amylase). Therefore, the long-term association between these biomarkers of emotional stress and IVF treatment outcomes is still not well understood (An et al., 2013; Zhou et al., 2019). Indeed, although emotional distress is hypothesised to influence IVF treatment outcomes, the precise pathways by which this effect occurs remain unclear (Massey et al., 2016; C Santa-Cruz et al., 2020). While it is important to assess the role of the HPA axis, it may be appropriate for future research to consider a broader range of biomarkers to ascertain the involvement of all the primary pathways that connect emotional experiences to health (Facchinetti et al., 1997; Gallinelli et al., 2001; Smeenk et al., 2005; Wu et al., 2018). It is also possible that hair cortisol is limited in its ability to capture HPA axis dynamics (e.g., the ultradian activities of cortisol) thus higher levels of cortisol at crucial times (e.g., at oocyte retrieval) or different aspects of the diurnal rhythm are relevant (Russell et al., 2012). Future studies are also encouraged to measure these biomarkers at different time points over a longer period, which may provide more information about the long-term roles of these pathways in IVF treatment outcomes.

7.1.3 Is stress associated with pregnancy outcomes following *in vitro* fertilisation treatment?

A second possible explanation is that, as indicated by these data, pre-treatment chronic stress is not associated with treatment outcomes. Two previous systematic reviews provided evidence for this argument (see Chapter 2, Boivin et al. (2011); Nicoloro-SantaBarbara et al. (2018)). Boivin et al. (2011) suggested that pre-treatment emotional distress (i.e., anxiety and depression) was not signifi-

cantly associated with the likelihood of pregnancy following treatment. Nicoloro-SantaBarbara et al. (2018) suggested that pre-treatment stress, anxiety, and depression were not significantly associated with the likelihood of treatment success (i.e., achieving pregnancy or live birth). While these two systematic reviews examined evidence of self-reported stress provided by previous observational studies, the current study further provided evidence that the potential psychobiological pathway of cortisol was also not involved. However, while the evidence from these reviews may appear compelling, a causal relationship cannot be concluded from observational evidence.

In search of further answers relating to whether stress is associated with fertility treatment outcomes, evidence from trials must also be considered. The effect of psychosocial interventions on pregnancy rates in fertility treatments were examined in a previous systematic review in which the results of 35 randomised controlled trials, 10 non-randomised controlled trials, and eight uncontrolled trials were examined (Frederiksen et al., 2015). The meta-analysis based on data from 10 studies suggested that psychosocial interventions significantly increased the likelihood of achieving clinical pregnancy by 50%–100%. The meta-analysis also showed that psychosocial interventions significantly reduced levels of anxiety (23 studies, pooled effect size = 0.51) and depression (25 studies, pooled effect size = 0.73) in women undergoing fertility treatment. In addition, the meta-regression showed that larger reductions in anxiety were significantly associated with greater improvement in the likelihood of achieving clinical pregnancy. Although the types of interventions included in the review varied (e.g., CBT, mind–body intervention, counselling, support programmes, and stress management), the results suggested a robust evidence base for the significant positive effects of psychosocial interventions on stress reduction and the likelihood of achieving pregnancy in women undergoing fertility treatment (Frederiksen et al., 2015).

When compared with the findings from observational studies, evidence from trials

seems to disagree and support the beneficial effect of psychosocial interventions on treatment outcomes. According to stress and coping theory (see Chapter 2), the occurrence of stressors does not necessarily induce stress. Rather, an individual's cognitive appraisal of an event plays a critical role in stress. Whether an event is personally significant and exceeds the individual's resources to cope are both important in the appraisal of how stressful it is (Lazarus, 1984). Infertility and the procedure of fertility treatment are no doubt stressors that are personally significant to patients (Verhaak et al., 2007b). However, psychosocial interventions may be able to provide patients with practical tools (e.g., information regarding the treatment and relaxation techniques) and emotional support (e.g., counselling or through partners) to help them cope. Enriching the resources available for patients to help them cope with infertility and fertility treatments may therefore buffer the stress patients experience and any negative effects of stress on treatment outcomes (Nicoloro-SantaBarbara et al., 2018). According to Folkman (1997), there are four types of coping: positive reappraisal, goal-directed problem-focus coping, spiritual beliefs and practices, and the infusion of ordinary events with positive meaning, which all share the purpose of finding positive meaning in the current situation. These have been recognised as the theoretical frameworks in psychosocial interventions for patients undergoing fertility treatment. For example, Lancastle and Boivin (2008) conducted a feasibility trial in which patients received intervention cards to prompt them to think about positive aspects of their situation (i.e., positive reappraisal). When compared with the control group of patients who received cards with only positive mood phrases, patients who received positive reappraisal cards felt significantly more positive and were significantly more likely to recommend this intervention approach (Lancastle and Boivin, 2008). The later main trial further revealed that although the positive reappraisal card intervention did not significantly reduce anxiety and depression among patients undergoing IVF treatment and awaiting pregnancy results, it significantly improved patients' positive mood when compared with the control

group (Ockhuijsen et al., 2014). Chan et al. (2012) developed a mind–body–spirit intervention in which spiritual beliefs and practices were incorporated into the intervention to help patients develop spiritual strength against disruptions and the ability to grow in their situations. They showed that when compared with the non-intervention control group, patients in the intervention group had significantly lower anxiety after treatment (Chan et al., 2012). Unfortunately, they did not find a significant difference in pregnancy rates between the groups (Chan et al., 2012).

Other evidence also suggests positive effects of psychological interventions on treatment outcomes. For example, Domar et al. (2011) evaluated a mind–body intervention that focused on CBT, relaxation techniques, negative health behaviour modification, and social support for patients undergoing IVF over two treatment cycles. The results showed that the mind–body intervention did not significantly affect treatment outcomes for the first treatment cycle, but patients in the intervention group had significantly higher pregnancy rates in the second treatment cycle when compared with the non-intervention control group (Domar et al., 2011). They explored reasons for the non-significant intervention effects in the first cycle and significant effects in the second cycle. Because most of the participants were eager to start their first IVF cycle, the treatment cycles were started before all participants had completed the mind–body intervention programme. The data showed that more than half the participants (54%) had not attended any intervention sessions and only 9% had completed 6–10 sessions before starting their first cycle. However, only 5% of the participants had not attended any intervention session and 76% had attended 6–10 sessions before starting the second cycle. Domar et al. (2011) observed that patients in the intervention group started to report improvements in their psychological symptoms (e.g., anxiety and depression) from approximately midway through the intervention programme. Therefore, it is possible that the participants were not fully equipped with the skills provided by the intervention during their first cycle,

which contributed to the non-significant intervention effects (Domar et al., 2011). Together, the evidence is compelling that helping patients to cope and identify the positive meaning in their current situation is potentially beneficial not only to their emotional wellbeing but also to pregnancy outcomes (Frederiksen et al., 2015). Therefore, the value of such psychosocial interventions should not be dismissed, even though observational studies, including the current one, have failed to find a significant association between stress and pregnancy outcomes following fertility treatment.

The influence of psychological stress on IVF might not only occur before the treatment but also occur during the treatment. The current study focused on pre-treatment stress while the influence of chronic stress throughout the treatment on the treatment outcomes was not specifically investigated. However, previous literature has shown that the time around oocyte retrieval and pregnancy test is the most stressful part of a fertility treatment cycle (Verhaak et al., 2007a; Boivin and Takefman, 1995). Future studies are recommended to investigate the influence of during-treatment stress on various clinical outcomes (e.g., clinical pregnancy and live birth) for more empirical knowledge. The two-week waiting period (i.e., the period between embryo transfer and pregnancy test) has also been the primary focus of some psychological interventions (e.g., Lancastle and Boivin (2008); Ockhuijsen et al. (2014)) where positive reappraisal coping interventions were delivered to patients awaiting their pregnancy results to improve their mental health. Although these studies did not find a significant difference in pregnancy results between groups, patients in the intervention group reported significant improvement in positive mood. This evidence suggests that increasing positive mood among patients during these stressful periods within a treatment cycle might help patients to manage the demands of treatment. Future studies are also encouraged to further investigate the benefits of such interventions on several treatment outcomes such as clinical pregnancy and live birth.

As with the relationship between psychological stress and IVF treatment outcomes, the psychobiological pathways between cognitive appraisal and coping and IVF treatment outcomes should also be considered. The HPA axis may again be a target of interest for its critical role in stress. Unfortunately, possible psychobiological pathways have not been investigated intensively in interventional studies (Frederiksen et al., 2015). Therefore, it is unclear which biomarkers are involved in the effects of these psychosocial interventions on pregnancy outcomes. It is recommended that future studies collect biological samples (e.g., blood, saliva, urine, follicular fluid, and hair) to measure a range of biomarkers that may be important in the psychobiological pathways (e.g., cortisol, adrenaline, noradrenaline, and ACTH) along with a range of relevant psychological factors (e.g., appraisal and coping) throughout the intervention and treatment period. This approach would help uncover the prospective association between psychological factors (e.g., mood and coping), biomarkers (e.g., cortisol), and treatment outcomes. Such trials should both measure and clearly report (1) the psychosocial factors that the intervention targeted, (2) the biomarker that was measured and the hypothesised effect of the intervention on it, and (3) the hypothesised clinical outcomes. As pointed out in the systematic review in Chapter 4, the transparent reporting of these elements in trials is essential to determine whether an intervention was successful in (1) achieving its target of changing the psychosocial factor, (2) achieving its target of better clinical outcomes, and (3) assessing the mediating biological pathway reflected by the chosen biomarker.

7.1.4 Hair cortisol and biochemical pregnancy

A secondary aim of this study was to examine the associations between HCCs and biochemical pregnancy. HCCs from combined assays were not significantly associated with biochemical pregnancy outcomes in this study. For HCCs from the IBL assay, although a null association was found between 3 cm HCCs and biochemical

pregnancy in the univariate analyses, higher 3 cm HCCs were significantly associated with an increased likelihood of biochemical pregnancy after accounting for age and previous miscarriage. However, the same significant association was not found for the 1 cm HCCs from the IBL assay. No association was found between HCCs from the DRG assays and biochemical pregnancy outcomes. Sensitivity analyses using MI datasets also showed null associations between both the 3 cm HCCs and 1 cm HCCs and all three treatment outcomes. Together, these null findings suggest that a significant effect may again have been due to a type 1 error, as with the significant effect between 1 cm HCCs and clinical pregnancy, due to the small sample size (n = 75), which is likely because the confidence interval for the effect size of the IBL assay 3 cm HCCs on biochemical pregnancy was large (1.05–12.04).

C Santa-Cruz et al. (2020) conducted a similar study but did not find a significant association between pre-treatment 3 cm HCCs and biochemical pregnancy. Together with the results of the current study, this shows that cortisol may not be directly associated with the likelihood of achieving biochemical pregnancy. One possible explanation is that the biochemical pregnancy outcome in the current study was not markedly different from the clinical pregnancy outcome. Indeed, in the current study, only 10 participants (5%) who achieved biochemical pregnancy failed to achieve clinical pregnancy. Therefore, the majority of the cohort would fall into the same group (i.e., achieved pregnancy or not) for both the biochemical and clinical pregnancy outcomes. Therefore, statistical analyses would be expected to return similar results. However, theoretically, the effects of stress at implantation stage (i.e., biochemical pregnancy) are likely to carry through to later pregnancy (Boivin et al., 2011). Robust evidence has further suggested that the levels of stress in patients undergoing IVF treatment tend to increase during the treatment cycle and reach their highest levels at the clinical pregnancy test (Verhaak et al., 2007a). Specifically, an observational study in which women kept daily records of their stress throughout the whole treatment cycle

(i.e., from the first day of the cycle until after receiving the results of the clinical pregnancy test) found that the levels of stress did not change significantly from the first day until the end of the cycle just before the clinical pregnancy test, at which time, a significant rise in stress levels was observed (Boivin and Takefman, 1995). Therefore, although the current study did not find significant associations between pre-treatment cortisol levels and either biochemical pregnancy or clinical pregnancy outcomes, stress during treatment (especially shortly before the clinical pregnancy test) may influence pregnancy outcomes.

Evidence is currently unclear on whether post-implantation cortisol levels are associated with later pregnancy (Parker and Douglas, 2010). Some evidence has suggested that women who have elevated urinary cortisol levels during the first three weeks of conception are more likely to have a spontaneous abortion (Nepomnaschy et al., 2006). However, other evidence has suggested that elevated serum cortisol in the first trimester is not associated with later miscarriage (Parker and Douglas, 2016). Similarly, other data indicate that because natural killer cells can threaten foetal development, the suppressive effect of cortisol (i.e., inhibiting the activity of natural killer cells) may benefit foetal development (Arck, 2001). Therefore, the role of cortisol levels in early pregnancy in later pregnancy development is still very unclear. In addition, to date, no study has investigated post-implantation cortisol levels and clinical pregnancy or live birth outcomes in patients undergoing IVF treatment, which may be another suggestion for future work.

7.1.5 The association between hair cortisol and discontinued treatment

The third part of research aim one was to examine whether pre-treatment hair cortisol (1 cm and 3 cm) was associated with the likelihood of discontinuation

of treatment due to complications occurring at any stage of the treatment (e.g., OHSS or no oocytes collected). It was found that hair cortisol pre-treatment was not associated with the likelihood of complications that resulted in the treatment being discontinued (e.g., no oocytes retrieved, failed fertilisation, failed implantation, or loss of pregnancy after hCG pregnancy test).

It is possible that levels of cortisol in hair are lower than the levels that promote oocyte maturation and fertilisation and embryo implantation and development when compared with levels in urine and follicular fluid. For example, previous evidence has shown that increased levels of intrafollicular cortisol may be associated with fertilised oocytes and implantation of embryos in IVF (Keay et al., 2002) while lower levels of total intra-follicular cortisol may be associated with reduced fertilisation capacity of oocytes (Smith et al., 2002). This evidence suggests that cortisol measured in follicular fluid may be a more appropriate indicator of HPA axis activity in the context of fertility (Andersen, 2002). However, the correlation between hair cortisol and cortisol measured in follicular fluid was not mentioned in either of the two recent reviews of the methodology of using hair cortisol as a biomarker of stress (Kalliokoski et al., 2019; Greff et al., 2019). Therefore, the comparability of hair cortisol to follicular fluid cortisol has not been investigated. It may be helpful for future studies to collect both follicular fluid samples and, later hair samples to examine this issue. Such studies should note that follicular fluid samples may need to be collected at multiple time points to adjust for the circadian and ultradian rhythm of cortisol secretion (see Chapter 2), and hair samples need to be collected later in order to reflect cortisol levels during the same time frame as reflected by the follicular fluid samples.

It is also possible that combining several endpoints (e.g., failed fertilisation and loss of pregnancy after biochemical test) into a broader concept of 'discontinued treatment' reduced the effect sizes to null. Indeed, existing evidence on the role of cortisol in IVF treatment complications is limited. Massey et al. (2014) conducted

a review and found only a limited number of studies investigating the association between cortisol and oocyte number, fertilisation rate, cleavage, and miscarriage rates in IVF treatment, and they reported mixed findings. However, evidence is now emerging regarding cortisol and the risks of OHSS during IVF. For example, in a prospective study, Chistyakova et al. (2014) reported that women who experienced symptoms of OHSS had significantly higher cortisol levels than women who did not. However, the study did not specify when cortisol levels were measured or what type and index of cortisol were measured. Therefore, it is difficult to conclude whether it was higher cortisol levels pre-treatment or during treatment that were associated with risks of OHSS during IVF. In an observational study, Kopijka et al. (2017) reported that women who were at increased risk of OHSS had higher levels of serum cortisol levels immediately after stimulation of the ovaries during treatment than women who were not. However, since the study measured acute levels of cortisol immediately after stimulation, cortisol could only act as an indicator of OHSS rather than a risk factor. Unfortunately, the current study was unable to examine the association between pre-treatment hair cortisol and the risk of OHSS alone due to the very small sample size that experienced OHSS (n = 20). Ovarian hyperstimulation syndrome is related to elevated responses to the administration of gonadotrophin during treatment (Kopijka et al., 2017). Therefore, in theory, cortisol may be associated with the risk of OHSS because it can inhibit the secretion of GnRH from the hypothalamus and the secretion of gonadotrophins from the pituitary (Kopijka et al., 2017). Although there is some evidence suggesting that the administration of glucocorticoids does not reduce the rate of OHSS (Tan et al., 1992), a clear conclusion cannot be drawn due to the small sample size of this work (Delvigne and Rozenberg, 2002). It may be helpful to reveal whether pre-treatment levels of cortisol may be an indicator of whether patients are at risk of OHSS. The gathering of more evidence on whether variations in cortisol levels during ovarian stimulation are associated with the risk of OHSS could be another area of enquiry. To further investigate

these research questions, it is recommended that future prospective studies collect cortisol samples pre-treatment and/or during treatment. If the studies were to use acute measures of cortisol (e.g., blood, saliva, urine, or follicular fluid), it is recommended that samples be collected at different time points throughout the day for several consecutive days to account for the circadian and ultradian rhythm of cortisol secretion.

7.2 The direct and indirect association between lifestyle factors and *in vitro* fertilisation treatment outcomes

The first parts of aim two and aim three of this thesis were to examine the association between several lifestyle factors (i.e., sleep, physical activity, caffeine consumption, alcohol consumption, and smoking) and treatment outcomes (i.e., clinical pregnancy, biochemical pregnancy, and discontinued treatment), as well as the possible indirect associations through hair cortisol. These results will now be discussed.

7.2.1 Sleep

The first observation of the results was that in the cohort examined in this study, female patients undergoing IVF treatment had low sleep quality in general according to the cut-off of the scale (PQSI). Similarly, other studies have reported that 57%–90% of women undergoing IVF treatment reported poor sleep quality prior to the treatment (Philipsen et al., 2021; Goldstein et al., 2017) and a high prevalence of poor sleep quality during treatment (Philipsen et al., 2021; Goldstein et al., 2017). The evidence suggests that women undergoing fertility

treatment are likely to experience poor sleep quality prior to treatment, which is likely to persist through different stages of the treatment.

Therefore, these findings lead to the question of whether the poor sleep quality experienced by patients influences the likelihood of achieving favourable IVF treatment outcomes. No significant associations were found between sleep quality and clinical pregnancy, biochemical pregnancy, or discontinuation of treatment in this study. In addition, no association was found between sleep quality and either hair cortisol levels three months pre-treatment or hair cortisol levels one month pre-treatment among the patients undergoing IVF treatment in this study. These findings suggest that in this cohort, sleep quality was neither directly associated with IVF outcomes nor indirectly through cortisol. However, these findings are difficult to compare with existing evidence. For the association between sleep and treatment outcomes, there is considerable heterogeneity in how sleep was measured and how treatment endpoints were defined in other studies, with conflicting results reported (Caetano et al., 2021). For example, in a recent review, Caetano et al. (2021)found that evidence of an association between sleep and IVF pregnancy outcomes is highly limited, with only one study reporting an increased likelihood of achieving pregnancy among women who slept 7–8 hours per night when compared with those who slept 9–11 hours per night (Caetano et al., 2021).

For the association between sleep and cortisol, findings from the current study were also difficult to compare with previous work. In the systematic review reported in Chapter 4, no association was found between sleep and long-term levels of cortisol due to insufficient data. Indeed, evidence is more abundant regarding the effect of either self-reported sleep quality or sleep deprivation experiments on acute cortisol responses, while the association between sleep and hair cortisol is more limited (El Mlili et al., 2021). According to a recent narrative review paper, most of the available evidence to date on the association between sleep and hair cortisol has focused on older adults or shift workers but has produced

contradictory results (El Mlili et al., 2021). For example, in two similar studies, one showed a significant association between daytime sleeping and higher hair cortisol levels (Feller et al., 2014), and the other showed no association between daytime sleeping and hair cortisol levels in older adults (Lanfear et al., 2020). Therefore, it is difficult to conclude whether cortisol or the HPA axis is a pathway through which sleep quality influences health outcomes over the long term in healthy adults. In addition, no other study to date has examined the indirect association between sleep and IVF treatment outcomes through cortisol; therefore, it would be premature to determine the relationship between sleep, cortisol, and IVF treatment outcomes based on the current study alone.

Indeed, given the consistency of the data indicating poor sleep quality in women undergoing IVF treatment, it is important for future studies to examine the relationship further. Prospective studies should consider collecting information on multiple treatment endpoints, including fertilisation rate, clinical pregnancy, and live birth, which would help provide a more comprehensive picture of the whole treatment cycle. In addition, future studies could examine the association between sleep quality at different stages of the treatment and multiple treatment outcomes because, as suggested by other evidence (Philipsen et al., 2021; Goldstein et al., 2017), women's sleep quality may change throughout the treatment, and the effects of variations in sleep on treatment are unclear.

It is noteworthy that in this study, sleep quality was found to be significantly positively correlated with self-reported anxiety, depression, infertility-related negative mood, and non-specific negative mood, and was significantly positively correlated with positive mood ($p<.001$).

This finding is supported by other studies (Huang et al., 2019; Philipsen et al., 2021). It is possible that poor sleep quality contributes to emotional distress among IVF patients, or vice versa. In this case, answering questions relating to how poor sleep quality and emotional distress among patients are associated

would provide valuable information to help patients improve sleep or mood irrespective of the consequences for IVF treatment outcomes. Since observational evidence has suggested a possible association, interventional studies are encouraged to investigate whether improving sleep would contribute to better mood among patients or vice versa. Information from such trials would help healthcare providers to provide better guidance for patients and help patients to make lifestyle improvements for treatment (Gormack et al., 2015). It is also recommended that future studies explore the possible pathways for the influence of sleep on IVF treatment. Because of the significant association between poor sleep and poor mood, the HPA axis could again be a target of interest. Future studies, especially interventions, are recommended to measure cortisol levels in patients at baseline (i.e., before IVF treatment and intervention), during treatment/intervention, and after treatment/intervention to gain a more complete picture of whether and how cortisol changes in relation to IVF treatment and/or the intervention.

7.2.2 Physical activity

This study found that the studied cohort was engaged in moderate to high levels of physical activity, according to the cut-offs for low, moderate, and high levels on the scale (GPAQ). Similarly, Domar et al. (2015) observed a high proportion of women who engaged in regular exercise (46%–62%). Interestingly, Gormack et al. (2015) reported that 7% of the surveyed female patients reduced their levels of physical activity and 28% increased their levels of physical activity during treatment. However, another study found that women undergoing IVF treatment tended to engage in low to moderate levels of physical activity measured by surveys delivered via mobile phones (Cesta et al., 2018). Together, this evidence does not suggest a clear pattern of the engagement in physical activity among women undergoing IVF treatment. However, it is possible that women who perceive the

positive influences of physical activity on treatment outcomes (e.g., higher likelihood of achieving pregnancy) are more likely to adopt these lifestyle changes. In fact, a previous study among 138 couples undergoing IVF treatment found that 76% of patients believed that it would be helpful for women to exercise, and 62% of patients believed that exercise for more than 4 hours a week would be helpful for achieving favourable treatment outcomes (Rossi et al., 2016).

However, the present study did not find any significant direct associations between the amount of physical activity and clinical pregnancy, biochemical pregnancy, or discontinued treatment. A null association between total physical activity (minutes/month) and hair cortisol was also found. These results suggest that in the current cohort, levels of physical activity were neither directly associated with treatment outcomes nor indirectly through hair cortisol.

These findings contradict previous literature. For example, in a previous systematic review of seven studies, Rao et al. (2018) found that women who regularly engaged in physical activity were more likely to achieve clinical pregnancy when compared with those who did not. However, in these seven studies, four studies measured regular exercise instead of physical activity. As was explained in Chapter 3, exercise is a sub-concept of physical activity which involves physical activities that are planned, structured, repetitive, and aim to improve or maintain physical fitness (Caspersen et al., 1985). Of the other three studies, only one measured total physical activity among women using the same tool as was used in this study (i.e., GPAQ) with a larger sample size (n = 216) and participants grouped into regular or not regular physical activity (Palomba et al., 2014). However, in the current study, instead of grouping the participants in two groups, the overall energy expenditure for each respondent was calculated to represent their total physical activity, which was then used in all subsequent analyses. The heterogeneity in physical activity measures used across studies made it difficult to compare the findings of this study with others. Together, these findings suggest

that the current evidence is unclear regarding how physical activity is associated with IVF treatment outcomes. Because a proportion of IVF patients tend to believe that physical activity is beneficial to their treatment (Rossi et al., 2016), it is important for future studies to examine how physical activity may deliver the hypothesised benefits. A few research questions could be considered: (1) Is physical activity associated with favourable IVF treatment outcomes? If so, at what levels of physical activity? (2) Are changes in the level of physical activity before and during treatment associated with favourable IVF treatment outcomes? Observational studies in which the habitual levels of physical activity are measured may be able to address the first question, and interventional studies that examine the effect of physical activity interventions on IVF treatment outcomes may be able to address the second question. Answers to these questions would provide more information to help patients decide whether to modify their level of physical activity, and if so, to what extent.

The current cohort study did not find a significant association between physical activity and HCC. This means that the hypothesised indirect effect of physical activity on IVF treatment outcomes through hair cortisol was not confirmed. Similarly, Cesta et al. (2018) found no association between levels of physical activity and the likelihood of not progressing in oocyte aspiration or embryo transfer or achieving clinical pregnancy, and physical activity was also not found to be associated with cortisol levels. However, only morning and evening cortisol levels were measured in the saliva on one day at around the start of the treatment (Cesta et al., 2018). Due to differences in measurements of both cortisol and physical activity, no comparison between the findings in the current study and the aforementioned findings (Cesta et al., 2018) can be made. To date, no other study has investigated the possible mediating role of hair cortisol in the association between physical activity and treatment outcomes in patients undergoing IVF treatment. However, based on other evidence, it is possible for physical activity to influence treatment outcomes indirectly through cortisol over the long term.

In fact, the association between physical activity and long-term levels of cortisol was confirmed in the systematic review presented in Chapter 4, with a significant pooled correlation of 0.16 ($p<0.001$). Future studies can further examine this pathway and update the evidence base by (1) using robust measures of the total amount of physical activity that patients undergoing fertility treatment engage in to draw a clearer pattern of how frequently they engage in physical activity and to what level and (2) using long-term measures (e.g., hair cortisol) to capture long-term cortisol levels, which may mediate the association between physical activity and treatment outcomes over the long term.

Another mechanism by which physical activity may influence IVF treatment is through weight management (Watson, 2015). As described in Chapter 3, it is well documented that BMI or bodyweight are significant predictors of treatment success, with higher BMI reducing the likelihood of achieving clinical pregnancy and live birth and increasing the likelihood of miscarriage (Rittenberg et al., 2011). However, a significant association between BMI and treatment outcomes was not found in this study. As BMI collectively reflects other lifestyle factors, such as diet and physical activity (Farhud, 2015), the moderate to high levels of physical activity that women engaged in may have contributed to the overall healthy BMI observed in the current study. This was supported by the evidence base that regular physical activity is an important modifiable component of maintaining a healthy BMI, which may then benefit IVF treatment outcomes (Harrison et al., 2016; Rittenberg et al., 2011).

7.2.3 Caffeine consumption

For caffeine consumption, the results in this study showed that the average caffeine consumption, based on the calculation of habitual consumption of caffeine-containing food and beverages, was 2,329.7 mg/month (approximately 78 mg/-day), which is lower than the 300 mg/day recommended by the UK government's

Food Standards Agency (CARE Study Group and others, 2008). A previous prospective study of 12,800 patients undergoing IVF treatment in the US reported that a high proportion of women undergoing IVF consumed caffeine (62%–68%). Another observational study in New Zealand also reported that 87% of women undergoing IVF consumed caffeinated beverages (Gormack et al., 2015). The evidence and findings from the current study suggest that although women are likely to consume caffeine during treatment, the levels of their caffeine consumption may be low. Interestingly, Gormack et al. (2015) reported that approximately 40% of women abstained from or reduced caffeine consumption during treatment. It is possible that women undergoing IVF believe that caffeine may negatively influence their treatment and therefore restrict their habitual consumption of caffeine (Rossi et al., 2016).

Based on these results, this question should be considered: Is caffeine consumption associated with the likelihood of achieving favourable IVF treatment outcomes? Results of this study also showed that caffeine consumption in the month prior to IVF treatment was not associated with clinical pregnancy, biochemical pregnancy, or discontinuation of treatment. Similarly, previous systematic reviews and recent observational studies did not find significant associations between caffeine consumption and IVF treatment outcomes, including ovarian stimulation, fertilisation, embryo quality, implantation rate, clinical pregnancy, and live birth (Lyngsø et al., 2017; Abadia et al., 2017). The current study further found null associations between caffeine consumption and hair cortisol either three months pre-treatment or one month pre-treatment. Together, these results suggest that pre-treatment caffeine consumption is not associated with either cortisol or IVF treatment outcomes, with the latter findings potentially offering reassurance to patients given the consistency of this finding across multiple studies.

However, it needs to be noted that in the current study, a diary-like self-report measure of caffeine consumption was used. Participants were asked to retro-

spectively recall their consumption of caffeinated beverages and food over the preceding four weeks. This method could lead to errors in quantifying the exact amount of caffeine in beverages and food. Although in this study, the caffeine measure included most of the beverages and food that contain caffeine (e.g., coffee, tea, soft drink, chocolate, and chocolate flavoured desserts), it is still possible for some caffeinated beverages or food not to have been captured by this measure. Error is also likely to occur in women's self-reported caffeine consumption in relation to the brand, serving size, frequency, etc. due to misreporting (Poslusna et al., 2009). In a previous review, Poslusna et al. (2009) addressed misreporting in dietary assessment and found that in studies using 24-hour recall or weighted food records, 30% of participants were likely to underreport their food intake in 24-hour recalls, and the energy intake was likely to be underestimated by 15% in such studies (Poslusna et al., 2009). In the current study, the recall time frame was four weeks, which may have contributed to greater errors. Although the information about misreporting was limited to energy intake in this review, owing to the limited number of studies available, as caffeine consumption is a comparable human behaviour, it may be expected that self-report measures of caffeine consumption suffer from similar levels and sources of misreporting.

Another potential source of bias is that the quantity of dietary caffeine is difficult to estimate accurately due to limited information on how much caffeine each beverage and food contains. The estimated caffeine contents in participants' diets were calculated using their reported brand and portions and the caffeine contents in popular beverages and food at the Caffeine Informer website (www.caffeineinformer.com). However, if a brand was not given by the participant, an estimated amount was assigned, but caffeine contents vary across different foods and beverages (Chou and Bell, 2007). For example, a cup of brewed coffee may contain 64–272 mg caffeine, depending on the brewing technique, ground coffee used, variations in the coffee blend, etc. (www.caffeineinformer.com). Therefore, errors may occur in these estimates. To reduce the effect of mis-

reporting, future studies using food diaries for measuring caffeine consumption should (1) reduce the recall time frame (e.g., half a day or 24 hours) and (2) capture the self-reported habitual caffeine consumption frequency and amount for a participant over a period of time longer than one day (e.g., one week to several weeks).

It is also possible that high levels of caffeine consumption are associated with IVF outcomes but were undetected in the current study. Women who consume more than 300 mg/day (approximately 9,000 mg/month) were underrepresented in this study (n = 9, 4%). Therefore, the conclusions relating to whether high levels of caffeine consumption (i.e., more than 300 mg/day) would influence treatment outcomes may have been influenced by a floor effect.

In a previous study in which a significant and positive association was observed between caffeine consumption and aborted pregnancy following IVF treatment, the average caffeine consumption among participants was 455.82 mg/day (range: 3.71–3,561 mg/day) (Al-Saleh et al., 2010). This is higher than the average caffeine consumption among women in the current study (approximately 78 mg/-day). The average caffeine consumption in the current study was also lower than the average caffeine consumption among pregnant women in another cohort study (159 mg/day) (CARE Study Group and others, 2008). The study included in the review in Chapter 4 also suggested that caffeine intake of 300 mg/day after a period of caffeine abstinence was associated with elevated cortisol responses, which only lasted for six hours (Lovallo et al., 2005). These study findings hint that there may be a threshold for caffeine to influence both IVF treatment outcomes and cortisol. However, evidence to support this assumption is lacking mainly due to difficulties associated with measuring habitual caffeine consumption from all possible sources, as mentioned above.

A previous study attempted to measure both self-reported caffeine consumption and caffeine clearance measured in saliva using a caffeine challenge (CARE Study

Group and others, 2008). This suggests that using a food diary along with biological samples may be able to better capture the levels of habitual caffeine consumption among participants. However, acute samples, such as blood and saliva, could only capture the half-life of caffeine that has been consumed within hours, while long-term measures, such as hair samples, may be more insightful for investigating long-term caffeine consumption (Jiang et al., 2022). Future studies could consider measuring habitual caffeine consumption via both a food diary and hair samples over the same time frame to find out (1) how well they are correlated and (2) the accuracy of a food diary in reflecting total caffeine consumption. Future studies may also investigate the association between caffeine levels and cortisol measured in hair, which would provide valuable information on whether and how caffeine may influence health outcomes through cortisol.

In summary, current evidence does not support an association between caffeine consumption and IVF treatment outcomes. It is also still unclear whether cortisol levels measured in hair are involved in this association, if it exists. Given the limitations in the evidence, it is difficult to conclude whether the null associations observed in this study are truly due to a lack of effect or the limitations of the study itself. Current evidence and findings from this study do not support the assumption that women undergoing fertility treatment tend to reduce or limit their caffeine consumption. However, whether limiting caffeine consumption would truly benefit the treatment and through which mechanisms also requires further investigation.

7.2.4 Alcohol consumption

This study observed that women undergoing IVF treatment generally had low levels of alcohol consumption, with an average score of three on the AUDIT (with a score >8 indicating risky or hazardous levels of drinking behaviours). Previous observational studies have reported inconsistent proportions of women who

consumed alcohol during treatment (17%–23%, Domar et al. (2015) and 5.8%, Gormack et al. (2015)). However, the current study suggested that although women did engage in alcohol consumption during treatment, they were likely to consume at a low level. Similarly, Gormack et al. (2015) reported that 45% of women abstained from or reduced their alcohol consumption during treatment. A previous study reported that at least 69% of women undergoing fertility treatment believed that avoiding alcohol would be beneficial to the treatment Rossi et al. (2016); therefore, the low levels of alcohol consumption in this population were not surprising.

In terms of alcohol consumption and IVF treatment, this study did not find significant direct associations between alcohol consumption and clinical pregnancy, biochemical pregnancy, or discontinuation of treatment. This study also did not find a significant association between alcohol consumption and hair cortisol levels, which indicated that the indirect association between alcohol consumption and IVF treatment outcomes through cortisol may be null. These findings agree with findings from a previous study in which no significant association was found between alcohol consumption in female patients and number of oocytes, fertilisation rates, biochemical pregnancy, or loss of pregnancy both before and after adjusting for other covariates, such as age and BMI (Firns et al., 2015). However, the findings of this study contradict findings from other studies reporting significant associations between increased alcohol consumption and a reduced likelihood of achieving pregnancy (Klonoff-Cohen et al., 2003) and increased risks of fertilisation failure, implantation failure, and loss of pregnancy in IVF treatment (Rossi et al., 2011; Klonoff-Cohen et al., 2003). The sample sizes of the current study (n = 84–178 in the logistic regressions) were smaller than the two studies that reported significant associations, (n = 221, Klonoff-Cohen et al. (2003) and n = 2,545, Rossi et al. (2011)) but similar to the other study that reported non-significant findings (n = 68, Firns et al. (2015)). Indeed, it is possible that because women consumed relatively low levels of alcohol or tended to reduce their levels

of alcohol consumption before and during treatment, a larger sample with greater variability in alcohol consumption is required for a study to have the statistical power to detect an effect.

It is clear that the current evidence base is insufficient to draw a definite conclusion. The systematic review in Chapter 4 only found a handful of relevant studies reporting mixed results: One observational study reported no association between alcohol consumption and hair cortisol over six months in the general population, and two interventional studies reported blunted serum cortisol responses in alcoholic participants. It is fair to assume that the levels of alcohol consumption among women undergoing IVF treatment would be similar to or lower than the general population, rather than being similar to the alcoholic population; therefore, the association is likely to be null. This suggests that as with caffeine consumption, there may be a threshold for alcohol consumption to affect hormones like cortisol and other health outcomes, such as pregnancy, in IVF treatment. There is some evidence for this assumption. For example, Stalder et al. (2010) compared hair cortisol levels among alcoholic and non-alcoholic participants and found that hair cortisol levels that reflected the period when alcoholic participants were drinking were three-fold higher than the cortisol levels of participants who consumed no or only a modest amount of alcohol (Stalder et al., 2010).

Taking all the evidence together, there are several implications that need to be considered. Although the evidence relating to whether and how alcohol consumption is associated with IVF treatment outcomes is unclear, it should still be recommended to patients who are trying to conceive and those who are undergoing IVF treatment to minimise alcohol consumption where possible due to its proven negative effects on health in general (Burton and Sheron, 2018) and female reproductive functions (Emanuele et al., 2002).

7.2.5 Smoking

Only 2% of the women in the cohort reported smoking, which is consistent with findings from other studies (2%–7%, Domar et al. (2015) and 2%, Gormack et al. (2015)). As described in Chapter 3, evidence has shown that cigarette smoking can negatively affect reproductive health, e.g., causing delayed conception, increasing the risk of miscarriage, and reducing the likelihood of achieving favourable outcomes following fertility treatment (Wallach et al., 1996). Therefore, it is unsurprising that a low percentage of women reported smoking before and during treatment in different studies. In fact, a previous study reported that nearly 72% of women undergoing fertility treatment believed that it would be helpful for women not to smoke and 86% believed that second-hand smoke exposure was harmful to IVF treatment (Rossi et al., 2016). These findings again suggested that women undergoing IVF treatment were likely to maintain a relatively healthy lifestyle or to improve their lifestyle with the hope that these lifestyle choices would benefit their treatment (Rossi et al., 2016).

Due to the low number of women who reported smoking, the direct association between smoking behaviours and IVF treatment outcomes and the indirect association through cortisol could not be examined in the current study. However, other evidence exists in support of the negative impact of smoking on IVF treatment outcomes. For example, several systematic reviews have consistently shown that smoking is a significant risk factor for delays in natural conception (Augood et al., 1998), miscarriage and ectopic pregnancy in IVF treatment, and the reduced likelihood of IVF treatment success (e.g., clinical pregnancy and live birth) (Waylen et al., 2008; Klonoff-Cohen, 2005). However, whether cortisol mediates the effect of smoking on IVF treatment outcomes is less clear. In the systematic review reported in Chapter 4, contradictory evidence was found from observational and interventional studies: Three observational studies reported no association between habitual smoking behaviours and hair cortisol, and one inter-

ventional study reported reduced hair cortisol among participants of a smoking cessation intervention. This evidence suggests that smoking behaviours are perhaps unlikely to affect hair cortisol, but smoking cessation may have a beneficial effect on reducing hair cortisol. A previous study also suggested that smoking is likely to have a short-term effect on the HPA axis (Badrick et al., 2007). This evidence may mean that although smoking is likely to affect acute levels of cortisol, habitual smoking is less likely to affect cortisol levels over the long term. Therefore, cortisol may be unlikely to be the main pathway by which smoking influences IVF treatment outcomes. Future research may benefit from considering other possible pathways, for instance, oxidative stress and DNA damage (Konstantinidou et al., 2021; Al-Saleh et al., 2010).

7.3 Mood among women undergoing assisted reproductive treatment and treatment outcomes

In Chapter 2, it was discussed how psychological stress may influence fertility and IVF treatment outcomes and the inconsistency in the evidence regarding this association was highlighted. Therefore, as part of aims two and three, this study also investigated the direct association between several mood factors as affective responses to stressors (i.e., anxiety, infertility-related negative mood, non-specific negative mood, depression, and positive mood) and treatment outcomes (i.e., clinical pregnancy, biochemical pregnancy, and discontinued treatment). The associations between these mood factors and pre-treatment hair cortisol levels were also examined. The findings are discussed below.

7.3.1 Direct association between mood and treatment outcomes

No significant associations were found between any of the mood measures (i.e., depression, anxiety, infertility-related negative mood, non-specific negative mood, and positive mood) and any of the treatment outcomes (i.e., clinical pregnancy, biochemical pregnancy, and discontinued treatment). These results contradict previous findings. For example, in a review, Matthiesen et al. (2011) found significant pooled associations between stress and anxiety and the likelihood of achieving clinical pregnancy. However, the findings of the present study echo findings from a previous systematic review in which Boivin et al. (2011) found no significant pooled association between pre-treatment psychological distress (i.e., depression and anxiety) and the likelihood of achieving pregnancy (i.e., biochemical or clinical pregnancy). Similar null associations were also reported in more recent cohort studies (An et al., 2013; Cesta et al., 2018; Cheung et al., 2019; Miller et al., 2019). The observational evidence regarding the relationship between mood and IVF treatment outcomes remains unclear. However, as explained previously, evidence from trials suggests that psychological interventions that improved mood significantly benefited treatment outcomes, such as the clinical pregnancy rate (Frederiksen et al., 2015). This seemingly contradictory evidence suggests that there may be other factors that mediate the association between mood and IVF treatment outcomes, such as coping and resilience, which were not necessarily examined in the observational studies. Future studies, especially observational studies, should consider examining the role of these factors (e.g., coping) in the association between mood and treatment outcomes using a mediation analysis.

7.3.2 Indirect association between mood and treatment outcomes

It is also possible that the effect of mood on treatment outcomes is not direct but mediated by a psychobiological pathway. The HPA axis, and specifically cortisol, was measured here. However, results from univariate regression analyses showed that apart from pre-treatment depression, which was significantly associated with 3 cm HCCs from the IBL assay, none of the other mood factors were significantly associated with 3 cm HCCs from either assay. None of the mood factors were associated with 1 cm HCCs from either of the assays. The significant association between higher levels of depression and higher 3 cm HCCs may suggest that cortisol is a robust biomarker of depression. This is supported by evidence in which a significant positive association between depression and long-term levels of cortisol was found (see Chapter 4). However, this significant association cannot lead to the conclusion that stress is associated with HCCs. Although often correlated, different mood factors (e.g., depression and anxiety) are qualitatively different and reflect different aspects of stress (Lazarus, 1993). For example, anxiety often reflects the emotional response to uncertainty and perceived threat (Lazarus, 1993), and depression often overwhelmingly consists of sadness (Lazarus, 1991). The nature and sources of these mood experiences may be important for the relationship between stress and cortisol (Staufenbiel et al., 2013). For example, the literature indicated that hair cortisol seemed to increase in patients with depression but decrease in patients with anxiety (Staufenbiel et al., 2013). In addition, the current study did not find associations between any other mood measure and cortisol. However, the current study did not measure perceived stress among patients, which has been found to be significantly associated with long-term levels of cortisol (see Chapter 4). Based on this evidence, a few recommendations for future studies can be made. First, when examining stress among patients undergoing IVF treatment, researchers should be mindful

of and try to capture the different aspects of emotional responses (e.g., perceived stress, anxiety, and depression). Second, while examining the association between mood and biomarkers, such as cortisol, the different aspects of mood should also be considered.

A path analysis was conducted to further examine whether high levels of depression were indirectly associated with pregnancy outcomes through hair cortisol. Two different models were built based on the data and the original thesis model. In the first model (suggested by data), age and previous miscarriage were covariates of both 3 cm HCCs and biochemical pregnancy (see Figure 6.2). This hypothesis was based on previous evidence that age may be associated with HCC and be a critical predictor of fertility treatment success (Wosu et al., 2013; van Loendersloot et al., 2010), and the experience of previous miscarriage may be associated with emotional distress (Qu et al., 2017) and IVF treatment outcomes (Bates Jr and Ginsburg, 2002; Kupka et al., 2004). In the second model (sensitivity analysis based on the original thesis model), age and previous miscarriage were the only covariates considered for biochemical pregnancy (see Figure 6.4). The biochemical pregnancy outcome was used in these models for bigger sample sizes and based on the possible association between these outcomes and HCCs that was found earlier.

In the first model, the overall indirect effect of depression on biochemical pregnancy through 3 cm HCCs was significant, with a small effect size of 1.44 after controlling for age and previous miscarriage. Both the path from depression to 3 cm HCCs (effect size = 0.91) and the path from 3 cm HCCs to biochemical pregnancy (effect size = 1.58) were small but significant. In this model, the direct path from depression to biochemical pregnancy was not significant, which indicated that women who reported a one-unit higher level of depression were estimated to have higher HCCs, which was then associated with an increased likelihood of achieving biochemical pregnancy following treatment. Although

this result suggested a significant indirect effect, it contradicts the literature and the hypothesis that depression would be negatively associated with pregnancy outcomes. The relatively small sample size in the path analysis (n = 62) suggests that caution is warranted in interpreting these results, as they may be due to a type I error. Although the bootstrapping method used in the path analysis is generally accurate for type I errors, bias is still likely to occur when sample sizes or effect sizes are small (Williams and MacKinnon, 2008). This assumption was supported by the sensitivity analysis. In the second model (sensitivity analysis), the overall indirect effect of depression on biochemical pregnancy through 3 cm HCCs was non-significant (effect size = 1.13) after controlling for age and previous miscarriage. The direct path from depression to 3 cm HCCs (effect size = 0.71) was also not significant, but the direct path from 3 cm HCCs to biochemical pregnancy (effect size = 1.58) reached a significant level. These results indicated that a null indirect effect from depression to biochemical pregnancy through 3 cm HCCs may exist.

From a methodological point of view, it is possible that because levels of depression were measured over the preceding four weeks and the 3 cm HCCs reflect approximately three months preceding the hair sample collection in the current study, they were not well correlated in time in the first place. As was pointed out in the systematic review, imprecision can be introduced when mood and cortisol are measured over different time frames. This means that the significant association between depression and 3 cm HCCs from the IBL assay could also be due to a type 1 error. This appears to be more likely if we consider the association between depression measured over the past four weeks and 1 cm HCCs, which reflects cortisol levels over the same time frame. No significant associations were found between these measures. Therefore, while not consistent with previous work, it appears most prudent to conclude that the relationship between depression and cortisol in the current study should be interpreted with caution.

7.4 High levels of depression among women undergoing assisted reproductive treatment

In the current study, a high proportion of women reported high levels of depression prior to treatment: 70% of the whole cohort reported levels of depression higher than the cut-off for clinical depression (mean total score = 18.8, *SD* = 7.0, cut-off score = 16). This result is supported by findings of another study using the same measure of depression (i.e., CED-S) in which 80% of women undergoing IVF treatment had CES-D scores >16 (Aimagambetova et al., 2020). These similar findings suggest that it is common for women to report relatively high levels of depression at the start of treatment, which further indicates the benefits of psychological screening before starting fertility treatments to identify those who may be at risk of experiencing further emotional distress (Verhaak et al., 2010). Although no significant association between depression and IVF treatment outcomes was found in the current study, robust evidence has suggested that the psychological burden that patients experience during treatment (e.g., anxiety and depression) is an important factor that contributes to patients' decision to discontinue treatment at all stages (Dawson et al., 2005; Gameiro et al., 2012). This again highlights the importance of (1) psychological screening to identify who might be at risk of experiencing psychological distress during treatment and (2) support (e.g., psychosocial intervention) to help patients to cope with their emotional distress to enhance both the compliance with treatment and patients' experience of the treatment (Gameiro et al., 2012; Boivin et al., 2011).

7.4.1 Differences in hair cortisol values from the two assays

A further consideration while interpreting the results of this study is the differences in the HCC values from the two assays. It was necessary for HCC values in this study to be determined by two types of ELISA kit (i.e., IBL and DRG) due to assay errors (see Chapter 5). The results showed that although the HCCs from the IBL ELISA and DRG ELISA were significantly correlated, the IBL assay results were significantly higher than the DRG assay results for both the 3 cm HCCs and 1 cm HCCs. This finding is likely due to the variability in the steroid cross-reactivity of the immunoassay manufacturers. Indeed, although immunoassays have been widely used for analysing cortisol concentrations in hair due to their high sensitivity and relatively easy procedure, the inter-assay variability among commercially available immunoassays can be significant (Russell et al., 2015). In an earlier study, Russell et al. (2015) compared methods adopted by various laboratories and assessed the consistency of determining HCCs using different assay methods. In the study, four different laboratories analysed hair cortisol from a common pool of hair samples using four different immunoassays including Alpco ELISA, IBL Luminescence immunoassay, DRG ELISA, and Salimetrics® ELISA and followed relatively similar laboratory protocols. They also compared these results to HCCs analysed by liquid chromatograph-mass spectrometry (LC-MS/MS), which is considered to be the gold standard with excellent sensitivity and specificity (Gao et al., 2016). Russell et al. (2015) found that the absolute values of HCC determined by each laboratory differed significantly, which was likely due to the variability from the immunoassay manufacturers. This finding suggested that using two different immunoassays from two manufacturers, as in the current study, is likely to result in significantly different HCC results by immunoassay. The differences in HCC results by immunoassay may further contribute to imprecision of the results in the current study.

7.4. High levels of depression among women undergoing assisted reproductive treatment

A further disadvantage of immunoassays is that they are likely to have low specificity for the antibody, which may result in lower cortisol concentrations and overestimates of other hormones in the hair, other than cortisol itself. Previous literature in serum oestradiol and progesterone immunoassays showed that the functional sensitivity (i.e., the lowest concentration of serum oestradiol or progesterone that can be measured with a run-to-run imprecision of 20%) was two- to four-fold higher than the detection limits that were defined by the manufacturer of the tests (Taieb et al., 2002). Similarly, Moal et al. (2007) assessed serum testosterone concentrations using five different immunoassays and LC-MS/MS and found that the serum testosterone concentrations obtained by the immunoassays were higher than those obtained by LC-MS/MS. Kirschbaum et al. (2009) found that the IBL immunoassay significantly overestimated cortisol concentrations by 15% in the hair samples (n = 195) when compared with LC-MS/MS values. In another study, HCCs measured by immunoassays were also invariably higher than LC-MS/MS values (Russell et al., 2015). It is hypothesised that the substances in hair cross-react with other steroids, lipids, and parts of the hair matrix (Kirschbaum et al., 2009; Gao et al., 2016). It is possible that the immunoassays report cortisol-like immunoreactivity while all steroids may have been washed out or degraded after sample preparations. This evidence suggests that it is possible that HCCs reported in the current study were over-estimated. The range of HCCs reported in the current study was large for both assays (3 cm IBL: 28.4–235.4 and 3 cm GRD: 5.4–366.1 pg/ml). Evidence suggests that apart from low specificity, immunoassays are also likely to suffer from poor method optimisation over a large concentration range (Taieb et al., 2002; Moal et al., 2007).

These methodological limitations contributed to difficulties in comparing HCCs from the two different ELISA kits and possible problems with further analyses that combined HCCs from both ELISA kits in this study. Although this study attempted to use Z-score transformed HCCs to standardise and combine val-

ues from the two ELISA kits, results from these analyses should be interpreted with caution. The variability in the immunoassays limits the interpretation of the results from different assays (Russell et al., 2015). One way to tackle these methodological caveats is to include a reference group in research attempting to address the effects of stress on cortisol or other hormones (Russell et al., 2015). Indeed, the establishment of an international benchmark is critical in determining a universal low-to-high cut-off for HCCs. Russell et al. (2015) suggested that laboratories use a correction factor for each immunoassay brand to transfer the original HCCs to the LC-MS/MS equivalent. However, to date, the amount of reference material for determining an international benchmark for HCCs is still insufficient (Greff et al., 2019). Future work should invest in the establishment of an international benchmark in order to move HCCs toward clinical utilisation. Given that previous studies have shown significant associations between hair cortisol levels and the likelihood of achieving pregnancy following IVF treatment but with different directions of the association (Massey et al., 2016; C Santa-Cruz et al., 2020), future studies are recommended to further investigate the ability to use pre-treatment hair cortisol to predict treatment outcomes, using more consistent assay procedures or analytical procedures that yield better sensitivity and specificity, such as LC-MS/MS.

7.5 Study limitations

There are several limitations that must be considered. First, the observational nature of the study means no causal relationships can be concluded. Even with the path analysis as a statistical tool, the data only allowed this study to examine the associations rather than causality between mood, cortisol, and IVF treatment outcomes. Future interventions among patients undergoing IVF treatment in which both mood and cortisol are measured may help address this issue. Second, the hair samples were assayed using two ELISA kits from different man-

ufacturers, which resulted in different ranges of cortisol values. Although the Z-score transformation was applied to hair cortisol values, the imprecision introduced by the two ELISA kits may have affected our results. Given the limitations in immunoassays, it is recommended that future studies measuring hair cortisol use LC-MS/MS, which is considered the "gold-standard" method for determining cortisol levels in hair. Future studies may also consider determining cortisone levels in hair, which are highly correlated with levels of free cortisol and less affected by other contaminants, such as cortisol-containing creams (Wang et al., 2019). Third, self-report measures were used to measure mood and lifestyle, which asked participants to recall their mood and lifestyle over the preceding month. Although these measures were validated, recall errors may occur. Future studies may consider using self-reported measures over a shorter recall window or in combination with other physical measures (e.g., activity monitors) where possible. However, research should also consider the duration and convenience of these measures for participants who may already be experiencing stress from the treatment. Another limitation in relation to the generalisability of the study findings is that all patients were recruited from a single fertility clinic in the United Kingdom, with a high proportion of White British women (85%); therefore, the findings may not be generalisable to other populations or in other countries.

7.6 Alternative designs

If a future study was to be conducted following this work, several study designs should be considered. Because the current study found that women undergoing IVF treatment were likely to report poor sleep and high levels of depression, an intervention to improve sleep quality and depressive symptoms among patients undergoing IVF should be considered first. For example, a cognitive behavioural intervention could be considered that challenges disadvantageous beliefs relating to sleep and insomnia, changes attitudes towards and behaviours surrounding

sleep, and educates participants with one or more behavioural, relaxation, and sleep hygiene techniques. The intervention type and these components have been found to be effective in improving sleep quality in other studies (De Niet et al., 2009; Murawski et al., 2018). The ideal participants for this intervention would be female patients from a single fertility clinic who are planning to undergo IVF treatment to control for any variations in treatment procedures or care patients may receive from different clinics. Both sleep quality and mood factors, including stress, anxiety, and depression, should be measured via self-reports. To investigate the potential mechanisms through which the intervention improves mood and influences treatment outcomes, a number of "biomarkers" of stress, such as hair cortisol, immunoglobulin A, lysozyme, melatonin, and *alpha*-amylase measured in saliva, may also be collected at the same time. If possible, participants may be given wearable devices to monitor their sleep physically throughout the treatment period. Collection of self-report data and biomarkers should be done pre-treatment and pre-intervention as a baseline and then weekly throughout the treatment period (until patients receive their clinical pregnancy results). The last data collection for mood, sleep, and biomarkers should be conducted before patients receive their clinical pregnancy results to avoid acute effects on these factors from hearing the pregnancy results. The minimum time frame for studies measuring cortisol in hair is one month (four weeks), and a typical IVF cycle usually takes four to six weeks. Therefore, the intervention should ideally last four weeks, starting with the IVF treatment or in the pre-treatment week. Potential participants may be approached at their initial consultations with IVF doctors before starting treatment. Participants who consent to take part should be randomly assigned to either the intervention group or the control group in which they receive all of the same assessments as the intervention group (including the wearable devices for sleep monitoring) but receive usual care. Clinical outcomes (e.g., biochemical pregnancy, clinical pregnancy, and live birth) should be collected directly from the fertility clinic when available. Other patient characteristics that

may be associated with mood, sleep, biomarker indices, and IVF treatment outcomes, such as age, BMI, previous miscarriage, and duration trying to conceive, should also be collected from the clinic. Data analyses would then examine (1) whether levels of stress, anxiety, depression, sleep quality, and biomarker indices significantly improve in the intervention group, (2) whether these parameters significantly differ between the intervention and control groups before and after the intervention, (3) whether IVF treatment outcomes differ between the intervention and control group, and (4) whether the biomarkers mediate the intervention effects on mood, sleep, and IVF treatment outcomes.

7.7 Future recommendations

Although the study did not find any significant associations between mood, lifestyle, and hair cortisol, it demonstrated the feasibility of measuring these factors in patients undergoing IVF treatment. A future direction would be to consider measuring hair cortisol in psychosocial interventions among patients undergoing fertility treatments to help underpin the possible pathways through which these interventions benefit patients' wellbeing and treatment outcomes. Another future direction would be to consider measuring these factors at different stages of the treatment, which will provide a more global picture of how patients' mood and lifestyle change with the treatment procedure and the role of cortisol in those changes. This study can also be viewed as a steppingstone towards understanding how patients' lifestyles and mood are associated with their treatment, both directly and indirectly. Future studies should build on the null findings in this study and consider (1) other factors that may be mediating the relationship between lifestyle and mood and treatment outcomes (e.g., coping strategies and social support) and (2) a broad range of mechanisms through which these associations may happen (e.g., hormones involved in the HPG axis, oxidative stress, and immune cells).

7.7.1 Emotional impact of the COVID-19 pandemic on women undergoing fertility treatment

Towards the end of this study, a global pandemic caused by the SARS-CoV-2 virus (COVID-19) was declared by the WHO on 2020, March 11 (Organization et al., 2020). The pandemic disrupted society, health services, and the economy. Compelling evidence has demonstrated the rapid negative effect of the pandemic on mental health. Early observational studies across the world have reported a relatively high prevalence of anxiety (6%–51%), depression (15%–48%), PTSD (7%–54%), psychological distress (34%–38%), and stress (8%–82%) (Xiong et al., 2020). Cross-sectional and longitudinal evidence have also shown rapid increases in depression and anxiety following enforced restrictions (e.g., lockdowns) in the UK among the general population, with young people, women, and those with pre-existing mental health disorders more at risk than the rest of the population (Fancourt et al., 2021; Jia et al., 2021, 2020). The COVID-19 pandemic and associated social restrictions were universal stressors for all. Patients undergoing fertility treatment experienced the additional stressor of treatment being delayed or cancelled during this time (Lazarus and Cohen, 1977). According to a report published by the European Society of Human Reproduction and Embryology in March–April 2020, most fertility treatment centres around the world reduced or suspended routine treatments to reduce the risk of infection and to comply with the social distancing guidelines from governments (ESHRE COVID-19 Working Group et al., 2021). In the UK, the Human Fertilisation and Embryology Authority (2019) issued a directive to end all treatments by mid-April, with exceptions for cancer patients. This meant that no patients were able to access treatment or had to abandon or convert ongoing cycles during this time. Therefore, the sudden closure of fertility clinics contributed to considerable emotional distress among patients. In a mixed-method study on 450 patients (446 women and four men) undergoing fertility treatment Boivin et al. (2020) reported that although

patients expressed understanding towards the closure of clinics, most of them appraised this situation as being extremely stressful, uncontrollable, and likely to have negative impacts on their lives. They also outlined the themes reported by patients in relation to emotional reactions to clinic closures, which for almost all patients included feeling stressed, worried, and frustrated about the uncertainty. Some patients reported feeling aggrieved, angry, and resentful. In addition, 12% of the cohort reported feelings of hopelessness and sadness and a deterioration in their wellbeing (Boivin et al., 2020). In study on 181 female patients in Canada, Haham et al. (2021) found high levels of sadness (66%), anxiety (60%), and helplessness (60%) among women. In a study on 92 women from Canada and the United States, Gordon and Balsom (2020) found a high prevalence of clinical levels of depressive symptoms (52%) among participants. In a study in France on 435 participants, Lablanche et al. (2022) found that more than half experienced stress and 22% reported symptoms of anxiety.

Although fertility treatments have restarted since May 2020 in the UK, the COVID-19 pandemic can be regarded as a chronic stressor (Elliott and Eisdorfer, 1982), which has not yet been declared to have ended. Therefore, the long-term impact of the COVID-19 pandemic, especially the emotional distress resulting from clinic closures, on patients' mental and physical health still needs to be addressed (Boivin et al., 2020). Following the methods in this study, one suggestion is to examine hair cortisol and mental and physical health outcomes of interest (e.g., self-reported depression and clinical pregnancy rate). As hair cortisol can reliably capture the accumulated cortisol levels over the past 5–6 months (Russell et al., 2012), it is a promising tool for measuring the possible psychobiological effects of pandemic-related experiences. Some early evidence has demonstrated increased hair cortisol among healthcare workers during the pandemic, with one study (n = 234) reporting a positive association between stress and hair cortisol and another smaller study (n = 67) reporting significantly increased hair cortisol during the pandemic when compared with pre-pandemic levels (Ibar et al., 2021;

Rajcani et al., 2021). Such research among the population undergoing fertility treatment may provide insightful knowledge into how the COVID-19 pandemic influenced HPA function and treatment outcomes.

7.8 Conclusion

Despite these limitations, the present study adds to the evidence on cortisol and mood in IVF treatment. Based on the findings of this prospective study and the systematic review reported in Chapter 4, a revised model is proposed (see Figure 7.2 below). Although this study does not support either a direct association between pre-treatment depression and treatment outcomes or an indirect association between depression and treatment outcomes through hair cortisol, it has been observed that women undergoing IVF treatment are likely to report poor sleep and high levels of depression. Sleep quality, but not other lifestyle factors, was positively correlated with anxiety, depression, and negative mood and negatively correlated with positive mood. The study has also shown a possible positive association between depression and hair cortisol, which is supported by findings from the systematic review. This hints at the negative effect of long-term stress on the functioning of the HPA axis. In order to fully understand the effect of long-term stress on reproductive health, especially in the current environment of a global pandemic that has existed for more than two years, tools such as hair cortisol, which reflects the psychobiological process of long-term stress, may be useful. In conclusion, the influence of emotional and behavioural factors on fertility treatment and the possible biological pathways continue to need further exploration.

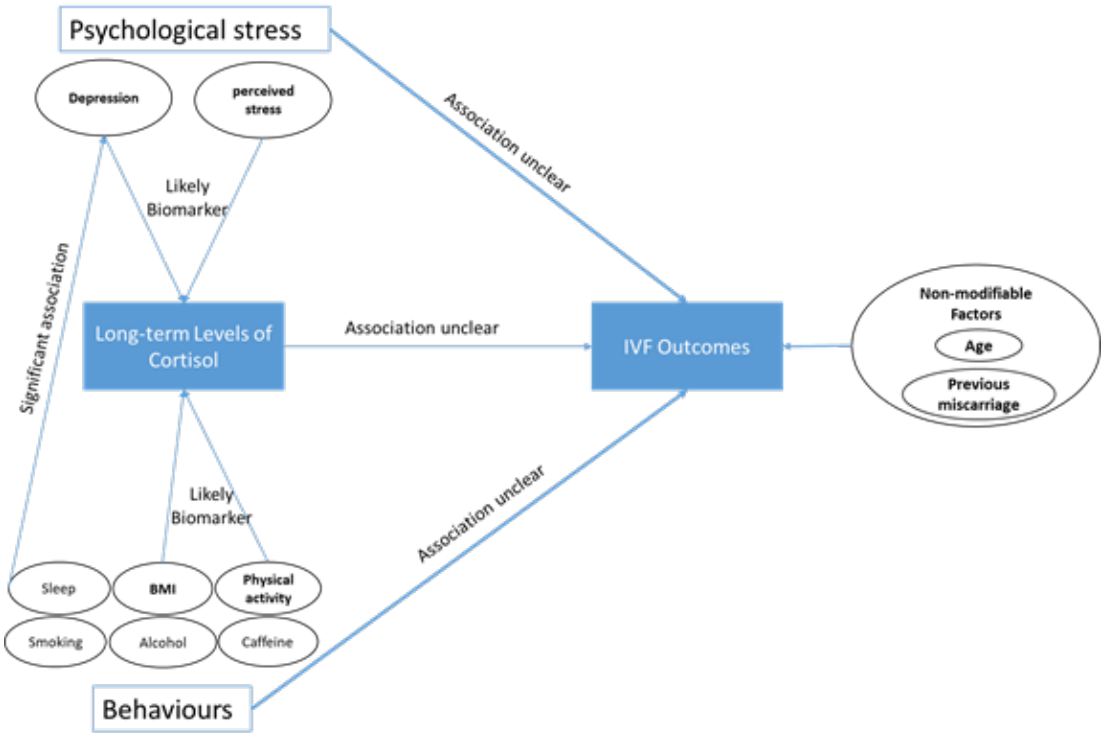


Figure 7.1 Modified thesis model based on findings from the cohort study and systematic review

References

Abadia, L., Chiu, Y.-H., Williams, P., Toth, T., Souter, I., Hauser, R., Chavarro, J., Gaskins, A., and Team, E. S. (2017). The association between pre-treatment maternal alcohol and caffeine intake and outcomes of assisted reproduction in a prospectively followed cohort. *Human Reproduction*, 32(9):1846–1854.

Abidin, R., Flens, J. R., and Austin, W. G. (2006). *The parenting stress index.* Lawrence Erlbaum Associates Publishers.

Adam, E. K., Doane, L. D., Zinbarg, R. E., Mineka, S., Craske, M. G., and Griffith, J. W. (2010). Prospective prediction of major depressive disorder from cortisol awakening responses in adolescence. *Psychoneuroendocrinology*, 35(6):921–931.

Adam, E. K., Quinn, M. E., Tavernier, R., McQuillan, M. T., Dahlke, K. A., and Gilbert, K. E. (2017). Diurnal cortisol slopes and mental and physical health outcomes: A systematic review and meta-analysis. *Psychoneuroendocrinology*, 83:25–41.

Adamson, G. D., de Mouzon, J., Chambers, G. M., Zegers-Hochschild, F., Mansour, R., Ishihara, O., Banker, M., and Dyer, S. (2018). International Committee for Monitoring Assisted Reproductive Technology: world report on assisted reproductive technology, 2011. *Fertility and sterility*, 110(6):1067–1080.

Aimagambetova, G., Issanov, A., Terzic, S., Bapayeva, G., Ukybassova, T., Baikoshkarova, S., Aldiyarova, A., Shauyen, F., and Terzic, M. (2020). The effect of psychological distress on ivf outcomes: Reality or speculations? *PloS one*, 15(12):e0242024.

Al-Saleh, I., El-Doush, I., Grisellhi, B., and Coskun, S. (2010). The effect of caffeine consumption on the success rate of pregnancy as well various performance parameters of in-vitro fertilization treatment. *Medical Science Monitor*, 16(12):CR598–CR605.

An, Y., Sun, Z., Li, L., Zhang, Y., and Ji, H. (2013). Relationship between psychological stress and reproductive outcome in women undergoing in vitro fertilization treatment: Psychological and neurohormonal assessment. *Journal of Assisted Reproduction and Genetics*, 30(1):35–41.

Andersen, C. Y. (2002). Possible new mechanism of cortisol action in female reproductive organs: physiological implications of the free hormone hypothesis. *The Journal of Endocrinology*, 173(2):211–217.

Arck, P. C. (2001). Stress and pregnancy loss: role of immune mediators, hormones and neurotransmitters. *American journal of reproductive immunology*, 46(2):117–123.

Armstrong, T. and Bull, F. (2006). Development of the world health organization global physical activity questionnaire (gpaq). *Journal of Public Health*, 14(2):66–70.

Augood, C., Duckitt, K., and Templeton, A. (1998). Smoking and female infertility: a systematic review and meta-analysis. *Human reproduction (Oxford, England)*, 13(6):1532–1539.

Ayers, S., Baum, A., McManus, C., Newman, S., Wallston, K., Weinman, J., and West, R. (2007). *Cambridge handbook of psychology, health and medicine*. Cambridge University Press.

Babu, A. S., Padmakumar, R., Maiya, A. G., Mohapatra, A. K., and Kamath, R. L. (2016). Effects of exercise training on exercise capacity in pulmonary arterial hypertension: a systematic review of clinical trials. *Heart, Lung and Circulation*, 25(4):333–341.

Badrick, E., Kirschbaum, C., and Kumari, M. (2007). The relationship between smoking status and cortisol secretion. *The Journal of Clinical Endocrinology & Metabolism*, 92(3):819–824.

Balen, A., Platteau, P., Andersen, A. N., Devroey, P., Sørensen, P., Helmgaard, L., and Arce, J.-C. (2006). The influence of body weight on response to ovulation induction with gonadotrophins in 335 women with world health organization group ii anovulatory infertility. *BJOG: An International Journal of Obstetrics & Gynaecology*, 113(10):1195–1202.

Bates, G. W., Bates, S. R., and Whitworth, N. S. (1982). Reproductive failure in women who practice weight control. *Fertility and sterility*, 37(3):373–378.

Bates Jr, G. W. and Ginsburg, E. S. (2002). Early pregnancy loss in in vitro fertilization (ivf) is a positive predictor of subsequent ivf success. *Fertility and sterility*, 77(2):337–341.

Batista, J. C., Souza, A. L., Ferreira, H. A., Canova, F., and Grassi-Kassisse, D. M. (2015). Acute and Chronic Effects of Tantric Yoga Practice on Distress Index. *The Journal of Alternative and Complementary Medicine*, 21(11):681–685.

Bergen-Cico, D., Possemato, K., and Pigeon, W. (2014). Reductions in Cortisol Associated With Primary Care Brief Mindfulness Program for Veterans With PTSD:. *Medical Care*, 52:S25–S31.

Boivin, J., Bunting, L., Collins, J. A., and Nygren, K. G. (2007). International estimates of infertility prevalence and treatment-seeking: Potential need and demand for infertility medical care. *Human Reproduction*, 22(6):1506–1512.

Boivin, J., Griffiths, E., and Venetis, C. A. (2011). Emotional distress in infertile women and failure of assisted reproductive technologies: Meta-analysis of prospective psychosocial studies. *BMJ*, 342(7795):481.

Boivin, J., Harrison, C., Mathur, R., Burns, G., Pericleous-Smith, A., and Gameiro, S. (2020). Patient experiences of fertility clinic closure during the covid-19 pandemic: appraisals, coping and emotions. *Human Reproduction*, 35(11):2556–2566.

Boivin, J. and Takefman, J. E. (1995). Stress level across stages of in vitro fertilization in subsequently pregnant and nonpregnant women. *Fertility and sterility*, 64(4):802–810.

Bolumar, F., Olsen, J., Rebagliato, M., Saez-Lloret, I., Bisanti, L., and the European Study Group on Infertility and Subfecundity (2000). Body Mass Index and Delayed Conception: A European Multicenter Study on Infertility and Subfecundity. *American Journal of Epidemiology*, 151(11):1072–1079.

Bonnet, M. H. and Arand, D. L. (2010). Hyperarousal and insomnia: state of the science. *Sleep medicine reviews*, 14(1):9–15.

Booij, S. H., Bos, E. H., Bouwmans, M. E., van Faassen, M., Kema, I. P., Oldehinkel, A. J., and de Jonge, P. (2015). Cortisol and α-amylase secretion patterns between and within depressed and non-depressed individuals. *PloS one*, 10(7):e0131002.

Boylan, S. M., Cade, J. E., Kirk, S. F. L., Greenwood, D. C., White, K. L. M., Shires, S., Simpson, N. A. B., Wild, C. P., and Hay, A. W. M. (2008). Assessing caffeine exposure in pregnant women. *British Journal of Nutrition*, 100(4):875–882.

Bradley, B. P. and Mogg, K. (1994). Mood and personality in recall of positive and negative information. *Behaviour research and therapy*, 32(1):137–141.

Braig, S., Grabher, F., Ntomchukwu, C., Reister, F., Stalder, T., Kirschbaum, C., Rothenbacher, D., and Genuneit, J. (2016). The Association of Hair Cortisol with Self-Reported Chronic Psychosocial Stress and Symptoms of Anxiety and Depression in Women Shortly after Delivery: Hair cortisol, self-reported chronic stress, and symptoms of anxiety and depression. *Paediatric and Perinatal Epidemiology*, 30(2):97–104.

Brannon, L. (2018). *Health psychology : an introduction to behavior and health / Linda Brannon, John A. Updegraff, Jess Feist.* Cengage Learning, Boston, Mass., 9th ed. edition.

Breen, K. M. and Karsch, F. J. (2004). Does cortisol inhibit pulsatile luteinizing hormone secretion at the hypothalamic or pituitary level? *Endocrinology*, 145(2):692–698.

Broekmans, F. J., Knauff, E. A., te Velde, E. R., Macklon, N. S., and Fauser, B. C. (2007). Female reproductive ageing: current knowledge and future trends. *Trends in Endocrinology & Metabolism*, 18(2):58–65.

Brown, B. D., Thomas, W., Hutchins, A., Martini, M. C., and Slavin, J. L. (2002). Types of Dietary Fat and Soy Minimally Affect Hormones and Biomarkers Associated With Breast Cancer Risk in Premenopausal Women. *Nutrition and Cancer*, 43(1):22–30.

Brown, K. W., Coogle, C. L., and Wegelin, J. (2016). A pilot randomized controlled trial of mindfulness-based stress reduction for caregivers of family members with dementia. *Aging & Mental Health*, 20(11):1157–1166.

Brugo-Olmedo, S., Chillik, C., and Kopelman, S. (2001). Definition and causes of infertility. *Reproductive BioMedicine Online*, 2(1):173–185.

Burton, R. and Sheron, N. (2018). No level of alcohol consumption improves health. *The Lancet*, 392(10152):987–988.

Buysse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R., and Kupfer, D. J. (1989). The pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry research*, 28(2):193–213.

Byun, E. (2013). *Effects of uncertainty on perceived and physiological stress and psychological outcomes in stroke-survivor caregivers*. University of Pennsylvania.

C Santa-Cruz, D., Caparros-Gonzalez, R. A., Romero-Gonzalez, B., Peralta-Ramirez, M. I., Gonzalez-Perez, R., and García-Velasco, J. A. (2020). Hair cortisol concentrations as a biomarker to predict a clinical pregnancy outcome after an ivf cycle: A pilot feasibility study. *International journal of environmental research and public health*, 17(9):3020.

Cacioppo, J. T., Tassinary, L. G., and Berntson, G. (2007). *Handbook of psychophysiology*. Cambridge university press.

Caetano, G., Bozinovic, I., Dupont, C., Léger, D., Lévy, R., and Sermondade, N. (2021). Impact of sleep on female and male reproductive functions: a systematic review. *Fertility and Sterility*, 115(3):715–731.

Cameron, A., Henley, D., Carrell, R., Zhou, A., Clarke, A., and Lightman, S. (2010). Temperature-responsive release of cortisol from its binding globulin: a protein thermocouple. *The Journal of Clinical Endocrinology & Metabolism*, 95(10):4689–4695.

Cannon, W. B. (1929). Organization for physiological homeostasis. *Physiological reviews*, 9(3):399–431.

Cannon, W. B. (1939). The wisdom of the body.

Caparros-Gonzalez, R. A., Romero-Gonzalez, B., Strivens-Vilchez, H., Gonzalez-Perez, R., Martinez-Augustin, O., and Peralta-Ramirez, M. I. (2017). Hair cortisol levels, psychological stress and psychopathological symptoms as predictors of postpartum depression. *PLOS ONE*, 12(8):e0182817.

Cappuccio, F. P., D'Elia, L., Strazzullo, P., and Miller, M. A. (2010). Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. *Sleep*, 33(5):585–592.

CARE Study Group and others (2008). Maternal caffeine intake during pregnancy and risk of fetal growth restriction: a large prospective observational study. *The BMJ*, 337.

Casals, G. and Hanzu, F. A. (2020). Cortisol measurements in cushing's syndrome: Immunoassay or mass spectrometry? *Annals of laboratory medicine*, 40(4):285–296.

Caspersen, C. J., Powell, K. E., and Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public health reports*, 100(2):126.

Cesta, C. E., Johansson, A. L., Hreinsson, J., Rodriguez-Wallberg, K. A., Olofsson, J. I., Holte, J., Wramsby, H., Wramsby, M., Cnattingius, S., Skalkidou, A., et al. (2018). A prospective investigation of perceived stress, infertility-related stress, and cortisol levels in women undergoing in vitro fertilization: influence on embryo quality and clinical pregnancy rate. *Acta obstetricia et gynecologica Scandinavica*, 97(3):258–268.

Chan, C. H., Chan, C. L., Ng, E. H., Ho, P. C., Chan, T. H., Lee, G. L., and Hui, W. (2012). Incorporating spirituality in psychosocial group intervention for women undergoing in vitro fertilization: a prospective randomized controlled study. *Psychology and Psychotherapy: Theory, Research and Practice*, 85(4):356–373.

Chen, P.-J., Yang, L., Chou, C.-C., Li, C.-C., Chang, Y.-C., and Liaw, J.-J. (2017). Effects of prenatal yoga on women's stress and immune function across pregnancy: A randomized controlled trial. *Complementary Therapies in Medicine*, 31:109–117.

Chen, X., Gelaye, B., Velez, J. C., Barbosa, C., Pepper, M., Andrade, A., Gao, W., Kirschbaum, C., and Williams, M. A. (2015). Caregivers' hair cortisol: a possible biomarker of chronic stress is associated with obesity measures among children with disabilities. *BMC Pediatrics*, 15(1):9.

Cheung, C. W. C., Saravelos, S. H., Chan, T. Y. A., Sahota, D. S., Wang, C. C., Chung, P. W., and Li, T. C. (2019). A prospective observational study on the stress levels at the time of embryo transfer and pregnancy testing following in vitro fertilisation treatment: a comparison between women with different treatment outcomes. *BJOG: An International Journal of Obstetrics & Gynaecology*, 126(2):271–279. Publisher: John Wiley & Sons, Ltd.

Chiodera, P., Volpi, R., Capretti, L., Speroni, G., Necchi-Ghiri, S., Caffarri, G., Colla, R., and Coiro, V. (1997). Abnormal effect of cigarette smoking on pituitary hormone secretions in insulin-dependent diabetes mellitus. *Clinical endocrinology*, 46(3):351–357.

Chistyakova, G., Remizova, I., Gazieva, I., and hermyaninova, O. (2014). Immunological and hemostasiological disorders in women with ovarian hyperstimulation syndrome. *Gynecological Endocrinology*, 30(sup1):39–42.

Chou, K.-H. and Bell, L. (2007). Caffeine content of prepackaged national-brand and private-label carbonated beverages. *Journal of food science*, 72(6):C337–C342.

Chourpiliadis, C. and Aeddula, N. R. (2020). Physiology, glucocorticoids. *StatPearls [Internet]*.

Cieslak, R., Benight, C. C., Luszczynska, A., and Laudenslager, M. L. (2011). Longitudinal relationships between self-efficacy, post-traumatic distress and salivary cortisol among motor vehicle accident survivors: Cortisol, Self-efficacy and PTSD. *Stress and Health*, 27(3):e261–e268.

Claman, P., On, O., Cheung, A. P., Goodrow, G. J., On, M., Hughes, E. G., On, H., Min, J. K., On, C., Roberts, J., Bc, B., Senikas, V., Chee, B., Wong, M., Ab, C., Young, D. C., and Ns, H. (2011). The Diagnosis and Management of Ovarian Hyperstimulation Syndrome. Technical report.

Clark, A., Roberts, B., Galletly, C., Tomlinson, L., and Norman, R. (2000). Maximizing weight loss in the overweight infertile patient-a prospective randomized controlled trial. In *Human Reproduction*, volume 15, pages 65–66. OXFORD UNIV PRESS GREAT CLARENDON ST, OXFORD OX2 6DP, ENGLAND.

Clow, A., Hucklebridge, F., and Thorn, L. (2010). The cortisol awakening response in context. *International review of neurobiology*, 93:153–175.

Clow, A., Thorn, L., Evans, P., and Hucklebridge, F. (2004). The awakening cortisol response: methodological issues and significance. *Stress*, 7(1):29–37.

Cohen, F. and Lazarus, R. S. (1973). Active coping processes, coping dispositions and recovery from surgery.

Cohen, S., Kamarck, T., and Mermelstein, R. (1983). A global measure of perceived stress. *Journal of health and social behavior*, pages 385–396.

Coiro, V., Casti, A., Saccani Jotti, G., Rubino, P., Manfredi, G., Maffei, M. L., Melani, A., Volta, E., and Chiodera, P. (2007). Adrenocorticotropic Hormone/Cortisol Response to Physical Exercise in Abstinent Alcoholic Patients. *Alcoholism: Clinical and Experimental Research*, 31(5):901–906.

Collins, J. A. and Rowe, T. C. (1989). Age of the female partner is a prognostic factor in prolonged unexplained infertility: A multicenter study. *Fertility and Sterility*, 52(1):15–20.

Coolens, J.-L., Van Baelen, H., and Heyns, W. (1987). Clinical use of unbound plasma cortisol as calculated from total cortisol and corticosteroid-binding globulin. *Journal of Steroid Biochemistry*, 26(2):197–202.

Coward, K. and Wells, D. (2013). *Textbook of clinical embryology*. Cambridge University Press.

Cui, D., Han, G., Shang, Y., Liu, C., Xia, L., Li, L., and Yi, S. (2015). Antisperm antibodies in infertile men and their effect on semen parameters: a systematic review and meta-analysis. *Clinica Chimica Acta*, 444:29–36.

Datta, J., Palmer, M. J., Tanton, C., Gibson, L. J., Jones, K. G., Macdowall, W., Glasier, A., Sonnenberg, P., Field, N., Mercer, C. H., Johnson, A. M., and Wellings, K. (2016). Prevalence of infertilityand help seeking among 15 000 women and men. *Human Reproduction*, 31(9):2108–2118.

Daubenmier, J., Kristeller, J., Hecht, F. M., Maninger, N., Kuwata, M., Jhaveri, K., Lustig, R. H., Kemeny, M., Karan, L., and Epel, E. (2011). Mindfulness intervention for stress eating to reduce cortisol and abdominal fat among overweight and obese women: an exploratory randomized controlled study. *Journal of obesity*, 2011.

Dawson, A., Diedrich, K., and Felberbaum, R. (2005). Why do couples refuse or discontinue art? *Archives of Gynecology and Obstetrics*, 273(1):3–11.

De, D., Roy, P. K., and Sarkhel, S. (2017). A psychological study of male, female related and unexplained infertility in Indian urban couples. *Journal of Reproductive and Infant Psychology*, 35(4):353–364.

De Geyter, C., Calhaz-Jorge, C., Kupka, M. S., Wyns, C., Mocanu, E., Motrenko, T., Scaravelli, G., Smeenk, J., Vidakovic, S., Goossens, V., The European IVF-monitoring Consortium (EIM) for the European Society of Human Reproduction and Embryology (ESHRE), Gliozheni, O., Strohmer, H., Obruca, Kreuz-Kinderwunschzentrum, S. P. G., Petrovskaya, E., Tishkevich, O., Wyns, C., Bogaerts, K., Balic, D., Sibincic, S., Antonova, I., Vrcic, H., Ljiljak, D., Pelekanos, M., Rezabek, K., Markova, J., Lemmen, J., Sõritsa, D., Gissler, M., Tiitinen, A., Royere, D., Tandler—Schneider, A., Kimmel, M.,

Antsaklis, A. J., Loutradis, D., Urbancsek, J., Kosztolanyi, G., Bjorgvinsson, H., Mocanu, E., Scaravelli, G., de Luca, R., Lokshin, V., Ravil, V., Magomedova, V., Gudleviciene, Z., Belo lopes, G., Petanovski, Z., Calleja-Agius, J., Xuereb, J., Moshin, V., Simic, T. M., Vukicevic, D., Romundstad, L. B., Janicka, A., Calhaz-Jorge, C., Laranjeira, A. R., Rugescu, I., Doroftei, B., Korsak, V., Radunovic, N., Tabs, N., Virant-Klun, I., Saiz, I. C., Mondéjar, F. P., Bergh, C., Weder, M., De Geyter, C., Smeenk, J. M. J., Gryshchenko, M., and Baranowski, R. (2018). ART in Europe, 2014: results generated from European registries by ESHRE†. *Human Reproduction*, 33(9):1586–1601.

de Kloet, E. R. and Sarabdjitsingh, R. A. (2008). Everything Has Rhythm: Focus on Glucocorticoid Pulsatility. *Endocrinology*, 149(7):3241–3243.

De Los Santos, M., Apter, S., Coticchio, G., Debrock, S., Lundin, K., Plancha, C., Prados, F., Rienzi, L., Verheyen, G., Woodward, B., and Vermeulen, N. (2015). ESHRE Guideline Group on Good Practice in IVF Labs. Technical report.

De Niet, G. J., Tiemens, B. G., Kloos, M. W., and Hutschemaekers, G. J. (2009). Review of systematic reviews about the efficacy of non-pharmacological interventions to improve sleep quality in insomnia. *International Journal of Evidence-Based Healthcare*, 7(4):233–242.

de Oliveira, M. d. G., Fonseca, P. P., and Del Porto, J. A. (2003). Inter-rater reliability of the brazilian version of the paykel scale of stressful life events. *Psychopathology*, 36(6):320–323.

Dean, K. (1989). Self-care components of lifestyles: The importance of gender, attitudes and the social situation. *Social Science & Medicine*, 29(2):137–152.

Deeks, J. J., Dinnes, J., D'Amico, R., Sowden, A. J., Sakarovitch, C., Song, F., Petticrew, M., Altman, D., et al. (2003). Evaluating non-randomised

intervention studies. *Health technology assessment (Winchester, England)*, 7(27):iii–173.

Delvigne, A. and Rozenberg, S. (2002). Epidemiology and prevention of ovarian hyperstimulation syndrome (ohss): a review. *Human reproduction update*, 8(6):559–577.

Dettenborn, L., Rosenloecher, F., and Kirschbaum, C. (2007). No effects of repeated forced wakings during three consecutive nights on morning cortisol awakening responses (car): a preliminary study. *Psychoneuroendocrinology*, 32(8-10):915–921.

DeVellis, R. F. (2003). *Scale development : theory and applications.* Applied social research methods series ; v. 26. SAGE, Thousand Oaks, Calif. ; London, 2nd ed. / robert f. devellis edition.

Devroey, P. and Van Steirteghem, A. (2004). A review of ten years experience of ICSI.

Diener, E., Larsen, R. J., and Emmons, R. (1984). Bias in mood recall in happy and unhappy persons. In *92nd Annual Convention of the American Psychological Association, Toronto, Ontario, Canada*.

Diment, B. C., Fortes, M. B., Greeves, J. P., Casey, A., Costa, R. J. S., Walters, R., and Walsh, N. P. (2012). Effect of daily mixed nutritional supplementation on immune indices in soldiers undertaking an 8-week arduous training programme. *European Journal of Applied Physiology*, 112(4):1411–1418.

Domar, A. D., Broome, A., Zuttermeister, P. C., Seibel, M., and Friedman, R. (1992). The prevalence and predictability of depression in infertile women**Supported by grant 5R03MH45591 from the National Institute of Mental Health, Bethesda, Maryland.††Presented at the Annual Meeting of The American Fertility Society, Orlando, Florida, October 21 to 24, 1991. *Fertility and Sterility*, 58(6):1158–1163.

Domar, A. D., Rooney, K. L., Milstein, M., and Conboy, L. (2015). Lifestyle habits of 12,800 ivf patients: Prevalence of negative lifestyle behaviors, and impact of region and insurance coverage. *Human Fertility*, 18(4):253–257.

Domar, A. D., Rooney, K. L., Wiegand, B., Orav, E. J., Alper, M. M., Berger, B. M., and Nikolovski, J. (2011). Impact of a group mind/body intervention on pregnancy rates in ivf patients. *Fertility and sterility*, 95(7):2269–2273.

Donnelly, J. E., Hillman, C. H., Castelli, D., Etnier, J. L., Lee, S., Tomporowski, P., Lambourne, K., and Szabo-Reed, A. N. (2016). Physical activity, fitness, cognitive function, and academic achievement in children: a systematic review. *Medicine and science in sports and exercise*, 48(6):1197.

Downs, S. H. and Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health*, 52(6):377–384.

Dubuisson, J. B., Aubriot, F. X., Mathieu, L., Foulot, H., Mandelbrot, L., and de Jolinière, J. B. (1991). Risk factors for ectopic pregnancy in 556 pregnancies after in vitro fertilization: implications for preventive management. *Fertility and Sterility*, 56(4):686–690.

Ebbesen, S., Zachariae, R., Mehlsen, M., Thomsen, D., Højgaard, A., Ottosen, L., Petersen, T., and Ingerslev, H. (2009). Stressful life events are associated with a poor in-vitro fertilization (ivf) outcome: a prospective study. *Human reproduction*, 24(9):2173–2182.

Edwards, R. G., Steptoe, P. C., and Purdy, J. M. (1980). Establishing full-term human pregnancies using cleaving embryos grown in vitro. *BJOG: An International Journal of Obstetrics & Gynaecology*, 87(9):737–756.

El Mlili, N., Ahabrach, H., and Cauli, O. (2021). Hair cortisol concentration as a biomarker of sleep quality and related disorders. *Life*, 11(2):81.

Elias, A. and Wilson, A. (1993). Exercise and gonadal function. *Human reproduction*, 8(10):1747–1761.

Ellenbogen, M. A., Hodgins, S., Linnen, A.-M., and Ostiguy, C. S. (2011). Elevated daytime cortisol levels: a biomarker of subsequent major affective disorder? *Journal of affective disorders*, 132(1-2):265–269.

Elliott, G. and Eisdorfer, C. (1982). Stress and human health: A study by the institute of medicine. *Washington, DC: National Academy of Sciences.[Google Scholar]*.

Ellis, J. (2007). *Sleep and health*, page 198–202. Cambridge University Press, 2 edition.

Emanuele, M. A., Wezeman, F., and Emanuele, N. V. (2002). Alcohol's effects on female reproductive function. *Alcohol Research & Health*, 26(4):274.

Eryilmaz, O. G., Devran, A., Sarikaya, E., Aksakal, F. N., Mollamahmutoğlu, L., and Cicek, N. (2011). Melatonin improves the oocyte and the embryo in IVF patients with sleep disturbances, but does not improve the sleeping problems. *Journal of Assisted Reproduction and Genetics*, 28(9):815–820.

ESHRE COVID-19 Working Group, Gianaroli, L., Ata, B., Lundin, K., Rautakallio-Hokkanen, S., Tapanainen, J. S., Vermeulen, N., Veiga, A., and Mocanu, E. (2021). The calm after the storm: re-starting art treatments safely in the wake of the covid-19 pandemic. *Human Reproduction*, 36(2):275–282.

Esposito, A. and Bianchi, V. (2012). *Cortisol [electronic resource] : physiology, regulation and health implications / Alonzo Esposito and Vito Bianchi, editors.* Human anatomy and physiology series. Nova Science Publishers, New York.

Eugster, A. and Vingerhoets, A. J. J. M. (1999). Psychological aspects of in vitro fertilization: a review. *Social Science & Medicine*, 5(48):579–589.

Everitt, B. S. (2016). Food, weight and lifestyle: Is eating too much and moving too little bad for your health? In *Health and Lifestyle*, pages 97–124. Springer.

Ewart, C. K. and Suchday, S. (2002). Discovering how urban poverty and violence affect health: development and validation of a neighborhood stress index. *Health Psychology*, 21(3):254.

Facchinetti, F., Volpe, A., Matteo, M. L., Genazzani, A. R., and Artini, G. P. (1997). An increased vulnerability to stress is associated with a poor outcome of in vitro fertilization-embryo transfer treatment. *Fertility and sterility*, 67(2):309–314.

Fancourt, D., Steptoe, A., and Bu, F. (2021). Trajectories of anxiety and depressive symptoms during enforced isolation due to covid-19 in england: a longitudinal observational study. *The Lancet Psychiatry*, 8(2):141–149.

Faresjö, Å., Theodorsson, E., Chatziarzenis, M., Sapouna, V., Claesson, H.-P., Koppner, J., and Faresjö, T. (2013). Higher perceived stress but lower cortisol levels found among young greek adults living in a stressful social environment in comparison with swedish young adults. *PloS one*, 8(9):e73828.

Farhud, D. D. (2015). Impact of lifestyle on health. *Iranian journal of public health*, 44(11):1442.

Fauser, B. C. and Edwards, R. G. (2005). The early days of IVF. *Human Reproduction Update*, 11(5):437–438.

Federenko, I. S., Nagamine, M., Hellhammer, D. H., Wadhwa, P. D., and Wust, S. (2004). The heritability of hypothalamus pituitary adrenal axis responses to psychosocial stress is context dependent. *The Journal of Clinical Endocrinology & Metabolism*, 89(12):6244–6250.

Fedorcsák, P., Dale, P. O., Storeng, R., Ertzeid, G., Bjercke, S., Oldereid, N., Omland, A. K., Åbyholm, T., and Tanbo, T. (2004). Impact of overweight

and underweight on assisted reproduction treatment. *Human Reproduction*, 19(11):2523–2528.

Feller, S., Vigl, M., Bergmann, M. M., Boeing, H., Kirschbaum, C., and Stalder, T. (2014). Predictors of hair cortisol concentrations in older adults. *Psychoneuroendocrinology*, 39:132–140.

Ferrie, J. E., Shipley, M. J., Cappuccio, F. P., Brunner, E., Miller, M. A., Kumari, M., and Marmot, M. G. (2007). A Prospective Study of Change in Sleep Duration: Associations with Mortality in the Whitehall II Cohort. *Sleep*, 30(12):1659–1666.

Ferrini, R. L. and Barrett-Connor, E. (1996). Caffeine intake and endogenous sex steroid levels in postmenopausal women the rancho bernardo study. *American journal of epidemiology*, 144(7):642–644.

Firns, S., Cruzat, V. F., Keane, K. N., Joesbury, K. A., Lee, A. H., Newsholme, P., and Yovich, J. L. (2015). The effect of cigarette smoking, alcohol consumption and fruit and vegetable consumption on ivf outcomes: a review and presentation of original data. *Reproductive Biology and Endocrinology*, 13(1):1–13.

Flegal, K. M., Kit, B. K., Orpana, H., and Graubard, B. I. (2013). Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *Jama*, 309(1):71–82.

Folkman, S. (1997). Positive psychological states and coping with severe stress. *Social science & medicine*, 45(8):1207–1221.

Folkman, S., Chesney, M., McKusick, L., Ironson, G., Johnson, D. S., and Coates, T. J. (1991). Translating coping theory into an intervention. In *The social context of coping*, pages 239–260. Springer.

Forouzanfar, M. H., Afshin, A., Alexander, L. T., Anderson, H. R., Bhutta, Z. A., Biryukov, S., Brauer, M., Burnett, R., Cercy, K., Charlson, F. J., et al. (2016). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the global burden of disease study 2015. *The lancet*, 388(10053):1659–1724.

Frederiksen, Y., Farver-Vestergaard, I., Skovgard, N. G., Ingerslev, H. J., and Zachariae, R. (2015). Efficacy of psychosocial interventions for psychological and pregnancy outcomes in infertile women and men: a systematic review and meta-analysis. *BMJ Open*, 5(1):e006592–e006592.

Fries, E., Dettenborn, L., and Kirschbaum, C. (2009). The cortisol awakening response (car): facts and future directions. *International journal of Psychophysiology*, 72(1):67–73.

Fukuda, S. and Morimoto, K. (2001). Lifestyle, stress and cortisol response: Review I: Mental stress. *Environmental Health and Preventive Medicine*, 6(1):9–14.

Fulkerson, W. and Tang, B. (1979). Ultradian and circadian rhythms in the plasma concentration of cortisol in sheep. *Journal of endocrinology*, 81(1):135–141.

Gallinelli, A., Roncaglia, R., Matteo, M. L., Ciaccio, I., Volpe, A., and Facchinetti, F. (2001). Immunological changes and stress are associated with different implantation rates in patients undergoing in vitro fertilization–embryo transfer. *Fertility and Sterility*, 76(1):85–91.

Gameiro, S., Boivin, J., Peronace, L., and Verhaak, C. M. (2012). Why do patients discontinue fertility treatment? a systematic review of reasons and predictors of discontinuation in fertility treatment. *Human reproduction update*, 18(6):652–669.

Gao, W., Kirschbaum, C., Grass, J., and Stalder, T. (2016). Lc–ms based analysis of endogenous steroid hormones in human hair. *The Journal of steroid biochemistry and molecular biology*, 162:92–99.

Garcia-Leon, M. A., Peralta-Ramirez, M. I., Arco-Garcia, L., Romero-Gonzalez, B., Caparros-Gonzalez, R. A., Saez-Sanz, N., Santos-Ruiz, A. M., Montero-Lopez, E., Gonzalez, A., and Gonzalez-Perez, R. (2018). Hair cortisol concentrations in a Spanish sample of healthy adults. *PLOS ONE*, 13(9):e0204807.

Geller, P. A., Kerns, D., and Klier, C. M. (2004). Anxiety following miscarriage and the subsequent pregnancy: a review of the literature and future directions. *Journal of psychosomatic research*, 56(1):35–45.

Geng, L., Xiang, P., Yang, J., Shen, H., and Sang, Z. (2016). Association between hair cortisol concentration and perceived stress in female methamphetamine addicts. *Journal of Psychosomatic Research*, 91:82–86.

Gerber, M., Kalak, N., Elliot, C., Holsboer-Trachsler, E., Pühse, U., and Brand, S. (2013). Both Hair Cortisol Levels and Perceived Stress Predict Increased Symptoms of Depression: An Exploratory Study in Young Adults. *Neuropsychobiology*, 68(2):100–109.

Giesbrecht, G. F., Poole, J. C., Letourneau, N., Campbell, T., and Kaplan, B. J. (2013). The Buffering Effect of Social Support on Hypothalamic-Pituitary-Adrenal Axis Function During Pregnancy:. *Psychosomatic Medicine*, 75(9):856–862.

Gleicher, N., Weghofer, A., and Barad, D. H. (2011). Defining ovarian reserve to better understand ovarian aging.

Goldberg, S. B., Manley, A. R., Smith, S. S., Greeson, J. M., Russell, E., Van Uum, S., Koren, G., and Davis, J. M. (2014). Hair Cortisol as a Biomarker of Stress in Mindfulness Training for Smokers. *The Journal of Alternative and Complementary Medicine*, 20(8):630–634.

Goldstein, C. A., Lanham, M. S., Smith, Y. R., and O'Brien, L. M. (2017). Sleep in women undergoing in vitro fertilization: a pilot study. *Sleep medicine*, 32:105–113.

Gonen, L. D. and Bokek-Cohen, Y. (2020). The experience of being an in vitro fertilization (IVF) patient in a pro-natalist society - a form of emotional labor. *Social Work in Health Care*, 59(5):273–299.

Gordon, J. L. and Balsom, A. A. (2020). The psychological impact of fertility treatment suspensions during the covid-19 pandemic. *PLoS One*, 15(9):e0239253.

Gormack, A. A., Peek, J. C., Derraik, J. G. B., Gluckman, P. D., Young, N. L., and Cutfield, W. S. (2015). Many women undergoing fertility treatment make poor lifestyle choices that may affect treatment outcome. *Human Reproduction*, 30(7):1617–1624.

Gourounti, K., Anagnostopoulos, F., Potamianos, G., Lykeridou, K., Schmidt, L., and Vaslamatzis, G. (2012). Perception of control, coping and psychological stress of infertile women undergoing IVF. *Reproductive BioMedicine Online*, 24(6):670–679.

Grant, S. P., Mayo-Wilson, E., Melendez-Torres, G. J., and Montgomery, P. (2013). Reporting Quality of Social and Psychological Intervention Trials: A Systematic Review of Reporting Guidelines and Trial Publications. *PLoS ONE*, 8(5):e65442.

Grassi-Oliveira, R., Pezzi, J. C., Daruy-Filho, L., Viola, T. W., Francke, I. D. A., Leite, C. E., and Brietzke, E. (2012). Hair Cortisol and Stressful Life Events Retrospective Assessment in Crack Cocaine Users. *The American Journal of Drug and Alcohol Abuse*, 38(6):535–538.

Greff, M. J., Levine, J. M., Abuzgaia, A. M., Elzagallaai, A. A., Rieder, M. J., and

van Uum, S. H. (2019). Hair cortisol analysis: an update on methodological considerations and clinical applications. *Clinical biochemistry*, 63:1–9.

Grosso, G., Godos, J., Galvano, F., and Giovannucci, E. L. (2017). Coffee, Caffeine, and Health Outcomes: An Umbrella Review. *Annual Review of Nutrition*, 37(1):131–156.

Haham, L. M., Youngster, M., Shani, A. K., Yee, S., Ben-Kimhy, R., Medina-Artom, T. R., Hourvitz, A., Kedem, A., and Librach, C. (2021). Suspension of fertility treatment during the covid-19 pandemic: views, emotional reactions and psychological distress among women undergoing fertility treatment. *Reproductive biomedicine online*, 42(4):849–858.

Hakimi, O. and Cameron, L.-C. (2017). Effect of Exercise on Ovulation: A Systematic Review. *Sports Medicine*, 47(8):1555–1567.

Haller, J., Halasz, J., Mikics, E., Kruk, M., and Makara, G. (2000). Ultradian corticosterone rhythm and the propensity to behave aggressively in male rats. *Journal of neuroendocrinology*, 12(10):937–940.

Hansen, M., Bower, C., Milne, E., de Klerk, N., and J.Kurinczuk, J. (2005). Assisted reproductive technologies and the risk of birth defects—a systematic review. *Human Reproduction*, 20(2):328–338.

Harris, T. O., Borsanyi, S., Messari, S., Stanford, K., Cleary, S. E., Shiers, H. M., Brown, G. W., and Herbert, J. (2000). Morning cortisol as a risk factor for subsequent major depressive disorder in adult women. *British Journal of Psychiatry*, 177(6):505–510.

Harrison, C. L., Brown, W. J., Hayman, M., Moran, L. J., and Redman, L. M. (2016). The role of physical activity in preconception, pregnancy and postpartum health. In *Seminars in reproductive medicine*, volume 34, pages e28–e37. Thieme Medical Publishers.

Hasanpoor-Azghdy, S. B., Simbar, M., and Vedadhir, A. (2014). The emotional-psychological consequences of infertility among infertile women seeking treatment: Results of a qualitative study. *Iranian journal of reproductive medicine*, 12(2):131–8.

Hassan, M. A. and Killick, S. R. (2004). Negative lifestyle is associated with a significant reduction in fecundity. *Fertility and sterility*, 81(2):384–392.

Hawkins, L. K., Rossi, B. V., Correia, K. F., Lipskind, S. T., Hornstein, M. D., and Missmer, S. A. (2014). Perceptions among infertile couples of lifestyle behaviors and in vitro fertilization (IVF) success. *Journal of Assisted Reproduction and Genetics*, 31(3):255–260.

Hayashi, K. T. and Moberg, G. P. (1989). Influence of the hypothalamic-pituitary-adrenal axis on the menstrual cycle and the pituitary responsiveness to estradlol in the female rhesus monkey (macaca mulatta). *Biology of reproduction*, 42(2):260–265.

Heinze, K., Lin, A., Reniers, R. L., and Wood, S. J. (2016). Longer-term increased cortisol levels in young people with mental health problems. *Psychiatry Research*, 236:98–104.

Hellhammer, D. H., Wüst, S., and Kudielka, B. M. (2009). Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology*, 34(2):163–171.

Henley, D. E., Leendertz, J., Russell, G., Wood, S., Taheri, S., Woltersdorf, W., and Lightman, S. (2009). Development of an automated blood sampling system for use in humans. *Journal of medical engineering & technology*, 33(3):199–208.

Higdon, J. V. and Frei, B. (2006). Coffee and Health: A Review of Recent Human Research. *Critical Reviews in Food Science and Nutrition*, 46(2):101–123.

Hoffman, M. C., Mazzoni, S. E., Wagner, B. D., Laudenslager, M. L., and Ross,

R. G. (2016). Measures of Maternal Stress and Mood in Relation to Preterm Birth. *Obstetrics & Gynecology*, 127(3):545–552.

Holaday, J. W., Martinez, H. M., and Natelson, B. H. (1977). Synchronized ultradian cortisol rhythms in monkeys: persistence during corticotropin infusion. *Science*, 198(4312):56–58.

Holt-Lunstad, J., Birmingham, W. A., and Light, K. C. (2008). Influence of a "Warm Touch" Support Enhancement Intervention Among Married Couples on Ambulatory Blood Pressure, Oxytocin, Alpha Amylase, and Cortisol:. *Psychosomatic Medicine*, 70(9):976–985.

Homan, G., Davies, M., and Norman, R. (2007). The impact of lifestyle factors on reproductive performance in the general population and those undergoing infertility treatment: a review. *Human Reproduction Update*, 13(3):209–223.

Horowitz, M., Wilner, N., and Alvarez, W. (1979). Impact of event scale: A measure of subjective stress. *Psychosomatic medicine*, 41(3):209–218.

Horrocks, P., Jones, A., Ratcliffe, W., Holder, G., White, A., Holder, R., Ratcliffe, J., and London, D. (1990). Patterns of acth and cortisol pulsatility over twenty-four hours in normal males and females. *Clinical endocrinology*, 32(1):127–134.

Huang, B., Hu, D., Qian, K., Ai, J., Li, Y., Jin, L., Zhu, G., and Zhang, H. (2014). Is frozen embryo transfer cycle associated with a significantly lower incidence of ectopic pregnancy? An analysis of more than 30,000 cycles. *Fertility and Sterility*, 102(5):1345–1349.

Huang, L.-H., Kuo, C.-P., Lu, Y.-C., Lee, M.-S., and Lee, S.-H. (2019). Association of emotional distress and quality of sleep among women receiving in-vitro fertilization treatment. *Taiwanese Journal of Obstetrics and Gynecology*, 58(1):168–172.

Huang, W., Taylor, A., Howie, J., and Robinson, N. (2012). Is the Diurnal Profile of Salivary Cortisol Concentration a Useful Marker for Measuring Reported Stress in Acupuncture Research? A Randomized Controlled Pilot Study. *The Journal of Alternative and Complementary Medicine*, 18(3):242–250.

Huffer, D., Hing, W., Newton, R., and Clair, M. (2017). Strength training for plantar fasciitis and the intrinsic foot musculature: A systematic review. *Physical Therapy in Sport*, 24:44–52.

Hull, M. G., North, K., Taylor, H., Farrow, A., Ford, W. C. L., et al. (2000). Delayed conception and active and passive smoking. *Fertility and sterility*, 74(4):725–733.

Human Fertilisation and Embryology Authority (2018). Fertility treatment 2014–2016: Trends and figures. *London: HFEA*.

Humeniuk, R., Ali, R., Babor, T. F., Farrell, M., Formigoni, M. L., Jittiwutikarn, J., De Lacerda, R. B., Ling, W., Marsden, J., Monteiro, M., et al. (2008). Validation of the alcohol, smoking and substance involvement screening test (assist). *Addiction*, 103(6):1039–1047.

Ibar, C., Fortuna, F., Gonzalez, D., Jamardo, J., Jacobsen, D., Pugliese, L., Giraudo, L., Ceres, V., Mendoza, C., Repetto, E. M., et al. (2021). Evaluation of stress, burnout and hair cortisol levels in health workers at a university hospital during covid-19 pandemic. *Psychoneuroendocrinology*, 128:105213.

Ixart, G., Barbanel, G., Nouguier-Soule, J., and Assenmacher, I. (1991). A quantitative study of the pulsatile parameters of crh-41 secretion in unanesthetized free-moving rats. *Experimental brain research*, 87(1):153–158.

Jacobs, T. L., Shaver, P. R., Epel, E. S., Zanesco, A. P., Aichele, S. R., Bridwell, D. A., Rosenberg, E. L., King, B. G., MacLean, K. A., Sahdra, B. K., Kemeny, M. E., Ferrer, E., Wallace, B. A., and Saron, C. D. (2013). Self-

reported mindfulness and cortisol during a Shamatha meditation retreat. *Health Psychology*, 32(10):1104–1109.

Jasper, M. S. and Engeland, W. C. (1991). Synchronous ultradian rhythms in adrenocortical secretion detected by microdialysis in awake rats. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 261(5):R1257–R1268.

Jensen, T. K., Andersson, A.-M., Jørgensen, N., Andersen, A.-G., Carlsen, E., Skakkebæk, N. E., et al. (2004). Body mass index in relation to semen quality and reproductive hormonesamong 1,558 danish men. *Fertility and sterility*, 82(4):863–870.

Jensen, T. K., Hjollund, N. H. I., Henriksen, T. B., Scheike, T., Kolstad, H., Giwercman, A., Ernst, E., Bonde, J. P., Skakkebæk, N. E., et al. (1998). Does moderate alcohol consumption affect fertility? follow up study among couples planning first pregnancy. *Bmj*, 317(7157):505–510.

Jia, R., Ayling, K., Chalder, T., Massey, A., Broadbent, E., Coupland, C., and Vedhara, K. (2020). Mental health in the UK during the COVID-19 pandemic: cross-sectional analyses from a community cohort study. *BMJ open*, 10(9):e040620.

Jia, R., Ayling, K., Massey, A., Chalder, T., Gasteiger, N., Broadbent, E., Carol, C., and Vedhara, K. (2021). The prevalence, incidence, prognosis and risk factors for symptoms of depression and anxiety in a UK cohort during the COVID-19 pandemic (in press). *BJPsychOpen*.

Jia, R., Carlisle, S., and Vedhara, K. (2022). The association of lifestyle and mood with long-term levels of cortisol: A systematic review. *Cogent Psychology*, 9(1):2036487.

Jiang, S., Zhong, Y., Qiao, H., Di, B., Chen, J., and Su, M. (2022). Uplc-ms/ms method for the simultaneous quantification of caffeine and illicit psychoactive

drugs in hair using a single-step high-speed grinding extraction–insights into a cut-off value for caffeine abuse. *Journal of Pharmaceutical and Biomedical Analysis*, 209:114489.

Johnson, M. H. (2018). *Essential reproduction*. John Wiley & Sons.

Kalafatakis, K., Russell, G., Harmer, C., Munafo, M., Marchant, N., Wilson, A., Brooks, J., Durant, C., Thakrar, J., Murphy, P., et al. (2018). Ultradian rhythmicity of plasma cortisol is necessary for normal emotional and cognitive responses in man. *Proceedings of the National Academy of Sciences*, 115(17):E4091–E4100.

Kalliokoski, O., Jellestad, F. K., and Murison, R. (2019). A systematic review of studies utilizing hair glucocorticoids as a measure of stress suggests the marker is more appropriate for quantifying short-term stressors. *Scientific Reports*, 9(1):11997.

Kamphuis, E. I., Bhattacharya, S., van der Veen, F., Mol, B. W. J., Templeton, A., and for the Evidence Based IVF Group (2014). Are we overusing IVF? *BMJ*, 348(jan28 8):g252–g252.

Kapoor, A., Schultz-Darken, N., and Ziegler, T. E. (2018). Radiolabel validation of cortisol in the hair of rhesus monkeys. *Psychoneuroendocrinology*, 97:190–195.

Kapoor, D. and Jones, T. (2005). Smoking and hormones in health and endocrine disorders. *European journal of endocrinology*, 152(4):491–499.

Katz, D. and Meller, S. (2014). Can We Say What Diet Is Best for Health? *Annual Review of Public Health*, 35(1):83–103.

Keay, S., Harlow, C., Wood, P., Jenkins, J., and Cahill, D. (2002). Higher cortisol: cortisone ratios in the preovulatory follicle of completely unstimulated ivf cycles indicate oocytes with increased pregnancy potential. *Human reproduction*, 17(9):2410–2414.

Kemeny, M. E. (2003). The psychobiology of stress. *Current directions in psychological science*, 12(4):124–129.

Kennedy, R., Kingsland, C., Rutherford, A., Hamilton, M., and Ledger, W. (2006). Implementation of the NICE guideline - Recommendations from the British Fertility Society for national criteria for NHS funding of assisted conception. *Human Fertility*, 9(3):181–189.

Kickbusch, I. (1986). Life-styles and health. *Social science & medicine (1982)*, 22(2):117–124.

Kiesswetter, M., Marsoner, H., Luehwink, A., Fistarol, M., Mahlknecht, A., and Duschek, S. (2020). Impairments in life satisfaction in infertility: Associations with perceived stress, affectivity, partnership quality, social support and the desire to have a child. *Behavioral Medicine*, 46(2):130–141.

Kirschbaum, C. and Hellhammer, D. H. (1994). Salivary cortisol in psychoneuroendocrine research: recent developments and applications. *Psychoneuroendocrinology*, 19(4):313–333.

Kirschbaum, C., Tietze, A., Skoluda, N., and Dettenborn, L. (2009). Hair as a retrospective calendar of cortisol production—increased cortisol incorporation into hair in the third trimester of pregnancy. *Psychoneuroendocrinology*, 34(1):32–37.

Kline, T. (2005). *Psychological testing: A practical approach to design and evaluation*. Sage.

Klonoff-Cohen, H. (2005). Female and male lifestyle habits and IVF: what is known and unknown. *Human Reproduction Update*, 11(2):180–204.

Klonoff-Cohen, H., Lam-Kruglick, P., and Gonzalez, C. (2003). Effects of maternal and paternal alcohol consumption on the success rates of in vitro fertilization and gamete intrafallopian transfer. *Fertility and Sterility*, 79(2):330–339.

Kloss, J. D., Perlis, M. L., Zamzow, J. A., Culnan, E. J., and Gracia, C. R. (2015). Sleep, sleep disturbance, and fertility in women. *Sleep medicine reviews*, 22:78–87.

Knuth, U., Hull, M., and Jacobs, H. (1977). Amenorrhoea and loss of weight. *BJOG: An International Journal of Obstetrics & Gynaecology*, 84(11):801–807.

Konstantinidou, F., Budani, M. C., Sarra, A., Stuppia, L., Tiboni, G. M., and Gatta, V. (2021). Impact of cigarette smoking on the expression of oxidative stress-related genes in cumulus cells retrieved from healthy women undergoing ivf. *International journal of molecular sciences*, 22(23):13147.

Kopijka, V., Bekasova, O., and Ivanova, K. (2017). Parameters of peripheral blood in women at risk of ovarian hyperstimulation syndrome. *Biological Markers in Fundamental and Clinical Medicine (scientific journal)*, 1(4):58–60.

Kramer, M. S., Lydon, J., Seguin, L., Goulet, L., Kahn, S. R., McNamara, H., Genest, J., Dassa, C., Chen, M. F., Sharma, S., Meaney, M. J., Thomson, S., Van Uum, S., Koren, G., Dahhou, M., Lamoureux, J., and Platt, R. W. (2009). Stress Pathways to Spontaneous Preterm Birth: The Role of Stressors, Psychological Distress, and Stress Hormones. *American Journal of Epidemiology*, 169(11):1319–1326.

Kumar, P., Sait, S. F., Sharma, A., and Kumar, M. (2011). Ovarian hyperstimulation syndrome. *Journal of human reproductive sciences*, 4(2):70.

Kupka, M. S., Dorn, C., Montag, M., Felberbaum, R., Van der Ven, H., Kulczycki, A., and Friese, K. (2004). Previous miscarriages influence ivf and intracytoplasmatic sperm injection pregnancy outcome. *Reproductive BioMedicine Online*, 8(3):349–357.

Lablanche, O., Salle, B., Perie, M.-A., Labrune, E., Langlois-Jacques, C., and Fraison, E. (2022). Psychological effect of covid-19 pandemic among women

undergoing infertility care, a french cohort–psycovart psychological effect of covid-19: Psycovart. *Journal of Gynecology Obstetrics and Human Reproduction*, 51(1):102251.

Lack, L. C., Gradisar, M., Van Someren, E. J., Wright, H. R., and Lushington, K. (2008). The relationship between insomnia and body temperatures. *Sleep medicine reviews*, 12(4):307–317.

Lakatos, E., Szigeti, J. F., Ujma, P. P., Sexty, R., and Balog, P. (2017). Anxiety and depression among infertile women: a cross-sectional survey from Hungary. *BMC Women's Health*, 17(1):48.

Lancastle, D. and Boivin, J. (2008). A feasibility study of a brief coping intervention (prci) for the waiting period before a pregnancy test during fertility treatment. *Human Reproduction*, 23(10):2299–2307.

Lanfear, J. H., Voegel, C. D., Binz, T. M., and Paul, R. A. (2020). Hair cortisol measurement in older adults: influence of demographic and physiological factors and correlation with perceived stress. *Steroids*, 163:108712.

Larsen, S. C., Fahrenkrug, J., Olsen, N. J., and Heitmann, B. L. (2016). Association between Hair Cortisol Concentration and Adiposity Measures among Children and Parents from the "Healthy Start" Study. *PLOS ONE*, 11(9):e0163639.

Lazarus, R. S. (1966). Psychological stress and the coping process.

Lazarus, R. S. (1975). A cognitively oriented psychologist looks at biofeedback. *American psychologist*, 30(5):553.

Lazarus, R. S. (1984). *Stress, appraisal, and coping / Richard S. Lazarus, Susan Folkman.* Springer.

Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American psychologist*, 46(8):819.

Lazarus, R. S. (1993). From psychological stress to the emotions: A history of changing outlooks. *Annual review of psychology*, 44(1):1–22.

Lazarus, R. S. and Cohen, J. B. (1977). Environmental stress. In *Human behavior and environment*, pages 89–127. Springer.

Lazarus, R. S. and Folkman, S. (1987). Transactional theory and research on emotions and coping. *European Journal of personality*, 1(3):141–169.

Leivadi, S., Hernandez-Reif, M., Field, T., O'Rourke, M., D'Arienzo, S., Lewis, D., Pino, N. d., Schanberg, S., and Kuhn, C. (1999). Massage therapy and relaxation effects on university dance students. *Journal of dance medicine & science*, 3(3):108–112.

Les Dees, W., Hiney, J. K., and Srivastava, V. (1998). Alcohol's effects on female puberty: The role of insulin-like growth factor 1. *Alcohol health and research world*, 22(3):165.

Levine, A., Zagoory-Sharon, O., Feldman, R., Lewis, J. G., and Weller, A. (2007). Measuring cortisol in human psychobiological studies. *Physiology & Behavior*, 90(1):43–53.

Lewis, J. G., Bagley, C. J., Elder, P. A., Bachmann, A. W., and Torpy, D. J. (2005). Plasma free cortisol fraction reflects levels of functioning corticosteroid-binding globulin. *Clinica Chimica Acta*, page 6.

Li, X., Guo, Y., and Zhang, Y. (2006). Comment on" the reliability and validity of the effort-reward imbalance-the chinese version". *Zhonghua liu xing bing xue za zhi= Zhonghua liuxingbingxue zazhi*, 27(1):25–28.

Li, Y., Lin, H., Li, Y., and Cao, J. (2011). Association between socio-psycho-behavioral factors and male semen quality: systematic review and meta-analyses. *Fertility and Sterility*, 95(1):116–123.

Lightman, S. (2016). Rhythms within rhythms: the importance of oscillations for glucocorticoid hormones. *A time for metabolism and hormones*, pages 87–99.

Lightman, S. L. and Conway-Campbell, B. L. (2010). The crucial role of pulsatile activity of the HPA axis for continuous dynamic equilibration. *Nature Reviews Neuroscience*, 11(10):710–718.

Limm, H., Gundel, H., Heinmuller, M., Marten-Mittag, B., Nater, U. M., Siegrist, J., and Angerer, P. (2011). Stress management interventions in the workplace improve stress reactivity: a randomised controlled trial. *Occupational and Environmental Medicine*, 68(2):126–133.

Lin, Y.-H., Chueh, K.-H., and Lin, J.-L. (2016). Somatic symptoms, sleep disturbance and psychological distress among women undergoing oocyte pick-up and in vitro fertilisation–embryo transfer. *Journal of clinical nursing*, 25(11-12):1748–1756.

London, S., Willett, W., Longcope, C., and McKinlay, S. (1991). Alcohol and other dietary factors in relation to serum hormone concentrations in women at climacteric. *The American journal of clinical nutrition*, 53(1):166–171.

Lopresti, A. L., Hood, S. D., and Drummond, P. D. (2013). A review of lifestyle factors that contribute to important pathways associated with major depression: Diet, sleep and exercise. *Journal of Affective Disorders*, 148(1):12–27.

Louis, J. F., Thoma, M. E., Sørensen, D. N., Mclain, A. C., King, R. B., Sundaram, R., Keiding, N., and Buck Louis, G. M. (2013). The prevalence of couple infertility in the United States from a male perspective: Evidence from a nationally representative sample. *Andrology*, 1(5):741–748.

Lovallo, W. R., Al'Absi, M., Blick, K., Whitsett, T. L., and Wilson, M. F. (1996). Stress-like adrenocorticotropin responses to caffeine in young healthy men. *Pharmacology Biochemistry and Behavior*, 55(3):365–369.

Lovallo, W. R., Whitsett, T. L., al'Absi, M., Sung, B. H., Vincent, A. S., and Wilson, M. F. (2005). Caffeine Stimulation of Cortisol Secretion Across the Waking Hours in Relation to Caffeine Intake Levels:. *Psychosomatic Medicine*, 67(5):734–739.

Lovibond, P. F. and Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the depression anxiety stress scales (dass) with the beck depression and anxiety inventories. *Behaviour research and therapy*, 33(3):335–343.

Lucero, J., Harlow, B. L., Barbieri, R. L., Sluss, P., and Cramer, D. W. (2001). Early follicular phase hormone levels in relation to patterns of alcohol, tobacco, and coffee use. *Fertility and sterility*, 76(4):723–729.

Lupien, S. J., McEwen, B. S., Gunnar, M. R., and Heim, C. (2009). Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nature reviews neuroscience*, 10(6):434–445.

Lynch, S., Gander, M.-L., Kohls, N., Kudielka, B., and Walach, H. (2011). Mindfulness-based Coping with University Life: A Non-randomized Wait-list-controlled Pilot Evaluation: Mindfulness-based Coping with University Life. *Stress and Health*, 27(5):365–375.

Lyngsø, J., Ramlau-Hansen, C. H., Bay, B., Ingerslev, H. J., Hulman, A., and Kesmodel, U. S. (2017). Association between coffee or caffeine consumption and fecundity and fertility: a systematic review and dose–response meta-analysis. *Clinical epidemiology*, 9:699.

MacDougall, M. J., Tan, S. L., and Jacobs, H. S. (1992). In vitro fertilization and the ovarian hyperstimulation syndrome. *Human Reproduction*, 7(5):597–600.

Manenschijn, L., van Kruysbergen, R. G. P. M., de Jong, F. H., Koper, J. W., and van Rossum, E. F. C. (2011). Shift Work at Young Age Is Associated

with Elevated Long-Term Cortisol Levels and Body Mass Index. *The Journal of Clinical Endocrinology & Metabolism*, 96(11):E1862–E1865.

Mangine, G. T., Hoffman, J. R., Gonzalez, A. M., Townsend, J. R., Jajtner, A. R., Beyer, K. S., Boone, C. H., Wang, R., LaMonica, M. B., Fukuda, D. H., Witta, E. L., Ratamess, N. A., and Stout, J. R. (2016). EXERCISE-INDUCED HORMONE ELEVATIONS ARE RELATED TO MUSCLE GROWTH. page 25.

Marcus, M. T., Fine, P. M., Moeller, F. G., Khan, M. M., Pitts, K., Swank, P. R., and Liehr, P. (2003). Change in Stress Levels Following Mindfulness-based Stress Reduction in a Therapeutic Community:. *Addictive Disorders & Their Treatment*, 2(3):63–68.

Marcus, S. F. and Brinsden, P. R. (1995). Analysis of the incidence and risk factors associated with ectopic pregnancy following in-vitro fertilization and embryo transfer. *Human Reproduction*, 10(1):199–203.

Marmer, C. and Weiss, D. (1997). The impact of event scale-revised. *Assessing psychological trauma and PTSD. Guilford, New York*, pages 399–411.

Marteau, T. M. and Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, 31(3):301–306.

Mascarenhas, M. N., Flaxman, S. R., Boerma, T., Vanderpoel, S., and Stevens, G. A. (2012). National, Regional, and Global Trends in Infertility Prevalence Since 1990: A Systematic Analysis of 277 Health Surveys. *PLoS Medicine*, 9(12):e1001356.

Massey, A. J., Campbell, B., Raine-Fenning, N., Aujla, N., and Vedhara, K. (2014). The association of physiological cortisol and IVF treatment outcomes: a systematic review. *Reproductive medicine and biology*, 13(4):161–176. Publisher: Springer.

Massey, A. J., Campbell, B. K., Raine-Fenning, N., Pincott-Allen, C., Perry, J., and Vedhara, K. (2016). Relationship between hair and salivary cortisol and pregnancy in women undergoing IVF. *Psychoneuroendocrinology*, 74:397–405. Publisher: Elsevier.

Mastorakos, G., Pavlatou, M., Diamanti-Kandarakis, E., and Chrousos, G. P. (2005). Exercise and the stress system. *Hormones (Athens)*, 4(2):73–89.

Matheny Jr, A. P., Wachs, T. D., Ludwig, J. L., and Phillips, K. (1995). Bringing order out of chaos: Psychometric characteristics of the confusion, hubbub, and order scale. *Journal of applied developmental psychology*, 16(3):429–444.

Matthiesen, S., Frederiksen, Y., Ingerslev, H., and Zachariae, R. (2011). Stress, distress and outcome of assisted reproductive technology (ART): a meta-analysis. *Human Reproduction*, 26(10):2763–2776.

McDonald, S. D., Han, Z., Mulla, S., Murphy, K. E., Beyene, J., and Ohlsson, A. (2009). Preterm birth and low birth weight among in vitro fertilization singletons: A systematic review and meta-analyses. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 146(2):138–148.

McEwen, B. S. (1998). Stress, adaptation, and disease: Allostasis and allostatic load. *Annals of the New York academy of sciences*, 840(1):33–44.

McEwen, B. S. (2004). Protection and damage from acute and chronic stress: allostasis and allostatic overload and relevance to the pathophysiology of psychiatric disorders. *Annals of the New York Academy of Sciences*, 1032(1):1–7.

McEwen, B. S. (2005). Stressed or stressed out: what is the difference? *Journal of Psychiatry and Neuroscience*, 30(5):315.

McEwen, B. S. (2007). Physiology and neurobiology of stress and adaptation: central role of the brain. *Physiological reviews*, 87(3):873–904.

McEwen, B. S. and Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Hormones and behavior*, 43(1):2–15.

McLernon, D. J., Harrild, K., Bergh, C., Davies, M. J., De Neubourg, D., Dumoulin, J., Gerris, J., Kremer, J. A., Martikainen, H., Mol, B., et al. (2010). Clinical effectiveness of elective single versus double embryo transfer: meta-analysis of individual patient data from randomised trials. *Bmj*, 341.

McNair, D. and Heuchert, J. (2005). Profiles of mood states–technical update. toronto, on: Multi-health systems.

Meeusen, R., Duclos, M., Gleeson, M., Rietjens, G., Steinacker, J., and Urhausen, A. (2006). Prevention, diagnosis and treatment of the overtraining syndrome: Ecss position statement 'task force'. *European Journal of Sport Science*, 6(01):1–14.

Meijer, J. L., Bockting, C. L., Stolk, R. P., Kotov, R., Ormel, J., and Burger, H. (2014). Associations of life events during pregnancy with longitudinal change in symptoms of antenatal anxiety and depression. *Midwifery*, 30(5):526–531.

Mejia, E. G. d. and Ramirez-Mares, M. V. (2014). Impact of caffeine and coffee on our health. *Trends in Endocrinology & Metabolism*, 25(10):489–492.

Mendel, C. M., Kuhn, R. W., Weisiger, R. A., Cavalieri, R. R., Siiteri, P. K., Cunha, G. R., and Murai, J. T. (1989). Uptake of cortisol by the perfused rat liver: validity of the free hormone hypothesis applied to cortisol. *Endocrinology*, 124(1):468–476.

Meniru, G. I. (2001). *Cambridge guide to infertility management and assisted reproduction*. Cambridge University Press.

Mershon, J. L., Sehlhorst, C., Rebar, R. W., and Liu, J. H. (1992). Evidence of a corticotropin-releasing hormone pulse generator in the macaque hypothalamus. *Endocrinology*, 130(5):2991–2996.

Meyer, J., Novak, M., Hamel, A., and Rosenberg, K. (2014). Extraction and analysis of cortisol from human and monkey hair. *Journal of visualized experiments: JoVE*, (83).

Meyer, J. S. and Novak, M. A. (2012). Minireview: Hair Cortisol: A Novel Biomarker of Hypothalamic-Pituitary-Adrenocortical Activity. *Endocrinology*, 153(9):4120–4127.

Miller, C. B., Kyle, S. D., Gordon, C. J., Espie, C. A., Grunstein, R. R., Mullins, A. E., Postnova, S., and Bartlett, D. J. (2015). Physiological Markers of Arousal Change with Psychological Treatment for Insomnia: A Preliminary Investigation. *PLOS ONE*, 10(12):e0145317.

Miller, N., Herzberger, E. H., Pasternak, Y., Klement, A. H., Shavit, T., Yaniv, R. T., Ghetler, Y., Neumark, E., Eisenberg, M. M., Berkovitz, A., et al. (2019). Does stress affect IVF outcomes? A prospective study of physiological and psychological stress in women undergoing IVF. *Reproductive biomedicine online*, 39(1):93–101.

Moal, V., Mathieu, E., Reynier, P., Malthièry, Y., and Gallois, Y. (2007). Low serum testosterone assayed by liquid chromatography-tandem mass spectrometry. comparison with five immunoassay techniques. *Clinica chimica acta*, 386(1-2):12–19.

Moore, K. L., Persaud, T. V. N., and Torchia, M. G. (2015). *Before We are Born E-Book: Essentials of Embryology and Birth Defects with STUDENT CONSULT Online Access*. Elsevier Health Sciences.

Murawski, B., Wade, L., Plotnikoff, R. C., Lubans, D. R., and Duncan, M. J. (2018). A systematic review and meta-analysis of cognitive and behavioral interventions to improve sleep health in adults without sleep disorders. *Sleep medicine reviews*, 40:160–169.

Muzni, K., Groeger, J. A., Dijk, D.-J., and Lazar, A. S. (2021). Self-reported sleep quality is more closely associated with mental and physical health than chronotype and sleep duration in young adults: A multi-instrument analysis. *Journal of Sleep Research*, 30(1):e13152.

National Collaborating Centre for Women's and Children's Health (UK) (2013). Fertility: assessment and treatment for people with fertility problems.

Nelson, S. M. and Lawlor, D. A. (2011). Predicting Live Birth, Preterm Delivery, and Low Birth Weight in Infants Born from In Vitro Fertilisation: A Prospective Study of 144,018 Treatment Cycles. *PLoS Medicine*, 8(1):e1000386.

Nepomnaschy, P. A., Welch, K. B., McConnell, D. S., Low, B. S., Strassmann, B. I., and England, B. G. (2006). Cortisol levels and very early pregnancy loss in humans. *Proceedings of the National Academy of Sciences*, 103(10):3938–3942.

Nicoloro-SantaBarbara, J., Busso, C., Moyer, A., and Lobel, M. (2018). Just relax and you'll get pregnant? meta-analysis examining women's emotional distress and the outcome of assisted reproductive technology. *Social Science & Medicine*, 213:54–62.

Oba, S., Nagata, C., Nakamura, K., Fujii, K., Kawachi, T., Takatsuka, N., and Shimizu, H. (2010). Consumption of coffee, green tea, oolong tea, black tea, chocolate snacks and the caffeine content in relation to risk of diabetes in japanese men and women. *British Journal of Nutrition*, 103(3):453–459.

O'Brien, K. M., Tronick, E. Z., and Moore, C. L. (2013). Relationship between Hair Cortisol and Perceived Chronic Stress in a Diverse Sample. *Stress and Health*, 29(4):337–344.

Ockhuijsen, H., van den Hoogen, A., Eijkemans, M., Macklon, N., and Boivin, J. (2014). The impact of a self-administered coping intervention on emotional

well-being in women awaiting the outcome of ivf treatment: a randomized controlled trial. *Human Reproduction*, 29(7):1459–1470.

O'Connor, D. B., Thayer, J. F., and Vedhara, K. (2021). Stress and health: A review of psychobiological processes. *Annual review of psychology*, 72:663–688.

Ogden, J. (2019). *Health psychology / Jane Ogden.* McGraw Hill, sixth edition. edition.

Olsen, J. A. (2017). *Determinants of (inequalities in) health: an overview*, volume 1. Oxford University Press.

Organization, W. H. (2019). *Global status report on alcohol and health 2018.* World Health Organization.

Organization, W. H. et al. (1986). *Targets for health for all: Targets in support of the European regional strategy for health for all.* World Health Organization. Regional Office for Europe.

Organization, W. H. et al. (2009). *Global health risks: mortality and burden of disease attributable to selected major risks.* World Health Organization.

Organization, W. H. et al. (2020). Mental health and psychosocial considerations during the covid-19 outbreak, 18 march 2020. Technical report, World Health Organization.

Owen, N., Spathonis, K., and Leslie, E. (2007). Physical activity and health. *Cambridge handbook of psychology, health and medicine*, pages 155–161.

O'Leary, K., O'Neill, S., and Dockray, S. (2016). A systematic review of the effects of mindfulness interventions on cortisol. *Journal of Health Psychology*, 21(9):2108–2121.

Palermo, G., Joris, H., Devroey, P., and Lancet, A. V. S. (1992). Pregnancies after intracytoplasmic injection of single spermatozoon into an oocyte

10.1016/0140-6736(92)92425-F : The Lancet — ScienceDirect.com. *Elsevier*, pages 17–18.

Palomba, S., Falbo, A., Valli, B., Morini, D., Villani, M. T., Nicoli, A., and La Sala, G. B. (2014). Physical activity before ivf and icsi cycles in infertile obese women: an observational cohort study. *Reproductive biomedicine online*, 29(1):72–79.

Pandian, Z., Gibreel, A., and Bhattacharya, S. (2015). In vitro fertilisation for unexplained subfertility.

Papacosta, E., Gleeson, M., and Nassis, G. P. (2013). Salivary Hormones, IgA, and Performance During Intense Training and Tapering in Judo Athletes:. *Journal of Strength and Conditioning Research*, 27(9):2569–2580.

Parker, V. J. and Douglas, A. J. (2010). Stress in early pregnancy: maternal neuro-endocrine-immune responses and effects. *Journal of reproductive immunology*, 85(1):86–92.

Pasch, L. A., Gregorich, S. E., Katz, P. K., Millstein, S. G., Nachtigall, R. D., Bleil, M. E., and Adler, N. E. (2012). Psychological distress and in vitro fertilization outcome. *Fertility and sterility*, 98(2):459–464.

Patel, A., Sharma, P. S. V. N., Narayan, P., Binu, V. S., Dinesh, N., and Pai, P. J. (2016). Prevalence and predictors of infertility-specific stress in women diagnosed with primary infertility: A clinic-based study. *Journal of Human Reproductive Sciences*, 9(1):28–34.

Penelo, E., Viladrich, C., and Domènech, J. M. (2012). Adolescents' perceptions of parental behavior: psychometric properties of the short egna minnen beträffande uppfostran–adolescent version (s-embu-a) in a clinical sample. *Comprehensive Psychiatry*, 53(1):87–94.

Philipsen, M. T., Knudsen, U. B., Zachariae, R., Ingerslev, H. J., Hvidt, J.

E. M., and Frederiksen, Y. (2021). Sleep, psychological distress, and clinical pregnancy outcome in women and their partners undergoing in vitro or intracytoplasmic sperm injection fertility treatment. *Sleep Health*.

Poslusna, K., Ruprich, J., de Vries, J. H., Jakubikova, M., and van't Veer, P. (2009). Misreporting of energy and micronutrient intake estimated by food records and 24 hour recalls, control and adjustment methods in practice. *British Journal of Nutrition*, 101(S2):S73–S85.

Pruessner, J. C., Kirschbaum, C., Meinlschmid, G., and Hellhammer, D. H. (2003). Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology*, 28(7):916–931.

Purnell, J. Q., Brandon, D. D., Isabelle, L. M., Loriaux, D. L., and Samuels, M. H. (2004). Association of 24-hour cortisol production rates, cortisol-binding globulin, and plasma-free cortisol levels with body composition, leptin levels, and aging in adult men and women. *The Journal of Clinical Endocrinology & Metabolism*, 89(1):281–287.

Qi, X., Zhang, J., Liu, Y., Ji, S., Chen, Z., Sluiter, J. K., and Deng, H. (2014). Relationship between effort–reward imbalance and hair cortisol concentration in female kindergarten teachers. *Journal of Psychosomatic Research*, 76(4):329–332.

Qu, F., Wu, Y., Zhu, Y.-H., Barry, J., Ding, T., Baio, G., Muscat, R., Todd, B. K., Wang, F.-F., and Hardiman, P. J. (2017). The association between psychological stress and miscarriage: a systematic review and meta-analysis. *Scientific reports*, 7(1):1–8.

Radloff, L. S. (1977). The ces-d scale: A self-report depression scale for research in the general population. *Applied psychological measurement*, 1(3):385–401.

Rajcani, J., Vytykacova, S., Solarikova, P., and Brezina, I. (2021). Stress and hair cortisol concentrations in nurses during the first wave of the covid-19 pandemic. *Psychoneuroendocrinology*, 129:105245.

Rao, M., Zeng, Z., and Tang, L. (2018). Maternal physical activity before IVF/ICSI cycles improves clinical pregnancy rate and live birth rate: a systematic review and meta-analysis. *Reproductive Biology and Endocrinology*, 16(1):11.

Rector, J. L., Tay, L., Wiese, C. W., and Friedman, E. M. (2019). Relative sensitivity of cortisol indices to psychosocial and physical health factors. *PLOS ONE*, 14(4):e0213513.

Redshaw, M., Hockley, C., and Davidson, L. L. (2007). A qualitative study of the experience of treatment for infertility among women who successfully became pregnant. *Human Reproduction*, 22(1):295–304.

Reitsma, M. B., Fullman, N., Ng, M., Salama, J. S., Abajobir, A., Abate, K. H., Abbafati, C., Abera, S. F., Abraham, B., Abyu, G. Y., et al. (2017). Smoking prevalence and attributable disease burden in 195 countries and territories, 1990–2015: a systematic analysis from the global burden of disease study 2015. *The Lancet*, 389(10082):1885–1906.

Rich-Edwards, J. W., Goldman, M. B., Willett, W. C., Hunter, D. J., Stampfer, M. J., Colditz, G. A., and Manson, J. E. (1994). Adolescent body mass index and infertility caused by ovulatory disorder. *American journal of obstetrics and gynecology*, 171(1):171–177.

Riemann, D., Spiegelhalder, K., Feige, B., Voderholzer, U., Berger, M., Perlis, M., and Nissen, C. (2010). The hyperarousal model of insomnia: a review of the concept and its evidence. *Sleep medicine reviews*, 14(1):19–31.

Rittenberg, V., Seshadri, S., Sunkara, S. K., Sobaleva, S., Oteng-Ntim, E., and El-Toukhy, T. (2011). Effect of body mass index on ivf treatment outcome:

an updated systematic review and meta-analysis. *Reproductive biomedicine online*, 23(4):421–439.

Roberts, R. E. and Vernon, S. W. (1983). The center for epidemiological studies depression scale: its use in a community sample. *The American journal of psychiatry*.

Roberts, S., Hirst, W., Brison, D., and Vail, A. (2010). Embryo and uterine influences on IVF outcomes: an analysis of a UK multi-centre cohort. *Human Reproduction*, 25(11):2792–2802.

Rossi, B. V., Berry, K. F., Hornstein, M. D., Cramer, D. W., Ehrlich, S., and Missmer, S. A. (2011). Effect of Alcohol Consumption on In Vitro Fertilization. *Obstetrics & Gynecology*, 117(1):136–142.

Rossi, B. V., Bressler, L. H., Correia, K. F., Lipskind, S., Hornstein, M. D., and Missmer, S. A. (2016). Lifestyle and in vitro fertilization: what do patients believe? *Fertility Research and Practice*, 2(1):11.

Rothe, N., Steffen, J., Penz, M., Kirschbaum, C., and Walther, A. (2020). Examination of peripheral basal and reactive cortisol levels in major depressive disorder and the burnout syndrome: A systematic review. *Neuroscience & Biobehavioral Reviews*, 114:232–270.

Russell, E., Kirschbaum, C., Laudenslager, M. L., Stalder, T., de Rijke, Y., van Rossum, E. F., Van Uum, S., and Koren, G. (2015). Toward standardization of hair cortisol measurement: results of the first international interlaboratory round robin. *Therapeutic drug monitoring*, 37(1):71–75.

Russell, E., Koren, G., Rieder, M., and Van Uum, S. (2012). Hair cortisol as a biological marker of chronic stress: Current status, future directions and unanswered questions. *Psychoneuroendocrinology*, 37(5):589–601.

Sannes, T. S., Dolan, E., Albano, D., Ceballos, R. M., and McGregor, B. A. (2015). Stress management reduces intraindividual cortisol variability, while

not impacting other measures of cortisol rhythm, in a group of women at risk for breast cancer. *Journal of Psychosomatic Research*, 79(5):412–419.

Sarabdjitsingh, R. A., Conway-Campbell, B. L., Leggett, J. D., Waite, E. J., Meijer, O. C., De Kloet, E., and Lightman, S. L. (2010). Stress responsiveness varies over the ultradian glucocorticoid cycle in a brain-region-specific manner. *Endocrinology*, 151(11):5369–5379.

Sarnyai, Z., Veldhuis, J. D., Mello, N. K., Mendelson, J., Erös-Sarnyai, M., Mercer, G., Gelles, H., and Kelly, M. (1995). The concordance of pulsatile ultradian release of adrenocorticotropin and cortisol in male rhesus monkeys. *The Journal of Clinical Endocrinology & Metabolism*, 80(1):54–59.

Saunders, J. B., Aasland, O. G., Babor, T. F., De la Fuente, J. R., and Grant, M. (1993). Development of the alcohol use disorders identification test (audit): Who collaborative project on early detection of persons with harmful alcohol consumption-ii. *Addiction*, 88(6):791–804.

Sauvé, B., Koren, G., Walsh, G., Tokmakejian, S., and Van Uum, S. H. (2007). Measurement of cortisol in human Hair as a biomarker of systemic exposure. *Clinical & Investigative Medicine*, 30(5):183.

Saxbe, D. E. (2008). A field (researcher's) guide to cortisol: tracking hpa axis functioning in everyday life. *Health psychology review*, 2(2):163–190.

Sazonova, A., Källen, K., Thurin-Kjellberg, A., Wennerholm, U.-B., and Bergh, C. (2013). Neonatal and maternal outcomes comparing women undergoing two in vitro fertilization (IVF) singleton pregnancies and women undergoing one IVF twin pregnancy. *Fertility and Sterility*, 99(3):731–737.

Schalinski, I., Elbert, T., Steudte-Schmiedgen, S., and Kirschbaum, C. (2015). The Cortisol Paradox of Trauma-Related Disorders: Lower Phasic Responses but Higher Tonic Levels of Cortisol Are Associated with Sexual Abuse in Childhood. *PLOS ONE*, 10(8):e0136921.

Schenker, J. G., Meirow, D., and Schenker, E. (1992). Stress and human reproduction. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 45(1):1–8.

Schieve, L. A., Meikle, S. F., Ferre, C., Peterson, H. B., Jeng, G., and Wilcox, L. S. (2002). Low and very low birth weight in infants conceived with use of assisted reproductive technology. *New England Journal of Medicine*, 346(10):731–737.

Schlarb, A. (2015). Sleep and Health. In *International Encyclopedia of the Social & Behavioral Sciences*, pages 59–62. Elsevier.

Schlotz, W. (2008). Tics: Trierer inventar zum chronischen stress. In *Diagnostische Verfahren in der Rehabilitation-Verfahren zum krankheitsübergreifenden Einsatz und für ausgewählte Indikationsgebiete*, pages 139–143. Hogrefe.

Seli, E., Aydin, A., and ProQuest (2011). *Infertility [electronic resource] / edited by Emre Seli.* Gynecology in practice. Wiley-Blackwell, Chichester, West Sussex, UK.

Selye, H. (1956). *The stress of life.*

Selye, H. et al. (1980). *Selye's guide to stress research.* Van Nostrand Reinhold.

Sermondade, N., Faure, C., Fezeu, L., Shayeb, A. G., Bonde, J. P., Jensen, T. K., Van Wely, M., Cao, J., Martini, A. C., Eskandar, M., Chavarro, J. E., Koloszar, S., Twigt, J. M., Ramlau-Hansen, C. H., Borges, E., Lotti, F., Steegers-Theunissen, R. P. M., Zorn, B., Polotsky, A. J., La Vignera, S., Eskenazi, B., Tremellen, K., Magnusdottir, E. V., Fejes, I., Hercberg, S., Levy, R., and Czernichow, S. (2013). BMI in relation to sperm count: an updated systematic review and collaborative meta-analysis. *Human Reproduction Update*, 19(3):221–231.

Siegrist, J. (1996). Adverse health effects of high-effort/low-reward conditions. *Journal of occupational health psychology*, 1(1):27.

Sinha, R. (2011). Effects of Adrenal Sensitivity, Stress- and Cue-Induced Craving, and Anxiety on Subsequent Alcohol Relapse and Treatment Outcomes. *Archives of General Psychiatry*, 68(9):942.

Smeenk, J., Verhaak, C., Vingerhoets, A., Sweep, C., Merkus, J., Willemsen, S., Van Minnen, A., Straatman, H., and Braat, D. (2005). Stress and outcome success in ivf: the role of self-reports and endocrine variables. *Human Reproduction*, 20(4):991–996.

Smith, M., Keay, S., Margo, F., Harlow, C., Wood, P., Cahill, D., and Hull, M. (2002). Total cortisol levels are reduced in the periovulatory follicle of infertile women with minimal–mild endometriosis. *American Journal of Reproductive Immunology*, 47(1):52–56.

Spielberger, C. D., Gorsuch, R., Lushene, R., Vagg, P., and Jacobs, G. (1983). State-trait anxiety inventory. palo alto. *CA: Mind Garden*.

Stalder, T. and Kirschbaum, C. (2012). Analysis of cortisol in hair–state of the art and future directions. *Brain, behavior, and immunity*, 26(7):1019–1029.

Stalder, T., Kirschbaum, C., Heinze, K., Steudte, S., Foley, P., Tietze, A., and Dettenborn, L. (2010). Use of hair cortisol analysis to detect hypercortisolism during active drinking phases in alcohol-dependent individuals. *Biological Psychology*, 85(3):357–360.

Stalder, T., Steudte-Schmiedgen, S., Alexander, N., Klucken, T., Vater, A., Wichmann, S., Kirschbaum, C., and Miller, R. (2017). Stress-related and basic determinants of hair cortisol in humans: A meta-analysis. *Psychoneuroendocrinology*, 77:261–274.

Staufenbiel, S. M., Penninx, B. W., Spijker, A. T., Elzinga, B. M., and van Rossum, E. F. (2013). Hair cortisol, stress exposure, and mental health in humans: A systematic review. *Psychoneuroendocrinology*, 38(8):1220–1235.

Stoppelbein, L. and Greening, L. (2015). A longitudinal study of the role of cortisol in posttraumatic stress disorder symptom clusters. *Anxiety, Stress, & Coping*, 28(1):17–30.

Sumra, M. K. and Schillaci, M. A. (2015). Stress and the Multiple-Role Woman: Taking a Closer Look at the "Superwoman". *PLOS ONE*, 10(3):e0120952.

Sunkara, S. K., La Marca, A., Seed, P. T., and Khalaf, Y. (2015). Increased risk of preterm birth and low birthweight with very high number of oocytes following IVF: an analysis of 65 868 singleton live birth outcomes. *Human Reproduction*, 30(6):1473–1480.

Taieb, J., Benattar, C., Birr, A. S., and Lindenbaum, A. (2002). Limitations of steroid determination by direct immunoassay. *Clinical Chemistry*, 48(3):583–585.

Tam, C. S., Frost, E. A., Xie, W., Rood, J., Ravussin, E., and Redman, L. M. (2014). No Effect of Caloric Restriction on Salivary Cortisol Levels in Overweight Men and Women. *Metabolism*, 63(2):194–198.

Tan, S.-L., Balen, A., El Hussein, E., Campbell, S., and Jacobs, H. S. (1992). The administration of glucocorticoids for the prevention of ovarian hyperstimulation syndrome in in vitro fertilization: a prospective randomized study. *Fertility and sterility*, 58(2):378–383.

Tanskanen, M. M., Kyröläinen, H., Uusitalo, A. L., Huovinen, J., Nissilä, J., Kinnunen, H., Atalay, M., and Häkkinen, K. (2011). Serum Sex Hormone–Binding Globulin and Cortisol Concentrations are Associated With Overreaching During Strenuous Military Training:. *Journal of Strength and Conditioning Research*, 25(3):787–797.

Tavakol, M. and Dennick, R. (2011). Making sense of cronbach's alpha. *International journal of medical education*, 2:53.

Templeton, A. A. and Penney, G. C. (1982). The incidence, characteristics, and prognosis of patients whose infertility is unexplained. *Fertility and Sterility*, 37(2):175–182.

Thoma, M. E., McLain, A. C., Louis, J. F., King, R. B., Trumble, A. C., Sundaram, R., and Buck Louis, G. M. (2013). Prevalence of infertility in the United States as estimated by the current duration approach and a traditional constructed approach. *Fertility and Sterility*, 99(5).

Thonneau, P., Marchand, S., Tallec, A., Ferial, M.-L., Ducot, B., Lansac, J., Lopes, P., Tabaste, J.-M., and Spira, A. (1991). Incidence and main causes of infertility in a resident population (1 850 000) of three french regions (1988–1989). *Human reproduction*, 6(6):811–816.

Timon, R., Corvillo, M., Brazo, J., Robles, M. C., and Maynar, M. (2013). Strength training effects on urinary steroid profile across the menstrual cycle in healthy women. *European Journal of Applied Physiology*, 113(6):1469–1475.

Tomic, V. and Tomic, J. (2011). Neonatal outcome of IVF singletons versus naturally conceived in women aged 35 years and over. *Archives of Gynecology and Obstetrics*, 284(6):1411–1416.

Torres, S. J., Nowson, C. A., and Worsley, A. (2008). Dietary electrolytes are related to mood. *British Journal of Nutrition*, 100(5):1038–1045.

Tsigos, C. and Chrousos, G. P. (2002). Hypothalamic–pituitary–adrenal axis, neuroendocrine factors and stress. *Journal of psychosomatic research*, 53(4):865–871.

Tsiouli, E., Pavlopoulos, V., Alexopoulos, E. C., Chrousos, G., and Darviri, C. (2014). Short-Term Impact of a Stress Management and Health Promotion Program on Perceived Stress, Parental Stress, Health Locus of Control, and

Cortisol Levels in Parents of Children and Adolescents With Diabetes Type 1: A Pilot Randomized Controlled Trial. *EXPLORE*, 10(2):88–98.

Turpeinen, U. and Hämäläinen, E. (2013). Determination of cortisol in serum, saliva and urine. *Best Practice & Research Clinical Endocrinology & Metabolism*, 27(6):795–801.

Ullmann, E., Licinio, J., Barthel, A., Petrowski, K., Stalder, T., Bornstein, S. R., and Kirschbaum, C. (2017). Persistent LHPA Activation in German Individuals Raised in an Overprotective Parental Behavior. *Scientific Reports*, 7(1):2778.

van Dalfsen, J. H. and Markus, C. R. (2018). The influence of sleep on human hypothalamic–pituitary–adrenal (hpa) axis reactivity: A systematic review. *Sleep medicine reviews*, 39:187–194.

Van der Spuy, Z., Steer, P., McCusker, M., Steele, S., and Jacobs, H. (1988). Outcome of pregnancy in underweight women after spontaneous and induced ovulation. *Br Med J (Clin Res Ed)*, 296(6627):962–965.

van Holland, B. J., Frings-Dresen, M. H. W., and Sluiter, J. K. (2012). Measuring short-term and long-term physiological stress effects by cortisol reactivity in saliva and hair. *International Archives of Occupational and Environmental Health*, 85(8):849–852.

van Loendersloot, L., van Wely, M., Limpens, J., Bossuyt, P., Repping, S., and van der Veen, F. (2010). Predictive factors in in vitro fertilization (IVF): a systematic review and meta-analysis. *Human Reproduction Update*, 16(6):577–589.

van Zonneveld, P., Scheffer, G., Broekmans, F., and te Velde, E. (2001). Hormones and reproductive aging. *Maturitas*, 38(1):83–91.

Vedhara, K., Stra, J., Miles, J., Sanderman, R., and Ranchor, A. (2006). Psychosocial factors associated with indices of cortisol production in women with breast cancer and controls. *Psychoneuroendocrinology*, 31(3):299–311.

Verhaak, C., Smeenk, J., Evers, A., Kremer, J., Kraaimaat, F., and Braat, D. (2007a). Women's emotional adjustment to IVF: a systematic review of 25 years of research. *Human Reproduction Update*, 13(1):27–36.

Verhaak, C., Smeenk, J., Nahuis, M., Kremer, J., and Braat, D. (2007b). Long-term psychological adjustment to IVF/ICSI treatment in women. *Human Reproduction*, 22(1):305–308.

Verhaak, C. M., Lintsen, A., Evers, A., and Braat, D. (2010). Who is at risk of emotional problems and how do you know? screening of women going for ivf treatment. *Human Reproduction*, 25(5):1234–1240.

Verhaak, C. M., Smeenk, J. M., Van Minnen, A., Kremer, J. A., and Kraaimaat, F. W. (2005). A longitudinal, prospective study on emotional adjustment before, during and after consecutive fertility treatment cycles. *Human reproduction*, 20(8):2253–2260.

Vrshek-Schallhorn, S., Doane, L. D., Mineka, S., Zinbarg, R. E., Craske, M. G., and Adam, E. K. (2013). The cortisol awakening response predicts major depression: predictive stability over a 4-year follow-up and effect of depression history. *Psychological Medicine*, 43(3):483–493.

Wallach, E. E., Hughes, E. G., and Brennan, B. G. (1996). Does cigarette smoking impair natural or assisted fecundity? *Fertility and sterility*, 66(5):679–689.

Wallston, K. A., Strudler Wallston, B., and DeVellis, R. (1978). Development of the multidimensional health locus of control (mhlc) scales. *Health education monographs*, 6(1):160–170.

Wang, H.-L., Visovsky, C., Ji, M., and Groer, M. (2016). Stress-related biobehavioral responses, symptoms, and physical activity among female veterans in the community: An exploratory study. *Nurse Education Today*, 47:2–9.

Wang, X., Busch, J. R., Banner, J., Linnet, K., and Johansen, S. S. (2019). Hair testing for cortisol by uplc–ms/ms in a family: External cross-contamination from use of cortisol cream. *Forensic Science International*, 305:109968.

Warner, K. E. (2015). Smoking and Health. In *International Encyclopedia of the Social & Behavioral Sciences*, pages 125–131. Elsevier.

Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the panas scales. *Journal of personality and social psychology*, 54(6):1063.

Watson, R. R. (2015). *Handbook of fertility: Nutrition, diet, lifestyle and reproductive health*. Academic Press.

Waylen, A., Metwally, M., Jones, G., Wilkinson, A., and Ledger, W. (2008). Effects of cigarette smoking upon clinical outcomes of assisted reproduction: a meta-analysis. *Human Reproduction Update*, 15(1):31–44.

Wen, J., Jiang, J., Ding, C., Dai, J., Liu, Y., Xia, Y., Liu, J., and Hu, Z. (2012). Birth defects in children conceived by in vitro fertilization and intracytoplasmic sperm injection: a meta-analysis. *Fertility and Sterility*, 97(6):1331–1337.e4.

Whirledge, S. and Cidlowski, J. A. (2010). Glucocorticoids, stress, and fertility. *Minerva endocrinologica*, 35(2):109.

Whirledge, S. and Cidlowski, J. A. (2017). Glucocorticoids and Reproduction: Traffic Control on the Road to Reproduction. *Trends in Endocrinology & Metabolism*, 28(6):399–415.

Who, J. and Consultation, F. E. (2003). Diet, nutrition and the prevention of chronic diseases. *World Health Organ Tech Rep Ser*, 916(i-viii):1–149.

Wikenius, E., Moe, V., Kjellevold, M., Smith, L., Lyle, R., Waagbø, R., Page, C. M., and Myhre, A. M. (2016). The Association between Hair Cortisol and Self-Reported Symptoms of Depression in Pregnant Women. *PLOS ONE*, 11(9):e0161804.

Wikoff, D., Welsh, B. T., Henderson, R., Brorby, G. P., Britt, J., Myers, E., Goldberger, J., Lieberman, H. R., O'Brien, C., Peck, J., Tenenbein, M., Weaver, C., Harvey, S., Urban, J., and Doepker, C. (2017). Systematic review of the potential adverse effects of caffeine consumption in healthy adults, pregnant women, adolescents, and children. *Food and Chemical Toxicology*, 109:585–648.

Wilkes, S., Chinn, D. J., Murdoch, A., and Rubin, G. (2009). Epidemiology and management of infertility: A population-based study in UK primary care. *Family Practice*, 26(4):269–274.

Williams, J. and MacKinnon, D. P. (2008). Resampling and distribution of the product methods for testing indirect effects in complex models. *Structural equation modeling: a multidisciplinary journal*, 15(1):23–51.

Windle, R., Wood, S., Shanks, N., Lightman, S., and Ingram, C. (1998). Ultradian rhythm of basal corticosterone release in the female rat: dynamic interaction with the response to acute stress. *Endocrinology*, 139(2):443–450.

Woods-Giscombé, C. L., Lobel, M., and Crandell, J. L. (2010). The impact of miscarriage and parity on patterns of maternal distress in pregnancy. *Research in nursing & health*, 33(4):316–328.

World Health Organization, t. et al. (2010). *Global recommendations on physical activity for health*. World Health Organization.

Wosu, A. C., Valdimarsdóttir, U., Shields, A. E., Williams, D. R., and Williams, M. A. (2013). Correlates of cortisol in human hair: implications for epidemiologic studies on health effects of chronic stress. *Annals of Epidemiology*, 23(12):797–811.e2.

Wu, M.-H., Su, P.-F., Chen, K.-Y., Tie, T.-H., Ke, H.-C., Chen, H., Su, Y.-C., Su, Y.-C., and Ou, H.-T. (2018). Heart rate variability among women undergoing in vitro fertilization treatment: Its predictive ability for pregnancy. *PloS one*, 13(3):e0193899.

Xiong, J., Lipsitz, O., Nasri, F., Lui, L. M., Gill, H., Phan, L., Chen-Li, D., Iacobucci, M., Ho, R., Majeed, A., et al. (2020). Impact of covid-19 pandemic on mental health in the general population: A systematic review. *Journal of affective disorders*, 277:55–64.

Yehuda, R., Halligan, S. L., Yang, R. K., Guo, L. S., Makotkine, I., Singh, B., and Pickholtz, D. (2003). Relationship between 24-hour urinary-free cortisol excretion and salivary cortisol levels sampled from awakening to bedtime in healthy subjects. *Life Sciences*, 73(3):349–358.

Young, E., Abelson, J., and Lightman, S. (2004). Cortisol pulsatility and its role in stress regulation and health. *Frontiers in Neuroendocrinology*, 25(2):69–76.

Yuan, H.-J., Han, X., He, N., Wang, G.-L., Gong, S., Lin, J., Gao, M., and Tan, J.-H. (2016). Glucocorticoids impair oocyte developmental potential by triggering apoptosis of ovarian cells via activating the fas system. *Scientific reports*, 6(1):1–12.

Zaidouni, A., Fatima, O., Amal, B., Siham, A., Houyam, H., Jalal, K., and Rachid, B. (2018). Predictors of Infertility Stress among Couples Diagnosed in a Public Center for Assisted Reproductive Technology. *Journal of Human Reproductive Sciences*, 11(4):376–383.

Zain, M. M. and Norman, R. J. (2008). Impact of obesity on female fertility and fertility treatment. *Women's health*, 4(2):183–194.

Zegers-Hochschild, F., Adamson, G. D., Dyer, S., Racowsky, C., de Mouzon, J., Sokol, R., Rienzi, L., Sunde, A., Schmidt, L., Cooke, I. D., Simpson, J. L., and van der Poel, S. (2017). The International Glossary on Infertility and Fertility Care, 2017†‡§. *Human Reproduction*, 32(9):1786–1801.

Zegers-Hochschild, F., Mansour, R., Ishihara, O., Adamson, G. D., de Mouzon, J., Nygren, K. G., and Sullivan, E. A. (2014). International Committee for Monitoring Assisted Reproductive Technology: world report on assisted reproductive technology, 2005. *Fertility and sterility*, 101(2):366–378.

Zhou, F.-j., Cai, Y.-n., and Dong, Y.-z. (2019). Stress increases the risk of pregnancy failure in couples undergoing ivf. *Stress*, 22(4):414–420.

Appendices

Appendix A

Systematic Review Supplementary Materials

A.1 Summary of studies

A.1.1 Table S1 Characteristics of studies

Authors (Year)	Type of Study	Study Size	Cortisol Outcome	Lifestyle Outcome	Mood Outcome	Included in meta-analysis
Batista et al. (2015)	Intervention	22	Salivary cortisol	Yoga practice	N/A	No
Bergeb-Cico et al. (2014)	Intervention	40	Salivary cortisol	N/A	Depression, PTSD	No
Booij et al. (2015)	Observational	30	Salivary cortisol	N/A	Other indices of stress, anxiety, depression	No
Braig et al. (2016)	Observational	768	Hair cortisol	N/A	Other indices of stress, anxiety, depression	Yes
Brown, Thomas et al. (2002)	Intervention	26	Serum cortisol	Diet	N/A	No
Brown, Coogle et al. (2016)	Intervention	38	Salivary cortisol	N/A	Perceived stress, other indices of stress	No
Byun (2017)	Observational	63	Salivary cortisol	N/A	Perceived stress, depression	Yes
Caparros-Gonzalez et al. (2017)	Observational	44	Hair cortisol	N/A	Depression	No
Chen, Yang et al. (2017)	Intervention	101	Salivary cortisol	Yoga practice	N/A	No
Chen, Gelaye et al. (2015)	Observational	85	Hair cortisol	BMI, smoking	N/A	No
Cieslak et al. (2011)	Observational	30	Salivary cortisol	N/A	PTSD	No
Coiro et al. (2007)	Intervention	20	Serum cortisol	Alcohol abstinence	N/A	No
Daubenmier et al. (2011)	Intervention	47	Serum cortisol, salivary cortisol	N/A	Mindfulness, stress, anxiety	No
Diment et al. (2012)	Intervention	30	Salivary cortisol	Diet	N/A	No
Faresjö et al. (2013)	Observational	591	Hair cortisol	Smoking	Perceived stress, anxiety, depression	Yes
Garcia-Leon et al. (2018)	Observational	62	Hair cortisol	BMI, physical activity	N/A	Yes
Geng et al. (2016)	Observational	51	Hair cortisol	BMI	Perceived stress, other indices of stress	Yes

Gerber et al. (2013)	Observational	46	Hair cortisol	Physical activity	Perceived stress, depression	Yes
Giesbrecht et al. (2013)	Observational	85	Salivary cortisol	N/A	Other indices of stress	Yes
Goldberg et al. (2014)	Intervention	135	Hair cortisol	Smoking	N/A	No
Grassi-Oliveira et al. (2012)	Observational	23	Hair cortisol	N/A	Mood impact of stressors	Yes
Heinze et al. (2016)	Observational	58	Hair cortisol	BMI, smoking, alcohol consumption	Perceived stress, anxiety, depression	Yes
Hoffman et al., (2016)	Observational	92	Hair cortisol	N/A	Perceives stress, anxiety, depression	Yes
Holt-Lunstad et al. (2008)	Intervention	72	Salivary cortisol	N/A	Unclear	No
Huang et al. (2012)	Intervention	18	Salivary cortisol	N/A	Perceived stress	No
Jacobs et al. (2013)	Intervention	60	Salivary cortisol	N/A	Mindfulness	No
Kramer et al. (2009)	Observational	651	Hair cortisol	N/A	Perceived stress, anxiety, depression	No
Larsen et al. (2016)	Observatioonal	532	Hair cortisol	BMI	N/A	Yes
Leivadi et al. (1999)	Intervention	30	Salivary cortisol	N/A	Anxiety, mood	No
Limm et al. (2011)	Intervention	174	Salivary cortisol	N/A	Perceived stress, other indices of stress, anxiety	No
Lovallo et al. (2005)	Intervention	96	Salivary cortisol	Caffeine consumption	N/A	No
Lynch et al. (2011)	Intervention	60	Salivary cortisol	N/A	Mindfulness, perceived stress, anxiety, depression	No
Manenscijn et al. (2017)	Observational	89	Hair cortisol	BMI	N/A	Yes
Mangine et al, (2017)	Intervention	33	Serum cortisol	Physical activity	N/A	No

Marcus et al. (2003)	Intervention	21	Salivary cortisol	N/A	Perceives stress	No
Miller et al. (2015)	Intervention	11	Serum cortisol	Sleep	N/A	No
O'Brien et al. (2013)	Observational	135	Hair cortisol	N/A	Perceived stress, Mood impact of stressors	Yes
Papacosta et al. (2013)	Intervention	11	Salivary cortisol	Physical activity	N/A	No
Qi et al. (2014)	Observational	39	Hair cortisol	BMI	Other indices of stress	Yes
Sannes et al. (2015)	Intervention	91	Salivary cortisol	N/A	Unclear	No
Schalinski et al. (2015)	Observational	45	Hair cortisol	N/A	Depression, PTSD	
Sinha et al. (2011)	Intervention	72	Serum cortisol	Alcohol abstinence	N/A	No
Stoppelbein and Greening (2015)	Observational	27	Salivary cortisol	N/A	PTSD	No
Sumra et al. (2015)	Observational	31	Urinary cortisol, hair cortisol	Physical activity	Perceived stress	Yes
Tam et al. (2014)	Intervention	46	Salivary cortisol	Diet	N/A	No
Tanskanen et al. (2011)	Intervention	57	Serum cortisol	Physical activity	N/A	No
Timon et al. (2013)	Intervention	20	Urinary cortisol	Physical activity	N/A	No
Torres et al. (2008)	Intervention	94	Salivary cortisol	Diet	N/A	No
Tsiouli et al. (2014)	Intervention	80	Salivary cortisol	Sleep, eating habit, smoking habits, drug use, physical activity	Perceived stress, other indices of stress	No
Ullmann et al. (2017)	Observational	40	Hair cortisol	N/A	Mood impact of stressors	No
Wang et al. (2016)	Observational	82	Hair cortisol	BMI, sleep	Depression	Yes
Wikenius et al. (2016)	Observational	181	Hair cortisol	N/A	Depression	Yes

A.1.2 Table S2 Summary of studies investigating long-term cortisol and lifestyle factors

Author (Year/Region)	Study type	Hypothesised intervention mechanism	Population and Study Size	Type of samples (indices if specified)	Time-frame cortisol was measured	Lifestyle factors and measures	Time-frame lifestyle was measured	Results
Chen et al. (2015/United States)	Observational	N/A	Caregivers of children with disabilities, n=85 (female=75, male=10)	Hair (HCC)	3 months	BMI (measured by researcher) Smoking (Self-identification of being smoker or not)	One-time assessment	HCC was not correlated with obesity status. No association between smoking status and HCC was found.
Faresjö et al. (2013/Sweden, Greece)	Observational	NA	General population, n=591 (female=321, male=270)	Hair (HCC)	3 months	Smoking (Self-identification of being daily smoker or not)	One-time assessment	No association between smoking status and HCC was found.
Garcia-Leon et al. (2018/Spain)	Observational	N/A	Pregnant women (n=62)	Hair (HCC)	3 trimester for pregnant participants	BMI (self-reported) Physical activity (self-reported regular/non-regular participant of physical activity)	One-time assessment	HCC was not correlated with BMI in pregnant women. Physical activity was significantly correlated with HCC in general sample.
Geng et al. (2016/China)	Observational	N/A	Participant were addicted to methamphetamine, n=51 (female=51)	Hair (HCC)	1 month	BMI (measure not described)	One-time assessment	No significant association between HCC and BMI was found.
Gerber, Kalak et al.(2013/Switzerland)	Observational	N/A	General population, n=46 (female=26, male=20)	Hair (HCC)	3 months	Vigorous physical activity: assessed with an accelerometer worn around the hip. All activities with >5,724 counts/min were defined as vigorous.	7 consecutive days	Higher HCC was significantly associated with greater vigorous physical activity.
Heinze et al. (2016/UK)	Observational	N/A	n=30 were mental health help-seekers and n=28 were healthy controls, (female=52, male=6)	Hair (HCC)	6 months	BMI (measure not described) Smoking and alcohol consumption (Self-reported: Alcohol, Smoking and Substance Involvement Screening Test)	BMI: one-time assessment Smoking and alcohol consumption: over the past 3 months	No significant correlations between HCC and BMI. No significant associations between smoking and HCC either over the past 3 month or the 3 month before were found. No significant association between alcohol consumption and HCC over the past 6 month.

Larsen et al. (2016/Denmark)	Observational	N/A	Parents of children at risk of obesity, n=532 (female=301, male=231)	Hair (HCC)	1-2 months	BMI (measured by researcher)	One-time assessment	Higher HCC was significantly associated with higher BMI.
Manenscijn et al. (2017/The Netherlands)	Observational	N/A	n=89 were shift workers and n=33 were controls, (m=122)	Hair (HCC)	3 months	BMI (measured by researcher)	One-time assessment	In shift workers there was a significant correlation between elevated levels of HCC and increased BMI.
Qi et al. (2014/China)	Observational	N/A	Healthy women, n=39	Hair (HCC)	1 month	BMI (self-reported)	One-time assessment	No significant correlations between HCC and BMI.
Sumra and Schillaci (2015/Canada)	Observational	N/A	Healthy women, n=31	Urine, hair (HCC)	1 month for urinary cortisol; 3 months for HCC	Physical activity: self-reported frequency	frequency of exercise per week	Neither hair, nor mean urinary cortisol was correlated with any other primary variables including physical activity (correlation coefficient not reported).
Wang et al. (2016/United States)	Observational	N/A	Female veterans, n=82	Hair (HCC)	3 months	BMI (measured by researcher) Sleep quality (the Pittsburgh Sleep Quality Index)	BMI: One-time assessment Sleep quality: not described	No significant correlations between HCC and BMI. No significant relationship between self-reported sleep quality and HCC.
Batista et al. (2015/Brazil)	6-week Yoga intervention 50 minute sessions twice a week for 6 weeks. Postures included snake posutre, boat posture, yoga mudra, Karmasana, self massage and deep relaxation.	Yoga practice would reduce stress, as reflected by cortisol	General population n=22 (female=15, male=7)	Saliva (daily rhythm of cortisol production, morning and night-time values)	Saliva samples were collected at 7.30am, 9am, noon, 12pm for 6 weeks (4 samples before and after yoga over 6 weeks).	NA	NA	Over long-term, Yoga induced higher cortisol production in the morning and lower production in the evening.
Brown et al. (2002/United States)	Diet intervention Crossover design of diet treatment for 11 menstrual cycles in total. Baseline + diet (polyunsaturated fat, PUFA) +washout + diet (soy protein + high saturated fat); control (high saturated fat). Each	N/A	Healthy women n=26	Blood	11 menstrual cycles in total. Blood samples were collected for 2 menstrual cycles for each treatment. 2-day consecutive blood sample for mid-follicular phase (day 7&8 or 8&9)	Meals were provided to the subjects each day. The subject's free meal night food records were obtained each week during all diet treatments. Caloric intake was individualized for each subject. The standardised study diets were consumed with 37% of	NA	A high polyunsaturated fat diet may decrease cortisol concentrations by 81.8 nmol/l (p = 0.088) compared with a high saturated fat diet. Compared with the control diet, no significant changes of cortisol were found for the soy diet.

	treatment lasting two menstrual cycles, with a washout time of two menstrual cycles between each treatment.				and mid-luteal phase (day 21&22 or 22&23) after menses during the second cycle of each diet treatment period and the baseline month.	kilocalories as fat, 48% as carbohydrate, and 15% as protein. Only the type of fat and amount of soy protein were altered for the study diets.		
Chen et al. (2017/China (Taiwan))	20-week Yoga intervention Prospective randomized control trial with a longitudinal, repeated-measures design. Intervention was yoga (70 minutes, twice per week for 20 weeks in groups of 10-12 women).	Yoga practice would reduce stress, as reflected by cortisol	Pregnant women n=101 (female=101)	Saliva	Saliva samples were collected 10 minutes before yoga and immediately after and at same time points for control group at 16, 20, 24, 28, 32 and 36 weeks gestational age (12 samples per participant). Time of sampling unclear.	NA	NA	No significant long-term effect of yoga on salivary cortisol levels.
Coiro et al. (2007/Italy)	8-week Alcohol abstinence Participants exercised on an electrically braked cycle ergometer. An initial load of 50 W was increased by 50 W every 3 minutes until subjective exhaustion, for about 15 minutes. Alcoholic patients were tested at 4, 6 and 8 weeks of alcohol abstinence.	Not described	Alcoholics n=20 (male=20) and healthy controls	Blood	Blood samples were taken at 30 minutes before exercise, initiation of exercise, 10, 20, 30, 40, 50, 60 minutes. Alcoholic participants provided 8 samples at 4, 6, 8 weeks after abstinence. Healthy controls were only assessed once.	Abstinence from alcohol was monitored with daily measurement of ethanol levels in 24-hour urine samples.	8 weeks	Serum cortisol levels significantly rose during exercise in the healthy controls and peaked after 30 minutes after the exercise task. However, the cortisol levels among abstinent alcohol patients did not significantly change during the exercise task, after 4-weeks of alcohol abstinence. At 8 weeks post-abstinence, the cortisol responses to exercise task for the abstinent alcoholic patients were significantly higher than 2 weeks previously and was no longer distinguishable from non-alcoholic participants.
Diment et al. (2012/UK)	8-week Supplemented diet intervention		Healthy male soldiers n=30	Saliva	8 weeks Baseline, pre-field exercise at the end	A 24-hour dietary record from each participant on the day before their visits with study staff was acquired.	NA	No intervention effect on plasma cortisol was found.

	Normal (habitual) diet plus mixed nutritional supplement.				of week 6; post-field exercise at week 8. Saliva samples were collected between 5.30am and 6.00am at baseline, pre-field exercise, and post-field exercise.			
Goldberg et al. (2014/Canada)	7-week mindfulness smoking cessation intervention: mindfulness training or freedom from smoking enhanced Participants were randomised to 1 of 2 smoking cessation interventions: mindfulness training for smoking versus freedom from smoking enhanced. Both groups had 8 meetings (over 7 weeks) and made quit attempts 1 month after the 1st meeting.	Negative affect and increased cortisol output under stress were related to increased risk of smoking relapse and that stress reduction interventions designed to cultivate emotion regulation skills and target social-emotional processes could aid smoking cessation and may be marked by lower levels of cortisol in hair.	Smokers n=135 (female=55)	Hair (HCC)	A 3cm hair sample was taken from each participant at 1-month post-quit study visit, to determine the HCC (1) during the month before the intervention, and (2) during the month post-quit	NA	NA	Participants' HCC was lower at 1-month post-quit compared to the month before the intervention. Intervention participants who stayed abstinent at the 1-month post-quit study visit had significantly lower levels of HCC compared with that in participants who had relapsed post-quit. However, differences were not found between two intervention groups for either decreases in HCC or the prevalence of smoking abstinence at post-quit.
Lovallo et al. (2005/United States)	4-week crossover caffeine stimulation intervention On each week, participants were instructed to abstain for 5 days from dietary caffeine but took caffeine capsules, followed by a test day,	NA	General population n=96 (female=48, male=48)	Saliva	Eight saliva samples were collected from each participant on the test day each week, before and after the stress challenge and each administration of caffeine capsules. Saliva samples were collected at the	NA	NA	Taking caffeine capsules on the test day after five days of abstinence resulted in a significant elevation of cortisol throughout the entire day. A significant elevation of cortisol was only seen in the group that had moderate caffeine intake of 300mg, but not 600mg,

	and a crossover day to buffer sudden changes in intake between study weeks. On the test day, each participant was administrated caffeine capsules three times and underwent a stress challenge between the first two administrations. In order to investigate the effect of different levels of caffeine intake on cortisol responses to caffeine challenges, three different maintenance doses were administrated to participants, to mimic a range of consumption commonly found in the US diet: none (0 mg/day), moderate (300 mg/day), high (600 mg/day).				following time: 7:30am on arrival at the laboratory, 9:00am after the pre-drug baseline, 10:30am at the end of stress testing, 11:10am after stress recovery, 1:00pm after pre-drug baseline 2, 2:00pm after post-drug response 2, at home at 6:00pm before the third capsule, and again at 7:00pm.			for five days. The elevation of cortisol in the 300mg group lasted only for about 6 hours on the test day.
Mangine et al. (2017/United States)	8-week resistance training Participants randomly assigned to high intensity low volume training or high volume moderate intensity training groups. Both groups required at least 28 training sessions in 8 weeks. Food intake was monitored via 3-day food diaries collected weekly.	Not described	Healthy resistance-trained men n=33	Blood (AUC)	Blood samples were collected at week 1 and 8, before exercise and immediately after exercise.	NA	NA	At week 1 of training, greater cortisol levels in response to exercise in the high-volume moderate-intensity training group were observed, compared to the high-intensity low-volume group. The difference in cortisol levels between groups was still significant at week 8 after training.
Miller et al. (2015/Australia)	5-week sleep restriction therapy Sleep restriction therapy comprised of one 40-minute face to face session which set a minimum time in bed	The sleep-restriction therapy would reduce early night plasma cortisol	People with insomnia n=11 (female=9, male=12)	Blood (mean plasma cortisol concentration)	Blood samples were taken intravenously every half hour from 22.00 to 06.00 on both nights of data collection at baseline and after the intervention.	Other sleep-related variables measured: Sleep onset latency; wake-time after sleep-onset; Insomnia Severity Index; Sleep Quality; Number of Awakenings; Total Sleep Time; Time in Bed	NA	Cortisol results from six patients showed that cortisol levels in the early morning were significantly higher post-intervention compared to baseline but no changes were observed in the pre-

	and go to bed threshold time. Minimum time in bed and a 'go-to-bed' threshold time were set for each participant at the beginning of the intervention. All participants were instructed to not go to bed until after the threshold time and to only go to bed if they felt 'sleepy tired'. Up to 5 weekly phone calls were made to participants to review the minimum time in bed, which was modified weekly for each individual.							sleep phase or early part of the night at post-intervention.
Papacosta et al. (2013/Greece)	5-week exercise intervention Week 1 participants trained as normal. Weeks 2-3 participants trained at a higher intensity. Weeks 4-5 training was tapered to half of the normal training intensity (week 1 intensity).	Not described	Male judo athletes n=11	Saliva	2 saliva samples were collected at the end of each week during week 1-3, and everyday during week 4-5. Saliva samples were provided at the following time: at 7.00am within 20 minutes after waking up and at 19pm before training.	NA	NA	Evening cortisol fell below baseline at the beginning of the tapering week and then returned to baseline by the end of the taper training.
Sinha et al. (2011/United States)	Each participant who were part of a 1- month in-hospital abstinence intervention underwent the stress and alcohol cue tasks twice: at baseline and at 1-month-abstinence.	Not deescribed	Alcoholics n=72 (female=41, male=31)	Blood	Blood samples were taken from participants each time before and after the tasks, at baseline and follow-up at week 5.	NA	NA	Although there were no significant differences at baseline, alcoholic participants showed a lack of cortisol response to both stress and alcohol cues compared with non-alcoholic controls at 1-month-abstinence.
Tam et al. (2014/United States, Australia)	24-week exercise and Diet Intervention Randomised study with four groups: control		Overweight individuals n=46 (female=20, male=26)	Saliva (morning cortisol and diurnal cortisol)	6 months (8 samples at baseline and end of intervention)	Body composition (measured using dual x-ray absorptiometry and abdominal fat distribution by multi-slice computed tomography)	NA	At baseline and 24 weeks post caloric restriction, there were no associations between morning/diurnal cortisol levels and body

	(weight maintenance diet based on the American Heart Association Step 1 diet); calorie restriction 25%; calorie restriction 12.5% + exercise (supervised aerobics); low calorie diet until achieved 15% weight loss.			Morning cortisol was defined as the mean cortisol concentration at 08:00 and 08:30. Diurnal cortisol was calculated as the mean of the 8 cortisol measures across the day.	Saliva samples collected at 8 time points over the day: 8:00, 8:30, 11:00, 11:30, 12:30, 1300,16:00 and 16:30.			composition (fat mass, visceral adipose tissue), in the whole group, or when groups were examined separately (data not shown).
Tanskanen et al. (2011/Finland)	7-week Exercise (military training) Participants divided into 3 categories according to levels of physical activity before military service, and then randomly selected 19 from each category.	Not described	Healthy male soldiers n=57	Blood	3 blood samples were taken at baseline and 3 at the end of intervention. Blood drawn from an antecubital vein after an overnight fast, 2 hours after breakfast, and immediately after exercise.	NA	NA	They reported that basal cortisol decreased from week 4 to week 7 of the training. Basal serum cortisol levels after 8 weeks of training were significantly higher in participants who were overreaching, compared to non-overreaching participants.
Timon et al. (2013/Spain)	8-week strength training Strength training program during 8 weeks with sessions 3 times a week.	Not described	Healthy women n=20	Urine	The first morning urine was collected from all subjects in three different moments across the menstrual cycle. Urine samples were taken the 1st –2nd day (menstruation phase), the 11th –12th day (follicular phase) and the 21st–22nd day (luteal phase). These samples were collected both the month before training and the month after training.	NA	NA	Significant changes in cortisol in urine following the training were not observed.

Torres et al. (2008/Australia)	4-week Diet (crossover design: 2-week control and 4-week diet, followed by 2-week control and 4-week diet) (mood) Randomised cross over. Participants received two (of 3) 4 week diets (OD diet, high in potassium, magnesium and fibre and low in sodium and sat fat; LNAHK diet was lower in sodium and calcium; HC, high calcium), each preceeded by a 2 week control diet. Subjects were instructed to follow the diets. During the CD and all test diet periods, a maximum of four Caffeine-containing drinks (e.g. cola drinks, coffee and tea) and two standard alcoholic drinks (10 g alcohol) were permitted daily.		General population n=94 (female=38, male=56)	Saliva (AUC)	10 weeks (3 samples at baseline and end of study) Saliva samples were collected at the end of each dietary period (4 weeks of diets, 2 weeks of control). A saliva sample was collected at approximately 12pm, 16pm and 20pm.			Higher cortisol levels were weakly associated with greater vigour, lower fatigue, and higher levels of urinary potassium and magnesium. However, the overall effects of interventions on cortisol were not reported.

A.1.3 Table S3 Summary of studies investigating long-term cortisol and mood factors

Author (Year/Region)	Study type	Hypothesised intervention mechanism	Population and Study Size	Type of samples (indices if specified)	Time-frame cortisol was measured	Mood factors and measures	Time-frame mood was measured	Results
Booij et al. (2015/The Netherlands)	Observational	N/A	Adult with and without depression (n=15/15), n=30 (female=22, male=8)	Saliva (daily slope of cortisol)	3 times a day (in the morning, 6 hours later in the afternoon, 6 hours later in the evening) for 30 days.	Depression (the Beck Depression Inventory; depressed individuals were identified using the DSM-IV diagnose of Major Depressive Disorder)	Unclear	Cortisol in saliva was higher, and the daily cortisol slope steeper in the depressed participants compared to the non-depressed ones.
Braig et al. (2016/Germany)	Observational	N/A	Women shortly after delivery, n=768	Hair (HCC[a])	3 months	Chronic stress (the screening scale of the Tirer Inventory of Chronic Stress) Anxiety (the Pregnancy-Related Anxiety Questionnaire; the Hospital Anxiety and Depression scale) Depression (the Hospital Anxiety and Depression scale)	Over the past 3 months	No association between chronic stress and HCC. No association between anxiety and HCC. No association between depression and HCC.
Byun et al. (2013/United States)	Observational	N/A	Caregivers of stroke survivors, n=63 (female=42, male=21)	Saliva (indices not specified)	Over 4 weeks (T1=within 2 weeks post-stroke, T2=6 weeks post-stroke). Saliva samples were collected on waking and again at 9pm.	Perceived stress (Perceived Stress Scale) Depression (the Patient Health Questionnaire)	Over the past month	Salivary cortisol on waking was not correlated with perceived stress but salivary cortisol in the evening had a positive correlation with higher levels of perceived stress. Salivary cortisol was not significantly correlated with depression at the beginning of the study but elevated cortisol levels in saliva in the evening were significantly correlated with greater self-reported depressive symptoms four weeks later.
Caparros-Gonzalez et al. (2017/Spain)	Observational	N/A	Pregnant women, n=44	Hair (HCC)	9 months (over trimester 1, 2, and 3)	Depression (the Edinburgh Postnatal Depression Scale)	At postpartum but time-frame unclear	Higher HCC during 1st and 3rd trimester was prospectively associated with greater postpartum depression.

Chen et al. (2015/United States)	Observational	N/A	Caregivers of children with disabilities, n=85 (female=75, male=10)	Hair (HCC)	3 months	Perceived stress (Perceived Stress Scale)	Over the past month	No association between perceived stress and HCC.
Cieslak et al. (2011/United States)	Observational	N/A	People with PTSD[f], N=30 (F=17, M=13)	Saliva (AUCg, AUCi)	Saliva samples from participants four times a day (upon awakening, 1h, 4h, 12h after waking) at 1 week, 1 month, and 3 months after a traffic accident (i.e., trauma).	PTSD (the Impact of Event Scale)	At 1 week, 1 month, and 3 months after the accident	The impact of trauma at 1-month post-accident was negatively associated with AUC_i (reflecting changes in daytime cortisol secretions) at 3 months and the impact of trauma at 3 months was positively associated with AUC_g (reflecting total daytime cortisol release) at 1-month post-accident.
Faresjö et al. (2013/Sweden, Greece)	Observational	N/A	General population, n=591 (female=321, male=270)	Hair (HCC)	3 months	Perceived stress (Perceived Stress Scale) Anxiety (the Hospital Anxiety and Depression scale) Depression (the Hospital Anxiety and Depression scale)	Perceives stress: over the past month Anxiety and depression: unclear	No association between perceived stress and HCC. No association between anxiety and HCC. Higher HCC was significantly correlated with greater depression.
Geng et al. (2016/China)	Observational	N/A	Participant were addicted to methamphetamine, n=51 (female=51)	Hair (HCC)	1 month	Perceived stress (Perceived Stress Scale) Implicit perceived stress (Implicit Association Test)	Over the past month	No association between self-reported perceived stress measured by Perceived Stress Scale and HCC. Higher HCC was significantly associated with greater implicit perceived stress.
Gerber et al. (2013/Switzerland)	Observational	N/A	General population, n=46 (female=26, male=20)	Hair (HCC)	3 months	Perceived stress (Perceived Stress Scale) Depression (the Beck Depression Inventory)	Perceived stress: over the past month Depression: unclear	Higher HCC was significantly correlated with lower perceived stress, but the correlation was weak. Lower HCC was significantly correlated with greater depression.
Giesbrecht et al. (2013/Canada)	Observational	N/A	Pregnant women, n=85	Saliva (Cortisol awaking levels, CAR, diurnal slope)	3 trimesters of pregnancy (saliva samples were collected 5 times per day over 2	Distress (the Profile of Mood States)	At each sampling point over the same time frame.	Higher salivary cortisol was significantly associated with distress.

					consecutive days at approximately 14, 20, 32 weeks of gestation). Saliva samples were collected at the following time: upon waking (allowing for individualized wake times), 30 minutes after waking, and semi-randomly after the anchor times of 11am, 3.30pm, and 8pm.			
Grassi-Oliveira et al. (2012/Brazil)	Observational	N/A	Crack-cocaine-dependent women, n=23	Hair (HCC)	3 months	Mood impact of stressors (semi-structured interview with the Psykel Life Events Scale)	Over the past 90 days	Higher HCC was significantly correlated with the number of stressful events exposure in the previous 90 days and 30 days.
Heinze et al. (2016/UK)	Observational	N/A	n=30 were mental health help-seekers and n=28 were healthy controls, (female=52, male=6)	Hair (HCC)	6 months	Perceived stress (Perceived Stress Scale) Anxiety (the Overall Anxiety Severity and Impairment Scale) Depression (the Quick Inventory of Depression Symptoms; the Kessler Psychological Distress Scale)	Perceive stress: Over the past month Anxiety: 3 months within the 6 months when cortisol was measured Depression: over the past 7-30 days.	No association between perceived stress and HCC. No association between anxiety and HCC. No association between depression and HCC.
Hoffman et al. (2017/United States)	Observational	N/A	Pregnant women, n=92	Hair (HCC)	3 trimesters (at 16, 28, 40 weeks of gestation)	Perceived stress (Perceived stress scale) Anxiety (the State-Trait Anxiety Inventory) Depression (the Center for	3 trimesters (at 16, 22, 28, 34, 40 weeks of gestation for the past month at each time)	Higher HCC was prospectively correlated with higher levels of perceived stress at 40 weeks of gestation. Higher HCC at 2-trimester was significantly correlated with perceived stress at 16 weeks of gestation. Greater anxiety measured at 16 and 28 weeks of pregnancy was correlated with higher HCC in both the second and the third trimester.

						Epidemiologic Studies – Depression Scale)		Higher HCC during the 1st trimester was prospectively associated with greater depression at 40 weeks of gestation, higher HCC during the 2nd-trimester was significantly associated with greater depression at 16 and 28 weeks of gestation, higher HCC at the 3rd-trimester was significantly associated with depression at 16 and 28 weeks of gestation.
Kramer et al. (2009/Canada)	Observational	N/A	Pregnant women, n=651	Hair (HCC)	9 months	Perceived stress (Perceived Stress Scale) Anxiety (a 4-item scale) Depression (the Center for Epidemiologic Studies – Depression Scale)	Perceived stress: at 24-26 weeks of gestation for the past month. Anxiety: at 24-26 weeks of gestation Depression: at 24-26 weeks of gestation	No association between perceived stress and HCC. No association between anxiety and HCC. No association between depression and HCC.
O'Brien et al. (2013/United States)	Observational	N/A	General population, N=135 (F=88, M=47)	Hair (HCC)	3 months	Perceived stress (Perceived stress scale) Mood impact of stressors (chaos in the neighbourhood in the past year, City Stress Index; chaos/order of the home environment ,Chaos, Hubbub, and Order Scale)	Perceives stress: Over the past 3 months Mood impact of stressors: not specified	No association between perceived stress and HCC. No association between HCC and chaos either in the neighbours or in the home.
Qi et al. (2014/China)	Observational	N/A	Healthy women, n=39	Hair (HCC)	1 month	Work stress (the Effort-Reward Imbalance Scale)	Over the past month	Neither the effort nor the reward scores were correlated with HCC. However, the effort-reward imbalance (1.83*effort-score/reward-score) at work was positively correlated with HCC.
Schalinski et al. (2015/Germany)	Observational	N/A	Women with stress-related disorders, n=45	Hair (HCC)	3 months	Depression (the clinical-administrated Hamilton Rating Scale)	Unclear	No association between depression and HCC. No association between PTSD and HCC.

						PTSD (the Clinician-Administered PTSD Scale)		
Stoppelbein et al. (2015/United States)	Observational	N/A	Mothers of children diagnosed with cancer, n=27	Saliva (AUCg)	Saliva samples were collected between 3-5 pm on the first day, and between 6-8 am and 3-5 pm on the following day. This saliva collection protocol was repeated monthly for 12 months.	PTSD (the Posttraumatic Stress Disorder Checklist)	Over 12 months	Higher levels of cortisol were significantly associated with one type of self-reported PTSD symptoms: numbing symptoms
Sumra and Schillaci (2015/Canada)	Observational	N/A	Healthy women, n=31	Urine, hair (HCC)	1 month for urinary cortisol; 3 months for HCC	Perceived stress (perceived stress scale)	Over the past month	No association between perceived stress and HCC.
Ullmann et al. (2017/Germany)	Observational	N/A	General population, N=40 (F=22, M=18)	Hair (HCC)	3 months	Mood impact of stressors (mental burdening events, the Recalled Parental Rearing Behaviour Questionnaire)	Over the past 3 months	Those with mental burdening in the previous 3 months had higher cortisol compared with those without mentally burdening events.
Wang et al. (2016/United States)	Observational	N/A	Female veterans, n=82	Hair (HCC)	3 months	Depression (the Center for Epidemiologic Studies – Depression Scale)	Unclear	No association between depression and HCC.
Wikenius et al. (2016/Norway)	Observational	N/A	Pregnant women, n=181	Hair (HCC)	1-3 month (Second trimester)	Depression (the Edinburgh Postnatal Depression Scale)	Over the past week	No association between depression and HCC.
Bergen-Cico et al. (2014/United States)	4-week mindfulness intervention Weekly in person 90 minute mindfulness based meditation group sessions for 4 consecutive weeks. Participants were given a book and audio CD with guided practice.	Engagement in the stress-reduction intervention would reduce diurnal cortisol output and CAR among veterans with PTSD	Veterans n=40 (female=4, male=36)	Saliva(AUC_g, AUC_i, CAR)	Saliva samples were collected at waking, 45min after waking, 8am, 12pm, and bedtime each day for 2 consecutive days, before treatment at baseline and after intervention.	Depression (Patient Health Questionnaire) PTSD (the Military PTSD checklist).	At baseline only	Significant changes in cortisol levels which was associated with engagement and number of sessions completed of the intervention. Participants who completed all four sessions had a significant decrease in the CAR and increase in AUC, compared with the control group. Changes in either PTSD or stress were not described.

Brown et al. (2016/United States)	8-week mindfulness intervention Participants attended weekly classes of 1.5-2 hours in length.	The intervention was to reduce psychological and neuroendocrine markers of stress.	Adult caregivers n=38 (female=32,male=6)	Saliva (diurnal cortisol)	Saliva samples were collected at 6 time point a day at 3 study periods (pre- and post-intervention, 3 month follow up). Saliva samples were collected at the following time: before rising in the morning, 45 min after awakening, 2.5 h after awakening, 8 h after awakening, 12 h after awakening, and bedtime.	Perceived stress (Perceived Stress Scale) Mood (Profile of Mood States)	NA	The mindfulness intervention significantly improved participants' perceived stress levels and mood. However, the improvements in mood, perceived stress, and burden among participants at 3-month follow-up were not significantly different between the mindfulness intervention and social support group. Neither conditions showed changes in diurnal cortisol response curve across the study time-frame.
Daubenmier et al. (2011/United States)	4-month mindfulness intervention Twenty-four participants were asked to attend a weekly 2.5h class for nine weeks and a 7-hour silent day of guided meditation practice after class 6.	Mindfulness training would enhance awareness of and responsiveness to bodily sensations and reduce psychological stress, emotional eating, and cortisol secretion, which would lead to reductions in abdominal fat	Stressed women n=47	Blood; Saliva(CAR, cortisol slope)	Saliva sampled were collected upon awakening, 30 minutes after and before bed, on four days before and after the intervention. Blood samples were collected pre- and post-intervention.	Mindfulness (the Kentucky Inventory of Mindfulness Skills) Stress (Wheaton Chronic Stress Inventory; Perceived Stress Scale) Anxiety (The State-Trait Anxiety Scale)	NA	Participants in the intervention group showed significant improvements in mindfulness, anxiety and external-based eating after the intervention, compared to control participants. However, differences between the intervention and control groups were not significant for emotional eating, the average cortisol awakening response, overall weight, abdominal fat, or fat distribution over time. Neither groups showed significant changes in the cortisol slope or morning serum cortisol concentrations. Improvement of mindfulness, chronic stress, and Subgroup of obese participants in the intervention group showed significant reductions in the cortisol awakening response and maintained body weight while obese participants in the control group had a stable cortisol awakening response and gained weight.
Holt-Lundstad et al. (2008/United States)	4-week support enhancement intervention	Warm touch enhancement may improve mood and stress and	n=36 couples (female=36, male=36)	Saliva (AUC)	Saliva samples were collected five times throughout a day at pre- and post-	Unclear	Unclear	No significant intervention effect on salivary cortisol.

	Twenty couples (n=40) in the intervention group were asked to attend two 1-hour group sessions for couple contact enhancement techniques. These couples were also instructed to practice the techniques for 30 minutes, 3 times per week for 4 weeks. A behaviour monitoring control group was designed where subjects were instructed not to change anything about their normal behaviours with their spouse and to simply keep a diary of their physical affection and mood.	result in reduced cortisol levels.			intervention. Saliva samples were collected at 7am, 12pm, 5pm, 10pm, and on wakening before participants got out of bed.			
Huang et al. (2012/UK)	5-week stress reduction acupuncture intervention The intervention group received weekly acupuncture sessions for 5 weeks. The relaxation-only group received weekly relaxation (without acupuncture) sessions for 5 weeks. The waitlist control group did not receive any treatment for 5 weeks.	Traditional Chinese acupuncture to treat chronic stress could change diurnal salivary cortisol profile in healthy individuals.	Stressed individuals n=18 (female=15, male=3)	Saliva (CAR, daily profile)	Saliva samples were collected four times a day on 9 days pre-, post-intervention and at follow-up after 2 weeks. Saliva samples were collected immediately upon waking in the morning, 30 minutes after waking, 3 hours after waking, and 12 hours after waking.	Perceived stress (Perceived Stress Scale)	NA	Perceived stress was reduced in all groups however the differences between pre- and post-intervention and between- and within-groups were not significant. Non-significant increased CAR was seen for both the intervention and relaxation-only groups. No significant changes in the decline in cortisol or the daily average cortisol concentrations were found between groups.
Jacobs et al. (2013/United States)	8-week mindfulness meditation intervention	Whether self-reported mindfulness, as well as changes in	General population n=60 (female=32, male=28)	Saliva (p.m. cortisol[e])	Saliva samples were collected for 3 consecutive days at pre-retreat (2 weeks after	Self-reported mindfulness (the 37-item Five Factor Mindfulness Questionnaire)	NA	Self-reported mindfulness significantly increased after the intervention. No significant changes were found in cortisol levels. Changes in mindfulness were

	The intervention group met twice per day for 1-hour sessions for meditation practice.	mindfulness following the intervention, were related to resting cortisol. Direction of the intervention effect on cortisol not hypothesised.			arrival) and post-retreat (10 weeks after arrival). Samples were collected in the afternoon (2 hr after lunch) and at bedtime (right before sleep).			inversely associated with changes in resting p.m. cortisol levels.
Leivadi et al. (1999/United States)	5-week massage intervention Participants randomised to massage received a 30-minute session twice a week for 5 weeks. The relaxation control group received a 30-minute session twice a week for 5 weeks.	Massage therapy would improve mood, decrease cortisol and other stressors such as anxiety and neck, shoulder, and back pain.	General population n=30 (massage group=15, relaxation group=15)	Saliva	Saliva samples were collected twice a day pre- and post-intervention. Time of sampling unclear.	Anxiety (State Anxiety Inventory) Mood (Profile Of Mood States)	NA	Both groups reported improvements in mood and anxiety after sessions. Changes in mood, anxiety, or cortisol levels were not observed over time. Participants reported reduced n neck, shoulder, and back pain after the treatment sessions.F.
Limm et al. (2011/Germany)	One-year work-site stress-management intervention The intervention consisted of two full-day group sessions followed by two refresher sessions within the next 8 months. At each group session, participants were asked to remember and share individual stressful situations at work. Several tools to cope with difficult and stressful situations were introduced by the trainer at the group session.	Hypothesised intervention effects on cortisol not described.	General population n=174 (female=22, male=152)	Saliva	Saliva samples to measure cortisol were collected seven times throughout a day both at pre-intervention and at 1-year.	Perceived stress reactivity (Stress Reactivity Scale) Effort-reward imbalance (the Effort-Reward Imbalance Scale) Anxiety and depression (the Hospital Anxiety and Depression Scale)	NA	Perceived stress reactivity reduced in both groups but a significantly greater reduction was seen in the intervention group at 1-year follow-up. No intervention effect on cortisol was observed. Improvements were seen for anxiety, depression, effort-reward imbalance and while greater in the intervention group compared to the control group, they were not significant.

Lynch et al. (2011/UK)	8-week mindfulness-based coping intervention Participants attended one 1.5h meeting per week for eight consecutive weeks.	Whether the intervention reduce stress which may be reflected by the biomarkers of stress. Direction of the intervention effect on cortisol not hypothesised.	General population n=60	Saliva	Saliva samples were collected 8 times throughout the day for 2 consecutive days at pre- and post-intervention. Saliva samples were collected at the following time: upon waking, 30 and 60 min later, and then at defined times throughout the day (8am, 11am, 3pm and 8pm).	Mindfulness (Five Factor Mindfulness Questionnaire) Perceived stress (Perceived Stress Scale) Anxiety and depression (the Hospital Anxiety and Depression Scale)	NA	Mindfulness intervention significantly improved perceived stress, anxiety, depression among participants, along with increased self-reported mindfulness. However, cortisol levels did not significantly change after the intervention.
Marcus et al. (2003/United States)	8-week mindfulness-based stress reduction intervention Participants attended a mindfulness meditation session for 2.5 hours per week for eight consecutive weeks. They were also asked to practice mindfulness outside the sessions for 45-60 minutes for 6 days per week.	Whether the intervention reduce stress which may be reflected by the biomarkers of stress. Direction of the intervention effect on cortisol not hypothesised.	People in theraputic community n=21 (female=18, male=3)	Saliva (AUC)	Saliva samples were collected at 0, 30, 45 and 60 minutes after awakening on the first morning pre intervention and post intervention.	Perceived stress (Perceived Stress Scale)	NA	A non-significant decrease in self-reported perceived stress after intervention among participants was found. The awakening salivary cortisol levels were significantly lower following the intervention.
Sannes et al. (2015/United States)	10-week cognitive behavioural stress management The intervention involved 10 two-hour sessions in groups of 4-10 on a weekly basis, which included education on awareness of the effects of stress, cognitive reframing, cognitive coping skills training etc.	Women in the intervention group would demonstrate lower intraindividual cortisol variability (i.e., the degree to which an individual's cortisol output may be erratic on a given day).	Healthy women at risk of breast cancer n=91	Saliva (cortisol slope, AUCg, CAR, individual cortisol variability)	Saliva samples were collected 5 times a day for 2 consecutive days at pre- and post-intervention, and at 1-month follow-up. Saliva samples were collected at awakening, 30min following awakening, 11am, 5pm and at bedtime.	Unclear	Unclear	Lower levels of intraindividual cortisol variability among intervention participants were found, compared to the waitlist control group. Improvements towards a steeper cortisol slope and less overall cortisol output (area under the curve with respect to ground, AUCg) were also observed in the intervention group. Significant changes in other cortisol indices were not found.

Tsiouli et al. (2014/Greece)	8-week stress management Intervention group were trained in progressive muscle relaxation and diaphragmatic breathing for 37 minutes. They were asked to perform these techniques twice a day for 8 weeks, and asked to adapt a healthy lifestyle.	The intervention could reduce perceived and parenting stress, increase internal locus of control, promote healthy lifestyle, and normalise cortisol levels. However the direction of changes in cortisol was not specified.	Caregivers of children and adolescents with type 1 diabetes n=80	Saliva (diurnal cortisol)	Samples collected at 5 time points throughout a day pre- and post-intervention. Saliva samples were collected at 8am, 12pm, 3pm, 6pm, 9pm, and at bedtime.	Lifestyle (questionnaire including quality of sleep, eating habit, smoking habits, drug use, physical activity etc.) Perceived stress (Perceived Stress Scale) Parental Stress (Parenting Stress Index-Short Form)	NA	No statistically significant improvement was found in cortisol levels in either the intervention or the control group. Stress was significantly reduced in the intervention group. Improvement in lifestyle factors such as sleep, stress-related symptoms, and eat habits were improved in both groups.

A.2 Risk of bias assessment

A.2.1 Table S4 Summary of the risk of bias of studies assessed by the 27-item Downs and Black checklist (Downs and Black 1998)

	Reporting	External Validity	Internal Validity (bias)	Internal Validity (selection bias)	Power	Total Score	Overall quality rating
Batista et al. (2015)	7	0	3	0	0	10	Limited
Bergen-Cico et al. (2014)	7	1	3	1	0	12	Limited
Booij et al. (2015)	7	0	3	1	0	11	Moderate
Braig et al. (2015)	7	0	3	2	0	12	Moderate
Brown et al. (2002)	9	0	3	2	0	14	Moderate
Brown et al. (2016)	8	0	3	2	0	13	Limited
Byun et al. (2013)	9	0	3	1	1	14	Moderate
Caparros-Gonzalez et al. (2017)	6	0	3	1	0	10	Moderate
Chen et al. (2015)	6	0	2	2	0	10	Moderate
Chen et al. (2017)	9	0	4	4	1	17	Moderate
Cieslak et al. (2010)	6	1	3	1	0	11	Limited
Coiro et al. (2007)	4	0	3	0	0	7	Limited
Daubenmier et al. (2011)	10	0	3	4	0	17	Moderate
Diment et al. (2012)	4	1	3	2	0	10	Limited
Faresjo et al. (2013)	6	1	2	0	0	9	Limited
Garcia-Leon et al. (2018)	3	0	2	0	0	5	Low
Gerber et al. (2016)	5	0	2	1	0	8	Limited
Giesbrecht et al. (2012)	7	0	3	2	0	11	Moderate
Geng et al. (2013)	4	0	2	0	0	6	Limited
Goldberg et al. (2014)	6	0	4	0	0	10	Limited
Grassi-Oliveira et al. (2012)	5	0	2	0	0	7	Limited
Heinze et al. (2015)	7	1	2	1	0	11	Moderate
Hoffman et al. (2016)	7	1	3	2	0	13	Moderate
Holt-Lundstad et al. (2008)	8	0	3	3	0	14	Moderate
Huang et al. (2012)	8	0	4	1	0	13	Limited
Jacobs et al. (2013)	5	0	3	1	0	9	Limited
Kramer et al. (2009)	5	0	2	2	0	6	Limited
Larsen et al. (2016)	8	0	2	1	0	9	Limited
Leivadi et al. (1999)	5	0	3	1	0	7	Low
Limm et al. (2010)	9	0	4	5	0	17	Moderate
Lovallo et al. (2005)	6	0	6	2	0	11	Limited
Lynch et al. (2011)	8	0	3	1	0	11	Limited
Manenscijn et al. (2011)	4	0	2	1	0	6	Limited
Mangine et al. (2016)	6	0	3	0	0	9	Limited
Marcus et al. (2003)	6	1	3	1	0	11	Limited
Miller et al. (2015)	7	0	4	1	0	12	Limited
O'Brien et al. (2013)	2	1	2	0	0	5	Limited
Papacosta et al. (2013)	5	1	3	0	0	9	Limited
Qi et al. (2014)	5	0	2	0	0	7	Limited
Sannes et al. (2015)	8	0	4	2	0	14	Moderate
Schalinski et al. (2015)	5	0	3	0	0	7	Limited
Sinha et al. (2011)	5	0	2	1	0	7	Low
Stoppelbein et al. (2015)	4	0	3	1	0	7	Limited
Sumra et al. (2015)	7	0	3	2	0	12	Moderate
Tam et al. (2014)	6	0	3	0	0	9	Limited
Tanskanen et al. (2011)	5	0	3	1	0	9	Limited
Timon et al. (2013)	6	0	3	0	0	9	Limited
Torres et al. (2008)	4	0	3	1	0	8	Limited
Tsiouli et al. (2014)	9	1	3	2	0	15	Moderate
Ullmann et al. (2017)	5	0	3	1	0	9	Limited
Wang et al. (2016)	8	0	2	1	0	11	Moderate
Wikenius et al. (2016)	8	0	2	1	0	11	Moderate

A.2.2 Table S5 Categorisation of the total risk of bias scores assessed by the 27-item Downs and Black checklist (Downs and Black 1998)[1]

Quality rating	**Percentage**	**Observational study total scores**	**Interventional study total scores**
High	75% and above	15 and above	21 and above
Moderate	50-74%	10-14	14-20
Limited	26-49%	6-9	7-13
Low	25% and below	5 and below	7 and below

[1]Categorisation of the total risk of bias scores were adapted from Huffer et al. (2017).

A.3 Search Terms

A.3.1 Table S6 Terms of cortisol, lifestyle and mood factors for database Search

	Terms
Cortisol	Cortisol reactivity
	Cortisol awakening response (CAR)
	Area under the curve with respect to ground (AUCg)
	Diurnal slope
	Long-term cortisol output
lifestyle factors	Lifestyle
	Sleep
	Exercises or physical activity
	Diet
	BMI/Body weight
	Caffeine intake
	Alcohol intake
	Smoking
Psychological factors	Mood or emotions
	Distress
	Anxiety
	Depressive symptom
	Psychological stress
	Depression
	Posttraumatic stress disorder (PTSD)

A.3.2 Search terms used for MEDLINE, Embase, PsycInfo

1. cortisol.mp. or *Hydrocortisone/

2. exp lifestyle/

3. (physical activity or physical activities).mp.

4. (exercise or exercises).mp.

5. (caffeine intake or caffeine consumption).mp.

6. (alcohol intake or alcohol consumption or alcohol drinking).mp.

7. diet.mp. or Diet/

8. Smoking/ or smoking.mp. or Marijuana Smoking/

9. Body Mass Index/ or BMI.mp. or Body Weight/

10. sleep.mp. or Sleep/ or Sleep Hygiene/

11. 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10

12. mood.mp.

13. (emotion or emotions).mp.

14. emotional status.mp.

15. *Stress, Psychological/ or psychological distress.mp.

16. psychological stress.mp.

17. *Anxiety/ or anxiety.mp.

18. depression.mp. or *Depression/

19. PTSD.mp. or *Stress Disorders, Post-Traumatic/

20. 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19

21. exp Adult/ or general population.mp.

22. (healthy volunteers or healthy people or healthy participant).mp.

23. 21 or 22

24. 11 or 20

25. 1 and 24

26. 23 and 25

27. limit 26 to (english language and full text)

Appendix B

Research Ethics

B.1 Patient Information Sheet

CARE fertility

University of Nottingham
UK | CHINA | MALAYSIA

Does the hormone cortisol affect fertility?

FEMALE patient information sheet



HELP US TO HELP YOU!

Research Team: Dr. Maha Ragunath, Dr. Kanna Jayaprakasan, Dr. Walid Maalouf, Dr. Peter Marsters, Dr. Matt Tomlinson, Dr. Adam Massey, Dr. Grazziela Figueredo, Miss Ru Jia, Miss Sophie Dodds and Professor Kavita Vedhara

Your Invitation

We would like to invite you to take part in our research study. Before you decide we would like you to understand why the research is being done and what it would involve for you. One of our team will go through the information sheet with you and answer any questions you have. Talk to others about the study if you wish. Ask us if there is anything that is not clear.

What is the purpose of the study?

The purpose of this study is to explore whether the hormone cortisol affects fertility and the chances of becoming pregnant in patients who have In Vitro Fertilisation (IVF) or Intrauterine insemination (IUI) treatment. Recent research done here in Nottingham has suggested that cortisol may be an important factor. The purpose of this study is to allow us to be more certain of whether or not this is the case and to help us to develop treatments which could reduce cortisol and so improve chances of pregnancy. **The study entails** taking a small hair sample taken from the back of the head using scissors. This procedure is painless and your sample will be analysed within our laboratory to determine what your cortisol levels were over the 3 months prior to having treatment. We will also ask you to complete two lifestyle questionnaires before your treatment (measuring factors such as smoking, stress, exercise and use of class A drugs) so we can examine how these factors may affect your cortisol levels.

Why have I been invited?

You are being invited to take part because you are a women undergoing IVF or IUI treatment. We are inviting 200 women undergoing 'short protocol' IVF and 200 women undergoing 'long protocol' IVF. A long protocol starts with suppressing your hormones so that doctors can have control of when you ovulate. Doctors call this 'down regulation' which typically involves taking tablets, followed by taking daily injections to stimulate your ovaries. Some patients have short protocol IVF in which the tablets aren't required (i.e. they do not undergo down regulation) so the process is quicker. If you do not know which protocol you are on then please ask your doctor. We are also recruiting 200 women undergoing IUI and 200 men to explore factors that may influence male fertility too.

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. This would not affect your legal rights and will have no bearing on the treatment you receive from your clinic.



What will happen to me if I take part?

If you choose to take part we will ask you to complete a lifestyle questionnaire (this should take about 10-15 minutes to complete) during your initial routine visit to the clinic and again just before you start your treatment (IUI or IVF). The questionnaire can be completed electronically or paper based from home or within the clinic. The questionnaire will measure your lifestyle behaviors (e.g. stress, smoking, alcohol consumption, exercise, sleep quality, use of class A drugs) over the previous 4 weeks. Prior to undergoing your treatment a small hair sample will be taken (approximately half the width of a pencil) using scissors from the back of your head. Before you provide a sample we will ask you to complete a short 5-minute survey about your hair (e.g. use of any treatments/washing frequency etc.) Providing a hair sample is painless and will take a couple of minutes. You will have the opportunity to look or feel the amount of hair to be taken before it is taken.

We take care to make sure the sample is taken from a section that is not visible (i.e. you will not be left with a bald patch). The sample will be collected by a member of the research team. At the end of your treatment, we will collect information from your patient notes on your treatment outcomes. This information will be used by us to explore the relationship between cortisol and pregnancy. In total, it will take approximately 10 – 15 minutes for each lifestyle survey and less than a few minutes to provide the hair sample.

Expenses and payments

Participants will not be paid to participate in the study.

What are the possible disadvantages and risks of taking part?

There are no risks to your health likely to arise from taking part in this study. The hair sampling is non-invasive and painless.

What are the possible benefits of taking part?

We cannot promise the study will help you but the information we get from this study will help us to find out if cortisol affects fertility. In the future, this knowledge may be used to improve treatment outcomes for other patients.

What happens when the research study stops?

Having participated in the study you will be given the opportunity to receive summary findings when we know the outcome of the study. If you are interested in knowing the results you can ask any of the researchers for this information. At the end of the study, findings from the study will be published at academic conferences and in academic journals but all data collected as part of the study will be strictly confidential and anonymous.

What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions. The researchers contact details are given at the end of this information sheet. If you remain unhappy and wish to complain formally, you can do this by contacting Prof. Kavita Vedhara by email kavita.vedhara@nottingham.ac.uk or via telephone 0115 846 6931.

If this achieves no satisfactory outcome you should then contact the Ethics Committee Secretary, Mrs Louise Sabir, Division of Therapeutics and Molecular Medicine, D Floor, South Block, Queens Medical Centre, Nottingham, NG7 2UH. Tel 0115-8231063 or email louise.sabir@nottingham.ac.uk

If you are a NHS patient then you can also contact The Patient Advice and Liaison Service (PALS), which offers confidential advice, support and information on health-related matters. You can find your nearest PALS office on the NHS Choices website.

Will my taking part in the study be kept confidential?

We will follow ethical and legal practice and all information about you will be handled in confidence.

If you join the study, we will use information collected from you and your medical records during the course of the research. This information will be kept **strictly confidential**, stored in a secure and locked office, and on a password-protected database at the University of Nottingham. Under UK Data Protection laws the University is the Data Controller (legally responsible for the data security) and the Chief Investigator of this study (named above) is the Data Custodian (manages access to the data). This means we are responsible for looking after your information and using it properly. Your rights to access, change or move your information are limited as we need to manage your information in specific ways to comply with certain laws and for the research to be reliable and accurate. To safeguard your rights we will use the minimum personally – identifiable information possible. You can find out more about how we use your information and to read our privacy notice at:
https://www.nottingham.ac.uk/utilities/privacy.aspx.

Where possible, information about you will have your name and address removed and a unique code will be used so that you cannot be recognised from it, however sometimes we need to ensure that we can recognise you to link the research data with your hair sample so in these instances we will need to know your name and date of birth.

Your contact information will be kept by the University of Nottingham for 3 months after the end of the study so that we are able to contact you about the findings of the study and possible follow-up studies (unless you advise us that you do not wish to be contacted). This information will be kept separately from the research data collected and only those who need to will have access to it. All other data (research data) will be kept securely for 7 years. After this time your data will be disposed of securely. During this time all precautions will be taken by all those involved to maintain your confidentiality, only members of the research team given permission by the data custodian will have access to your personal data.

In accordance with the University of Nottingham's, the Government's and our funders' policies we may share our research data with researchers in other Universities and organisations, including those in other countries, for research in health and social care. Sharing research data is important to allow peer scrutiny, re-use (and therefore avoiding duplication of research) and to understand the bigger picture in particular areas of research. Data sharing in this way is usually anonymised (so that you could not be identified) but if we need to share identifiable information we will seek your consent for this and ensure it is secure. You will be made aware then if the data is to be shared with countries whose data protection laws differ to those of the UK and how we will protect your confidentiality.

What will happen if I don't want to carry on with the study?

Your participation is voluntary and you are free to withdraw at any time, without giving any reason, and without your legal rights being affected.

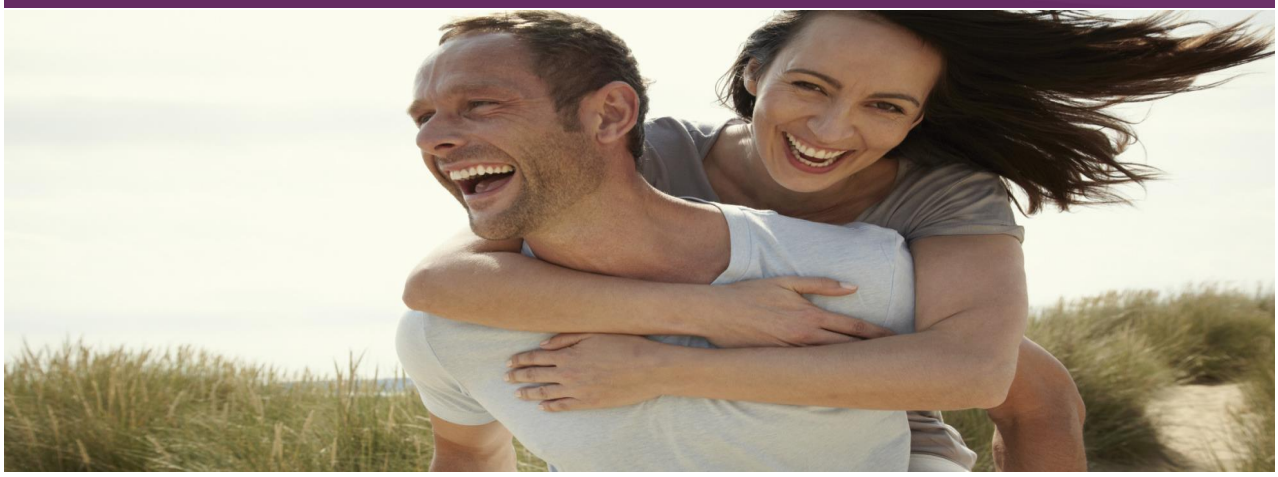
However, we will keep the information about you that we have already obtained as we are not allowed to tamper with study records and this information may have already been used in some analyses and may still be used in the final study analyses. To safeguard your rights, we will use the minimum personally-identifiable information possible.

What will happen to any samples I give?

Your hair sample will be taken and analysed in Division of Child Health, obstetrics and Genecology Laboratory to determine your cortisol levels of the previous 3 months. We would also like to seek your consent so that any remaining samples may be stored and used in possible future research – this is optional (please indicate you agree to this on the consent form). The samples will be stored with a code unique to you and securely at the University of Nottingham. Some of these future studies may be carried out by researchers other than current team of researchers who ran the first study, including researchers working for commercial companies. Any samples or data used will be anonymised, and you will not be identified in anyway. If you do not agree to this any remaining samples will be disposed of.

What will happen to the results of the research study?

The study results will be presented and published in national/international meetings or journals, but you will not be referred to by name or identified in any report or publication. If you wish, you will be provided with an overview of the results at the end of the study which we hope you will find interesting.



Who is organising and funding the research?

This research is being organised by the University of Nottingham, Division of Primary Care and is being funded by the University of Nottingham**.**

Who has reviewed the study?

All research in the NHS is looked at by independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by Newcastle and North Tyneside 1 [REC REF 17/NE/0278 IRAS REF 229947] NHS Research Ethics Committee

Further information and contact details
Chief Investigator:

Prof. Kavita Vedhara
Room 1305 Tower Building
University Park
Nottingham
NG7 2RD
UK
Tel: 0115 846 6931
kavita.vedhara@nottingham.ac.uk

What should I do next?

Following your consents appointment you will have the opportunity to have a brief chat with the researchers and decide if you want to take part. But, if you would like to speak to a member of the research team beforehand please call 0115 82 30511.

.

.

B.2 Participant Consent Form

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

CONSENT FORM
(Final version 1.0: 281117)

Title of Study:

Investigating the association between cortisol and fertility outcomes in patients undergoing in vitro fertilisation (IVF) and intrauterine insemination (IUI) treatment.

Short title: Does the hormone cortisol affect fertility?

IRAS Project ID: 229947

Name of Participant: **Please initial box**

1. I confirm that I have read and understand the information sheet version Final Version 1.0 dated 281117 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, and without my medical care or legal rights being affected. I understand that should I withdraw then the information collected so far cannot be erased and that this information may still be used in the project analysis. ☐

3. I understand that relevant sections of my medical notes and data collected in the study may be looked at by authorised individuals from the University of Nottingham, the research group and regulatory authorities where it is relevant to my taking part in this study. I give permission for these individuals to have access to these records and to collect, store, analyse and publish information obtained from my participation in this study. I understand that my personal details will be kept confidential. ☐

4. I understand and agree that a hair sample will be taken for analysis of cortisol. ☐

5. Consent for storage and use in possible future research (Optional)

   a) I agree that the hair sample I have provided can be stored by the University of Nottingham at the Division of Child Health, Obstetrics and Gynaecology for possible use in future studies. I understand this sample will be anonymised. ☐

   b) I also agree that all other information I have provided (e.g., lifestyle information) can be stored by the University of Nottingham for a period of 7 years after the research has ended. I understand these data will be anonymised. ☐

   c) I understand that some of these studies may be carried out by researchers other than the current team who ran the first study, including researchers working for commercial companies. Any samples or data used will be anonymised, and I will not be identified in anyway. ☐

The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

7. I agree to take part in the above study and confirm that I am not participating in any other research related to my fertility treatment at this time. ☐

8. I would like to receive an overview of the study findings.

YES ☐

NO ☐

_____	_____	_____
Name of Participant	Date	Signature
_____	_____	_____
Name of Person taking consent	Date	Signature

3 copies: 1 for participant, 1 for the project notes and 1 for the medical notes

B.3 Ethics approval letter: Research Ethics Committee

Health Research Authority
North East - Newcastle & North Tyneside 1 Research Ethics Committee
HRA Jarrow
Jarrow Business Centre
Room 001
Rolling Mill Road
Jarrow
NE32 3DT

Telephone: 0207 1048084

Please note: This is the favourable opinion of the REC only and does not allow you to start your study at NHS sites in England until you receive HRA Approval

10 August 2017

Professor Kavita Vedhara
Division of Primary Care
School of Medicine
Tower Building (13th floor, Room1305)
Nottingham University
Nottingham
NG7 2RD

Dear Professor Vedhara

Study title:	**Examining the association between cortisol and treatment outcome in In Vitro Fertilisation (IVF) and intrauterine insemination (IUI) treatment**
REC reference:	**17/NE/0278**
IRAS project ID:	**229947**

The Proportionate Review Sub-committee of the North East - Newcastle & North Tyneside 1 Research Ethics Committee reviewed the above application on 7 August 2017.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all studies that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact hra.studyregistration@nhs.net outlining the reasons for your request. Under very limited circumstances (e.g. for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the study.

Ethical opinion

On behalf of the Committee, the Sub-Committee gave a **Favourable** ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

B.4 Ethics approval letter: Health Research Authority

Health Research Authority

Prof Kavita Vedhara
Division of Primary Care,
School of Medicine
Tower Building (13th floor, Room1305),
NG7 2RD

Email: hra.approval@nhs.net

14 September 2017

Dear Prof Vedhara

Letter of HRA Approval

Study title:	**Examining the association between cortisol and treatment outcome in In Vitro Fertilisation (IVF) and intrauterine insemination (IUI) treatment**
IRAS project ID:	**229947**
REC reference:	**17/NE/0278**
Sponsor	**University of Nottingham**

I am pleased to confirm that **HRA Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications noted in this letter.

Participation of NHS Organisations in England

The sponsor should now provide a copy of this letter to all participating NHS organisations in England.

Appendix B provides important information for sponsors and participating NHS organisations in England for arranging and confirming capacity and capability. **Please read *Appendix B* carefully**, in particular the following sections:

- *Participating NHS organisations in England* – this clarifies the types of participating organisations in the study and whether or not all organisations will be undertaking the same activities
- *Confirmation of capacity and capability* - this confirms whether or not each type of participating NHS organisation in England is expected to give formal confirmation of capacity and capability. Where formal confirmation is not expected, the section also provides details on the time limit given to participating organisations to opt out of the study, or request additional time, before their participation is assumed.
- *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* - this provides detail on the form of agreement to be used in the study to confirm capacity and capability, where applicable.

Further information on funding, HR processes, and compliance with HRA criteria and standards is also provided.

Appendix C

Questionnaires

C.1 Lifestyle and mood questionnaires



Does the hormone cortisol affect fertility? (T1)

INTRODUCTION TO THE QUESTIONNAIRES

Thank you for taking part in this research. Please read these instructions carefully.

Before you start your treatment we would like to ask you some questions about your lifestyle. There are two questionnaires in your research pack. Each questionnaire should only take around 10-20 minutes of your time. You can complete the questionnaires at the clinic or in your own time, either online or by completing paper versions and returning them in the prepaid envelope provided. If you have any questions or you feel anything isn't clear then please ask us. Contact details are provided in your information leaflet.

Please complete 'pre-treatment questionnaire 1' as soon as possible following today's appointment. Please complete 'pre-treatment questionnaire 2' on the first day of your treatment. We will send you a reminder to complete the second questionnaire.

We understand that the experience of fertility difficulties can be upsetting, and talking about issues surrounding these difficulties can be too. In this questionnaire, we use terms that you may feel are upsetting or do not reflect your current situation (e.g. 'infertile' or 'my Infertility'). Please remember that taking part in the study is optional and you can withdraw at any time. We also have a counsellor you can talk to if completing this questionnaire causes you distress.

There are no right or wrong answers. Please answer each question as honestly as you can.

Basic Information

1. Please insert your Study Identity Code in the following format: first name initial + last name initial + date of birth [ddmmyyyy]. For example, if your name is John Smith and you were born on the 29th March 1986 you would enter JS29031986

Your answer should be no more than 12 characters long.

2. Date

Dates need to be in the format 'DD/MM/YYYY', for example 27/03/1980.

(dd/mm/yyyy)

About your lifestyle

Below are several questions about **lifestyle choices** some people make. Please select the appropriate answer for each question. Your answers will be kept **confidential** and will be used for this research only.

3. Have you ever used class A drugs in the past (e.g. heroin, cocaine, ecstasy)?

- YES
- NO

4. Do you smoke marijuana frequently (more than four times per week)?

- YES
- NO

5. To your knowledge, have you ever had a sexually transmitted infection?

- YES
- NO

6. Are you currently taking anabolic steroids (for non-medical uses)?

- YES
- NO

How do you feel at the moment in relation to your infertility?

7. Are you single patient or PGD patient (i.e., not infertile)? ✱ *Required*

- Yes
- No

A number of statements which people have used to describe themselves are given below. Read each statement and then select the most appropriate response for each statement to indicate **how you feel RIGHT NOW at this moment in relation to your infertility.** 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 your present feelings best.

8. I feel calm.

- Not at all
- Somewhat
- Moderately
- Very much

9. I feel tense.

- Not at all
- Somewhat
- Moderately
- Very much

10. I am upset.

- Not at all
- Somewhat
- Moderately
- Very much

11. I feel relaxed.

- Not at all
- Somewhat
- Moderately
- Very much

12. I feel content.

- Not at all
- Somewhat
- Moderately
- Very much

13. I am worried.

- Not at all
- Somewhat
- Moderately
- Very much

Your thoughts about IVF and infertility DURING THE PAST SEVEN DAYS

I**nfertility** and its **treatment** can often be stressful.

Below are comments made by people in relation to infertility and its treatment. Please read each item and select how frequently it applied to your thoughts about infertility **during the past seven days**. Please select "not at all" if it did not occur during this time.

14. I thought about my infertility when I didn't mean to.

- Not at All
- Rarely
- Sometimes
- Often

15. I avoided letting myself get upset when I thought about my infertility or was reminded of it.

- Not at All
- Rarely
- Sometimes
- Often

16. I tried to remove my infertility from my memory.

- Not at All
- Rarely
- Sometimes
- Often

17. I had trouble falling asleep or staying asleep, because of pictures or thoughts about my infertility that came into my mind.

- Not at All
- Rarely
- Sometimes
- Often

18. I had waves of strong feelings about my infertility.

- Not at All

- Rarely
- Sometimes
- Often

19. I had dreams about my infertility.

- Not at All
- Rarely
- Sometimes
- Often

20. I stayed away from reminders of my infertility.

- Not at All
- Rarely
- Sometimes
- Often

21. I felt as if my infertility hadn't happened or my infertility wasn't real.

- Not at All
- Rarely
- Sometimes
- Often

22. I tried not to talk about my infertility.

- Not at All
- Rarely
- Sometimes
- Often

23. Images regarding my infertility popped into my mind.

- Not at All
- Rarely
- Sometimes
- Often

24. Other things kept making me think about my infertility.

- Not at All
- Rarely
- Sometimes
- Often

25. I was aware that I still had a lot of feelings about my infertility, but I didn't deal with them.

- Not at All
- Rarely
- Sometimes
- Often

26. I tried not to think about my infertility.

- Not at All
- Rarely
- Sometimes
- Often

27. Any reminder brought back feelings about my infertility.

- Not at All
- Rarely
- Sometimes
- Often

28. I feel numb when thinking about my infertility.

- Not at All
- Rarely
- Sometimes
- Often

How you have felt or behaved during the PAST WEEK?

Below is a list of the ways you might have felt or behaved. Please tell us how often you have felt this way **during the past week** by selecting the statement that suits you most.

29. I was bothered by things that usually don't bother me.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

30. I did not feel like eating; my appetite was poor.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

31. I felt that I could not shake off the blues even with help from my family or friends.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

32. I felt I was just as good as other people.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

33. I had trouble keeping my mind on what I was doing.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

34. I felt depressed.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

35. I felt that everything I did was an effort.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

36. I felt hopeful about the future.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

37. I thought my life had been a failure.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

38. I felt fearful.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

39. My sleep was restless.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)

- Most or all of the time (5-7 days)

40. I was happy.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

41. I talked less than usual.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

42. I felt lonely.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

43. People were unfriendly.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

44. I enjoyed life.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

45. I had crying spells.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

46. I felt sad.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

47. I felt that people dislike me.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

48. I could not get “going”.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

The next part of the questionnaire lists a number of words that people often use to describe different feelings and emotions.

Read each item and then select the appropriate statement to indicate to what extent you generally feel this way, that is, **how you feel on the average**.

49. Upset

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

50. Hostile

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

51. Alert

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

52. Ashamed

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

53. Inspired

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

54. Nervous

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

55. Determined

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

56. Attentive

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

57. Afraid

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

58. Active

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

59. Distressed

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

60. Interested

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

61. Guilty

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

62. Strong

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

63. Jittery

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

64. Enthusiastic

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

65. Irritable

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

66. Proud

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

67. Scared

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

68. Excited

- Very slightly or not at all
- A little
- Moderately
- Quite a bit
- Extremely

Your sleep habits during the PAST MONTH ONLY.

The following questions relate to your usual sleep habits **during the past month only**. Please note that the answer should indicate your **typical** sleeping habit in the past month. Please answer all questions.

69. During the past month, thinking of a typical day, what time did you go to bed?

70. During the past month, thinking of a typical day, how long (in minutes) did it take you to fall asleep that night?

71. During the past month, thinking of a typical day, what time did you wake up in the morning?

72. During the past month, thinking of a typical day, how many hours of actual sleep did you get that night?

73. During the past month, how many hours were you in bed on a typical night?

The following questions (from 52 to 61) are going to ask you about **How Often** you have had trouble sleeping due to different REASONS.

Please select the answer that suits you most.

74. I could not get to sleep within 30 minutes.

- Not during the past month (0)

- Less than once a week (1)
- Once or twice a week (2)
- Three or more times a week (3)

75. I woke up in the middle of the night or early morning.

- Not during the past month (0)
- Less than once a week (1)
- Once or twice a week (2)
- Three or more times a week (3)

76. I had to get up to use the bathroom.

- Not during the past month (0)
- Less than once a week (1)
- Once or twice a week (2)
- Three or more times a week (3)

77. I could not breathe comfortably.

- Not during the past month (0)
- Less than once a week (1)
- Once or twice a week (2)
- Three or more times a week (3)

78. I coughed or snored loudly.

- Not during the past month (0)
- Less than once a week (1)
- Once or twice a week (2)
- Three or more times a week (3)

79. I felt too cold.

- Not during the past month (0)
- Less than once a week (1)
- Once or twice a week (2)
- Three or more times a week (3)

80. I felt too hot.

- Not during the past month (0)
- Less than once a week (1)
- Once or twice a week (2)
- Three or more times a week (3)

81. I had bad dreams.

- Not during the past month (0)
- Less than once a week (1)
- Once or twice a week (2)
- Three or more times a week (3)

82. I had pain.

- Not during the past month (0)
- Less than once a week (1)
- Once or twice a week (2)
- Three or more times a week (3)

83. Is there any other reason(s) that you have trouble sleeping?

- Yes
- No

83.a. Please describe the reason(s):

83.b. Please indicate how often you have had trouble sleeping because of this reason(s).

- Not during the past month (0)
- Less than once a week (1)
- Once or twice a week (2)
- Three or more times a week (3)

84. During the past month, how often have you taken medicine (prescribed or “over the counter”) to help you sleep?

- Not during the past month (0)
- Less than once a week (1)
- Once or twice a week (2)
- Three or more times a week (3)

85. During the past month, how often have you had trouble staying awake while driving, eating meals, or engaging in social activities?

- Not during the past month (0)
- Less than once a week (1)
- Once or twice a week (2)
- Three or more times a week (3)

86. During the past month, how much of a problem has it been for you to keep up the enthusiasm to get things done?

- Not during the past month (0)
- Less than once a week (1)
- Once or twice a week (2)
- Three or more times a week (3)

87. During the past month, how would you rate your sleep quality overall?

- Very good (0)
- Fairly good (1)
- Fairly bad (2)
- Very bad (3)

The time you spend doing different types of physical activities in a typical week

In the next section, we are going to ask you some questions about the **time you spend doing different types of physical activity in a typical week**. Please answer these questions even if you do not consider yourself to be a physically active person.

Think first about the time you spend doing work. Think of work as the things that you have to do such as paid or unpaid work, study/training, household chores etc. In answering the following questions, ***'vigorous-intensity activities'*** are activities that require hard physical effort and cause large increases in breathing or heart rate, ***'moderate-intensity activities'*** are activities that require moderate physical effort and cause small increases in breathing or heart rate.

88. Does your work involve vigorous-intensity activity that causes large increases in breathing or heart rate for **at least 10 minutes continuously**? (e.g., carrying or lifting heavy loads, digging or construction work)

- ○ Yes
- ○ No
- ○ Other

88.a. **If you selected other, please specify:**

88.b. In a typical week, on how many days do you do vigorous- intensity activities as part of your work? (Number of days)

88.c. In a typical week, on how many days do you do vigorous- intensity activities as part of your work? (Hours:Minutes)

89. Does your work involve moderate-intensity activity that causes small increases in breathing or heart rate for **at least 10 minutes continuously**? (e.g., brisk walking or carrying light loads)

- ○ Yes
- ○ No
- ○ Other

89.a. **If you selected other, please specify:**

89.b. In a typical week, on how many days do you do moderate- intensity activities as part of your work? (Number of days)

89.c. How much time do you spend doing moderate-intensity activities at work on a typical day? (Hours:Minutes)

Travel to and from places

The next question(s) exclude the physical activities at work that you have already mentioned. Now we would like to ask you about the usual way you travel to and from places. For example to work, for shopping, to market, to your place of worship.

90. Do you walk or use a bicycle (pedal cycle) for at least 10 minutes continuously to get to and from places?

- Yes
- No
- Other

90.a. **If you selected other, please specify:**

90.b. In a typical week, on how many days do you walk or bicycle for at least 10 minutes continuously to get to and from places? (Number of days)

90.c. How much time do you spend walking or bicycling for travel on a typical day? (Hours:Minutes)

Recreational activities

The next questions **exclude** the work and transport activities that you have already mentioned. Now we would like to ask you about sports, fitness and recreational activities (leisure).

91. Do you do any vigorous-intensity sports, fitness or recreational (leisure) activities that cause large increases in breathing or heart rate for **at least 10 minutes continuously**? (e.g., running or football)

- Yes
- No
- Other

91.a. **If you selected other, please specify:**

91.b. In a typical week, on how many days do you do vigorous- intensity sports, fitness or recreational (leisure) activities? (Number of days)

91.c. How much time do you spend doing vigorous-intensity sports, fitness or recreational activities on a typical day? (Hours:Minutes)

92. Do you do any moderate-intensity sports, fitness or recreational (leisure) activities that cause a small increase in breathing or heart rate for **at least 10 minutes continuously**? (e.g., brisk walking, cycling, swimming, volleyball)

- Yes
- No
- Other

92.a. **If you selected other, please specify:**

92.b. In a typical week, on how many days do you do moderate- intensity sports, fitness or recreational (leisure) activities? (Number of days)

92.c. How much time do you spend doing moderate-intensity sports, fitness or recreational (leisure) activities on a typical day? (Hours:Minutes)

Sedentary behaviour

The following question is about sitting or reclining at work, at home, getting to and from places, or with friends. This should include time spent sitting at a desk, sitting with friends, travelling in car/bus/train, reading, playing cards or watching television. Please **do not include** time spent sleeping.

93. How much time do you usually spend sitting or reclining on a typical day? (Hours:Minutes)

Alcohol consumption

Because alcohol use can affect your health and can interfere with certain medications and treatments, it is important that we ask some questions about your use of alcohol **during the last month**. Your answers will remain confidential so please be honest with the answers.

PLEASE SELECT THE STATEMENT THAT APPLIES TO YOU.

94. How often do you have a drink containing alcohol?

- Never
- Monthly or less
- 2-4 times a month
- 2-3 times a week
- 4 or more times a week

95. How many drinks containing alcohol do you have on a typical day when you are drinking?

- 1 or 2
- 3 or 4
- 5 or 6
- 7 to 9
- 10 or more
- Not at all

96. How often do you have six or more drinks on one occasion?

- Never
- Less than monthly
- Monthly
- Weekly
- Daily or almost daily

97. How often during the last month have you found that you were not able to stop drinking once you had started?

- Never
- Less than monthly
- Monthly
- Weekly
- Daily or almost daily

98. How often during the last year have you failed to do what was normally expected of you because of drinking?

- Never
- Less than monthly
- Monthly
- Weekly
- Daily or almost daily

99. How often during the last year have you needed a first drink in the morning to get yourself going after a heavy drinking session?

- Never
- Less than monthly
- Monthly
- Weekly
- Daily or almost daily

100. How often during the last year have you had a feeling of guilt or remorse after drinking?

- Never
- Less than monthly
- Monthly
- Weekly
- Daily or almost daily

101. How often during the last year have you been unable to remember what happened the night before because of your drinking?

- Never
- Less than monthly
- Monthly
- Weekly
- Daily or almost daily

102. Have you or someone else been injured because of your drinking?

- Never
- Yes, but not in the last year
- Yes, during the last year

Smoking and Caffeine consumption

103. Do you smoke?

- YES
- NO

103.a. How many cigarettes on average do you smoke per day.

103.b. What brand of cigarettes do you use?

In the next section, we are going to ask you some questions about your intake of foods and drinks that contain caffeine. For these foods and drinks, we ask you how much you eat or drink daily, weekly, and monthly. This is because we are not only interested in how much you eat or drink in a day, but also the patterns of how you consume these foods and drinks. Therefore, please write down how much you consume per day, per week, and per month. **Please do not simply add up the number.** It is fine if these numbers do not add up. We're just trying to capture the actual pattern of your consumption of these foods and drinks.

For the following questions, you also need to do the following:

1. Write **brand details** for *coffee, hot chocolate, and tea*, and *other Drinks* in the space provided;
2. Write down which **type and brand** of coffee you drink;
3. Tick the 'Decaf' box if you had a decaffeinated drink;
4. Tick the 'Diet' box if you had any diet soft drink;
5. Tick the 'Full size' or 'Mini bar' box to indicate the portion of chocolate bars you eat;
6. Tick the 'None' box if you do not eat or drink this product.

104. Coffee, Hot chocolate, and Tea

	Brand	Decaf	Normal	How many cups or mugs do you drink per day?	How many cups or mugs do you drink per week?	How many cups or mugs do you drink per month?	None
Instant Coffee							
Filter or cafetiere Coffee							
Espresso							

Hot chocolate		○	○				○
Tea bags		○	○				○
Tea leaves (loose tea)		○	○				○
Instant tea		○	○				○

104.a.

	Type (e.g. latte, mocha etc.)	Brand (e.g. Starbucks or Costa)	Decaf	Normal	How many cups or mugs do you drink per day?	How many cups or mugs do you drink per week?	How many cups or mugs do you drink per month?	None
Any other coffee (e.g. latte, mocha, cappuccino, Americano etc.)			○	○				○

105. Other drinks

	Brand	Diet	Normal	Bottle/can/glass	How many bottle/can/glass do you drink **per day**?	How many bottle/can/glass do you drink **per week**?	How many bottle/can/glass do you drink **per month**?	None
Iced tea (bottle)		○	○					○
Cola (Not Dr Pepper)		○	○					○
Energy drinks		○	○					○

105.a.

	Diet	Normal	Bottle/can/glass	How many bottle/can/glass do you drink **per day**?	How many bottle/can/glass do you drink **per week**?	How many bottle/can/glass do you drink **per month**?	None
Dr Pepper	○	○					○
Iron-Bru	○	○					○
Lucozade	○	○					○
Mountain Dew	○	○					○

106. Chocolate bars

	Portion		How many bars do you eat **per day**?	How many bars do you eat **per week**?	How many bars do you eat **per month**?	None
	Full size	Mini bar				

Energy bar (e.g. Boost)	○	○				○
Milk chocolate coated bars (e.g. Twix, Mars etc.)	○	○				○
Milk chocolate bars	○	○				○
Plain chocolate bars	○	○				○

107. Other chocolate products

	How much do you eat **per day**?	How much do you eat **per week**?	How much do you eat **per month**?	None
Chocolate coated biscuits				○
Chocolate brownie, cake				○
Chocolate mousse/ice-cream				○
Chocolate milk (shake) glass				○

Hair sampling questionnaire

Now we would like to ask you a few questions about your hair. This is to ensure that we allow for the influence that certain products or treatment(s) may have had when we assess the amount of cortisol in your hair.

108. What is your natural hair colour?

109. Do you dye your hair?

- Yes
- No

109.a. Has your hair been dyed in the previous 3 months?

- Yes
- No

110. How often do you wash your hair per week?

111. Have you had any other treatment(s) (e.g. perm/chemical treatments for straightening) done on your hair in the previous 3 months?

- Yes
- No

111.a. Please state what treatment(s) you had:

112. What, if any, hair products (e.g. hair spray, wax, gel etc.) have you used on your hair in the previous 3 months?