University of Groningen

Can hormones get you down?

de Wit, Anouk

DOI:
[10.33612/diss.178483963](https://doi.org/10.33612/diss.178483963)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
de Wit, A. (2021). *Can hormones get you down? studies on sex steroids and depression*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. <https://doi.org/10.33612/diss.178483963>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Can hormones get you down?

Studies on sex steroids and depression

Anouk Elisabeth de Wit

Can hormones get you down?

PhD thesis, University of Groningen, The Netherlands

Cover design: Loutje Hoekstra

Lay-out: Publiss | www.publiss.nl

Print: Ridderprint | www.ridderprint.nl

Gedeon Richter BENELUX

© Copyright 2021: Anouk Elisabeth de Wit

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, by photocopying, recording, or otherwise, without the prior written permission of the author.

university of
groningen

Can hormones get you down?

Studies on sex steroids and depression

PhD thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. C. Wijmenga
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Wednesday 22 September 2021 at 16.15 hours

by

Anouk Elisabeth de Wit

born on 4 August 1993
in Zeist, the Netherlands

Supervisor

Prof. R.A. Schoevers

Co-supervisors

Dr. M.K. de Boer

Dr. E.J. Giltay

Assessment Committee

Prof. E.F.C. van Rossum

Prof. N.J.A. van der Wee

Prof. D.J. Touw

Table of contents

Chapter 1	General Introduction	9
<i>Part I: Endogenous sex steroids and depression</i>		
Chapter 2	Testosterone in human studies: modest associations between plasma and salivary measurements	29
Chapter 3	Plasma androgens and the presence and course of depression in a large cohort of women	45
Chapter 4	Plasma androgens and the presence and course of depression in a large cohort of men	63
Chapter 5	Associations between testosterone and metabolic syndrome in depressed and non-depressed older men and women	81
Chapter 6	Associations of plasma androgens with suicidality among men and women: a 9-year longitudinal cohort study	99
Chapter 7	Impact of estradiol variability and progesterone on mood in perimenopausal women with depressive symptoms	115
Chapter 8	Predictors of irritability symptoms in mildly depressed perimenopausal women	131
<i>Part II: Hormonal contraceptives and depression</i>		
Chapter 9	Association of use of oral contraceptives with depressive symptoms among adolescents and young women	151
Chapter 10	Association between adolescent oral contraceptive use and future major depressive disorder: a prospective cohort study	167
Chapter 11	Hormonal contraceptive use and depressive symptoms: systematic review and network meta-analysis of randomized clinical trials	183
Chapter 12	Efficacy of oral contraceptive pills for depressive symptoms and overall symptomatology in premenstrual syndrome: pairwise and network meta-analysis of randomized trials	203

Part III: Researchers' role in science journalism

Chapter 13	From academia to society – improving the quality of science journalism	223
Chapter 14	Discussion	227
References		245
Supplements		281
Summary in Dutch / Nederlandse samenvatting		337
List of publications		353
Acknowledgements / Dankwoord		357
Curriculum Vitae		365

Abbreviations

5 α -DHT	5 α -dihydrotestosterone
ATC	World Health Organization Anatomical Therapeutic Chemical
AUC	Area under the curve
B	Beta-coefficient
β	Standardized beta-coefficient
BMI	Body mass index
CI	Confidence interval
CIDI	Composite International Diagnostic Interview
CrI	Credible interval
DHEAS	Dehydroepiandrosterone-sulphate
DSM	Diagnostic and Statistical Manual of Mental Disorders
FSH	Follicle-stimulating hormone
GABA	Glutamatergic and γ -aminobutyric acid
GEE	Generalized estimating equation
GRADE	Grading of Recommendations Assessment, Development, and Evaluation
HDL	High density lipoprotein
HPA-axis	Hypothalamic-pituitary-adrenal-axis
HPG-axis	Hypothalamic-pituitary-gonadal-axis
IQR	Interquartile range
LC-MS/MS	Liquid chromatography–tandem mass spectrometry
LEIDS-R	Revised Leiden Index of Depression Sensitivity
LH	Luteinizing hormone
LIDAS	Lifetime Depression Assessment Self-Report
MADRS	Montgomery-Åsberg Depression Rating Scale
MDD	Major depressive disorder
Mets	Metabolic syndrome
NESDA	Netherlands Study on Depression and Anxiety
NESDO	Netherlands Study on Depression in Older persons
OC	Oral contraceptive pill
OR	Odds ratio
PMDD	Premenstrual dysphoric disorder
PMS	Premenstrual syndrome
SD	Standard deviation
SHBG	Sex hormone binding globulin
SMD	Standardized mean difference
TRAILS	TRacking Adolescents' Individual Lives Survey
VMS	Vasomotor symptoms

Chapter 1

General Introduction

Depression

Depression is one of the most prevalent mental disorders, affecting millions of men and women worldwide.¹ It afflicts about 10% to 20% of people at some time in their life,^{2,4} and because of its often early onset,⁵ high relapse rate⁶ and the associated risk for suicide,⁷ depression has a high impact on the person affected and his or her environment.⁸

By definition, patients with depression suffer for at least two weeks from symptoms, including feeling depressed or anhedonia, that cause a significant degree of disability.⁹ However, people who suffer from depressive symptoms but do not meet these criteria, can still experience significant impairment.¹⁰⁻¹² When studying depression, it is therefore important to not only investigate the diagnosis, but also subsyndromal symptoms.

Sex differences

A consistent finding in epidemiological studies is that the prevalence of depression is twice as high in women as in men.^{1, 2} This sex disparity is noticeable across cultures and societies,¹³ and is most profound during women's reproductive years (see also *Figure 1.1*).¹⁴ Throughout these years women are especially at risk for depression during hormonal transition phases such as adolescence,^{15, 16} late pregnancy to postpartum,^{17, 18} and perimenopause.¹⁹⁻²²

Figure 1.1. Prevalence of depression across the life span in men and women.

The graph is based on several epidemiological studies.^{4, 19, 22-26} As a consequence of the wide variety of definitions of depression used, the scale on the y-axis is lacking.

Heterogeneity

Despite the availability of several treatment options, a third of the patients are treatment resistant.²⁷ This is likely due to the heterogeneity in mechanisms underlying the disorder. Examples of these mechanisms include genetic abnormalities, neurotransmitter deficiencies, disturbance of the hypothalamic-pituitary-adrenal (HPA)-axis, traumatic life experiences, low social support, and personality characteristics such as neuroticism.^{6,28-30} Prompted by the sex difference in vulnerability to depression, another presumed underlying mechanism is the influence of sex steroids. However, essential questions on the exact relationship between sex steroids and depression remain unanswered.

Central theme of this thesis

The aim of the research reported in this thesis was to define more precisely the association between sex steroids and depression. A better insight in these associations will provide cues for who are vulnerable to the effects of these hormones and, consequently, who may benefit from hormonal treatment options. To obtain this goal, the research reported for this thesis was conducted with data from four different well-phenotyped cohorts. Three of these cohorts were large population studies that followed participants for years. These cohorts focused either on adolescents (Tracking Adolescents' Individual Lives Survey [TRAILS]), adults (Netherlands Study on Depression and Anxiety [NESDA]), or older people (Netherlands Study on Depression in Older persons [NESDO]). The fourth cohort was an in-depth study of perimenopausal women (PERI). The perimenopause is characterized by frequent and unpredictable changes in sex steroids. Therefore, the women from this cohort were studied for a shorter period of time, but with more frequent assessments that would therefore better capture the dynamics of sex steroids. Altogether, the data from these cohorts provided the opportunity to test for several assumed, but often not thoroughly studied hypotheses on the associations between sex steroids and depression across different phases in life.

Below the next header "Sex steroids" more general background information will be given on sex steroids and their effects on the brain. This will provide readers who are not familiar with sex steroids a better understanding as to why these hormones may relate to depression. After the general introduction on sex steroids, more information on the current knowledge about the associations between sex steroids and depression will be provided below the two headers "Associations

between endogenous sex steroids and depression” and “Associations between hormonal contraceptives and depression”. For sake of simplicity, “depression” will be used in this thesis as an overarching term for everything from depressive symptoms to depression. If findings are specific to a certain disorder or symptom domain, the specific term will be used. In case of depression, this will be major depressive disorder (MDD). Thereafter, the role of researchers in science journalism will be discussed below the header “The role of researchers in science journalism”. This may seem out of topic at first sight, but it will become clear that good quality of science journalism is of crucial importance to the ultimate goal of science: to help society progress. Finally, the importance of the studies conducted in this thesis are introduced in more detail below the header “Aims of this thesis”.

Sex steroids

Endogenous sex steroids

Sex steroids are powerful substances that shape both body and mind. The most well-known compounds of sex steroids are estradiol, progesterone, and testosterone, yet many others are also involved. Sex steroids are primarily responsible for sexual differentiation, development and maintenance of secondary sex characteristics, and for reproduction.³¹ However, they are also vital players in maintaining the metabolic balance in the body,^{31, 32} and of paramount importance for this thesis, they also play an important role in processes in the brain.³³⁻³⁷

The synthesis of sex steroids occurs at multiple sites in the body, but sex steroids are predominantly synthesized in the gonads. This synthesis is primarily controlled by the hypothalamic-pituitary-gonadal (HPG)-axis (see also *Figure 1.2*). Another important peripheral production site are the adrenal glands where androgens, and to lesser extent estradiol, are also synthesized. This production is not under the control of luteinizing hormone (LH) and follicle stimulating hormone (FSH) but under the control of adrenocorticotrophic hormone as part of the HPA-axis.^{38, 39} Peripherally produced sex steroids enter the circulatory system so they can exert their effects throughout the body. In the circulatory system, sex steroids are mainly bound to plasma proteins such as sex hormone binding globulin (SHBG) or corticosteroid binding globulin.⁴⁰ SHBG binds to estrogens and androgens, but progesterone is bound by corticosteroid binding globulin. As the majority of sex steroids are bound, only a small fraction of these hormones is “free”. Yet, due to their lipophilicity, this small fraction of “free” sex steroids can easily transfer over

the blood-brain-barrier and exert effects in the brain.⁴¹ Apart from the peripheral synthesis of sex steroids, the brain itself can also synthesize sex steroids because it possesses the enzymes needed for this synthesis.⁴² Sex steroids produced in the brain include estradiol, testosterone and progesterone, as well as others such as pregnenolone, dehydroepiandrosterone and allopregnanolone.⁴³ Once sex steroids are active in the brain (either locally produced or transferred from the periphery), they are called "neurosteroids".⁴⁴

Hormonal contraceptives

Besides the endogenous production of sex steroids, an important exogenous source of sex steroids is the intake of synthetic steroids. An example includes the use of hormonal contraceptives. Hormonal contraceptives are a group of medications that are taken by means of preventing pregnancies or diminishing menstrual symptoms. The most well-known subgroup of hormonal contraceptives is the oral contraceptive pill (OC) which contain both a synthetic estradiol and a synthetic progestin. Other subgroups of hormonal contraceptives are hormonal intrauterine devices and progestin-only implants (or pills of injections) that only contain a synthetic progestin.

These synthetic variants of sex steroids affect endogenous estradiol, progesterone, and testosterone levels by binding to sex steroid receptors and initiating feedback loops.^{45, 46} For example, the intake of OCs decreases free testosterone levels by increasing hepatic production of SHBG thus binding and decreasing biologically active circulating androgens. Additionally, the use of synthetic progestin and estrogen inhibit the ovarian estradiol and progesterone production by suppressing LH and FSH centrally through negative feedback.⁴⁵ How and to what extent hormonal contraceptives interfere with the endogenous hormonal system depends on the specific formulation used and the route of administration (see also *Figure 1.3*).

Figure 1.2. Simplified negative feedback loop of the hypothalamic-pituitary-gonadal axis.

Abbreviations: DHEA, dehydroepiandrosterone; DHEAS, dehydroepiandrosterone-sulphate; 5 α -DHT, 5 α -dihydrotestosterone. The medial preoptic area of the hypothalamus releases gonadotropin-releasing hormone (GnRH) in response to stimuli from the cerebral cortex and limbic system via various neurotransmitters, such as kisspeptin. Subsequently, GnRH stimulates the secretion of luteinizing hormone (LH) and follicle-stimulating hormone (FSH) by the pituitary gland. LH and FSH in turn stimulate the sex steroid production from cholesterol. LH raises the synthesis and secretion of androgens including testosterone in both men and women by the gonads (testes in men and ovaries in women) and triggers ovulation in women. FSH stimulates the estradiol production by the ovaries in women. Testosterone, estradiol, and progesterone provide feedback to the hypothalamus and pituitary gland that results in adjustment of GnRH, LH and FSH secretion.⁴⁷

Figure 1.3. Impact of hormonal contraceptives on the menstrual cycle.

In users of vaginal rings, oral contraceptive pills, and most of the progesterone only contraceptives the majority of ovulations is inhibited.^{45, 46} In intrauterine device users however, only about 20% of the ovulations is inhibited.^{45, 46}

Effects of sex steroids on the brain

Besides their role in peripheral processes, sex steroids, including hormonal contraceptives, influence myriad processes in the brain. Influences of sex steroids on the brain start early on in fetal development and continue throughout adolescence and adulthood.³³⁻³⁷ Sex steroids exert their effects via sex steroid receptors that are located in almost every part of the central nervous system, but are more predominantly present in some regions such as the limbic system (see also *Figure 1.4*).⁴⁷⁻⁷⁶ A basic understanding of the effects of these sex steroid receptor bindings is important in order to grasp how sex steroids ultimately may be implicated in depression.

Effects of sex steroids binding to sex steroid receptors arise from slow genomic mechanisms (hours to days) in the case that they bind to nuclear receptors to regulate gene expression, or arise from faster non-genomic mechanisms when they bind to membrane receptors (seconds to minutes).^{34-37, 77} Together these bindings result in organizational and activational changes in the brain. Activational effects can temporarily modify neural activity.^{34-37, 77} In contrast, organizational

effects compose lasting alterations of the nervous system, like myelination or dendritic spine remodeling.^{34-37, 77}

The biological outcome of sex steroid bindings in the brain is context dependent.⁷⁸ This explains why there is no single mechanism in which changes by sex steroids can lead to depression. Not only can different sex steroids bind to different sex steroid receptors (with different affinities), the effect of the bindings also depends on the coregulatory proteins. The availability of these coregulators differs between specific tissues, and depending on the chemical modification (such as phosphorylation, methylation) of the coregulators, they have different effects. Moreover, sex steroids also affect and interact with important other regulatory systems that are implicated in depression. For example, testosterone, estradiol and progesterone can alter the presynaptic release of classical neurotransmitters and the responsiveness of postsynaptic receptors to these neurotransmitters.^{36, 48, 77, 79} In this way, sex steroids largely affect the dopaminergic, serotonergic, glutamatergic and γ -aminobutyric acid (GABA)-ergic neurotransmitter systems. Also, estradiol and androgens such as testosterone contribute to the immune response in a dose dependent manner. In general, low levels of estradiol stimulate activation of the immune system, while higher levels of estradiol and testosterone suppress the immune system.⁸⁰ Finally, sex steroids also play a key role in the establishment, activation, and regulation of the HPA-axis; the main stress system in the body.^{38, 39} On top of the basal sex differences in the stress response that have arisen during the initial development of the structures of this system,⁸¹ the HPG- and HPA-axis continuously interact.^{38, 39}

The above-described neurotransmitter system, the stress system, and immune system all play a role in the development and maintenance of depression. Together with the observation that sex steroids can exert direct effects via sex steroid bindings in areas of the limbic system which are important for emotion regulation, this suggests that sex steroids have the potential to affect processes that are of importance to the etiology of depression.⁸² Still, exact mechanisms are hardly understood. Sex steroids will probably have different effects on mood, ranging from a few depressive symptoms to MDD, depending on the context, such as phase of life and circumstances of the individual. More precise insights in how sex steroids associate with depression and how these associations depend on the context, is crucial to obtain insights in what factors increase the vulnerability for the effects of these hormones and ultimately who may benefit from hormonal treatment options. Previous observational studies have already provided some of these insights, as described in the following paragraphs.

Figure 1.4. Location of testosterone, estrogen and progesterone receptors in the intracranial central nervous system.

Dotted areas are located more laterally and therefore not visible in the mid sagittal plane. The figure is composed based on valuable publications of others.⁴⁷⁻⁷⁶ For the purpose of clarity, estrogen receptors are not subdivided by type of receptor (namely α and β). However, it is known that the expression of the specific receptors differ by location and also depend on the age of the individual.⁸³

Associations between endogenous sex steroids and depression

Testosterone

One of the first explanations for the sex difference in the prevalence of depression was that testosterone would serve a protective function for mood.⁸⁴ This idea was rooted in the observation that men have much higher levels of testosterone than women, and at the same time are less often diagnosed with depression. However, this theory is too simplistic given the diverse and complex impact of testosterone on the brain.

A more nuanced explanation for the sex disparity includes that supraphysiological testosterone levels in women and very low testosterone levels in men associate with depression (see also *Figure 1.5*). This theory is based on the fact that depression has a considerably greater prevalence in adult women with medical conditions that are associated with significantly elevated testosterone levels,⁸⁵ while in men, hypogonadal levels of testosterone have been associated with depression.⁸⁶⁻⁸⁸ Although this theory may be plausible, it's unknown whether

there is also an association between testosterone levels in the physiological range and depression. Moreover, the question remains whether the higher prevalence of depression in men with low testosterone levels, could be ascribed to symptoms such as sexual dysfunction which are also often present in these men, or also to core emotional symptoms such as sadness and anhedonia. Finally, previous studies were often limited by the determination of androgens by immunoassays rather than liquid chromatography–tandem mass spectrometry (LC-MS/MS). This is important as immunoassays are less reliable for the determination of lower testosterone levels such as are present in women and in some men.⁸⁹

Figure 1.5. Theoretical model that supraphysiological levels of testosterone in women and low levels in men associate with depression.

The x-axis gives a range of testosterone levels from low to high. Exact values are not given as these values depend on the assay used and are not important for the general idea that relatively high levels in women and relatively low levels in men are both associated with an increased risk for depression.^{85-88, 90, 91} The grey area on the left represents the normal distribution of testosterone levels in women, while the area on the right represents the normal distribution of testosterone levels in men. On average, men have about 14 times higher levels of testosterone than women.⁹² The area in between the two grey areas refers to levels that are relatively high for women and relatively low for men.

Another explanation suggests that depression is not primarily associated with testosterone, but that it increases the vulnerability for other diseases associated with aberrant testosterone levels. Especially at an older age, this may be an important issue as testosterone levels change and comorbidities become more prevalent. For example, some suggest that the higher prevalence of metabolic syndrome among older men and women,⁹³ may be related to altered testosterone levels. Low testosterone levels in men and high levels in women are thought to alter

adipose tissue function, with effects on lipid metabolism, insulin resistance and fat mass expansion. This may induce metabolic disadvantages.⁹⁴ Indeed, earlier research has shown that low testosterone levels in men and high testosterone levels in women may contribute to the development of metabolic syndrome, but that the magnitude of this effect depends on several other factors.⁹⁵⁻⁹⁷ That said, one such factor may be depression, as dysregulations in the immune response and the HPA-axis seen in depression,⁹⁸ closely respond to alterations in the HPG-axis. This not only alters testosterone levels but also impacts its function.^{36, 94, 99} Hence, this would suggest that patients with depression are more likely to develop a metabolic syndrome if they also have aberrant testosterone levels. Nevertheless, this is a hypothesis that has not been tested before.

A final, often suggested but hardly proven idea is that elevated testosterone levels in men and women contribute to increased likelihood of attempting suicide.¹⁰⁰⁻¹⁰⁷ However, the main problem with the current evidence is that most studies measured testosterone levels after the attempt. This measurement is likely biased given the acute stress that accompanies an attempt which will consequently alter testosterone levels.³⁸ Hence, from cross-sectional studies it cannot be concluded whether the association suggests a mechanism for suicidality or whether it is a result of experienced stress, thus reflecting reverse causation. Therefore, longitudinal studies that could confirm or refute this hypothesis are needed.

Estradiol and progesterone

The perimenopause marks a woman's final reproductive years. This period is characterized by a diminishing supply of ovarian follicles which causes estradiol to highly fluctuate daily and ovulations to occur unpredictively.^{20, 108, 109} At the same time, perimenopausal women are particularly vulnerable for depression and irritability.¹¹⁰⁻¹¹⁵ It is therefore suggested that fluctuations in estradiol and progesterone may trigger women to develop mood disturbances.

Some previous studies on the link between variability in estradiol levels and depressive symptoms during the perimenopause found an association,¹¹⁶⁻¹¹⁸ but others did not.¹¹⁹⁻¹²¹ However, in the majority of studies sex steroid levels were measured only a couple of times over multiple months. This does not reflect the high variability in estradiol that perimenopausal women are experiencing.^{116, 117, 119-121} Moreover, irritability has never been examined in relation to variability in estradiol

or progesterone. Hence, studies that measure sex steroids more frequently concurrently with depressive symptoms and irritability are crucial to provide insight in the role of variability in mood disturbances during the perimenopause.

Associations between hormonal contraceptives and depression

While hormonal contraceptives are generally considered safe,^{122, 123} several large-scale observational studies suggest that hormonal contraceptive use may increase a woman's risk for depression, especially during adolescence.^{124, 125} Alternatively, two systematic reviews have suggested that hormonal contraceptives might induce depression in women using higher doses of estrogen or androgenic progestins, but actual evidence for that remains lacking.^{126, 127} Due to these conflicting findings, the putative relationship remains disputed.

Several factors are likely to contribute to the discrepancy between previous findings. First, there might be subgroups of women especially at risk for side effects of hormonal contraceptives. For example, young women may be on such subgroup of women as puberty is a critical period for both the activational and organizational sex steroidal influences on the brain.^{35, 128} In contrast, women with premenstrual syndrome (PMS) may experience beneficial effects of OC use. Fluctuations in estradiol and progesterone likely trigger symptoms in women with PMS, and thus interventions such as OCs that abolish ovarian cyclicity are thought to effectively reduce symptoms.¹²⁹ Second, there are multitude of theoretically justifiable operationalizations and analytic choices to address the question of whether hormonal contraceptive use is linked to a higher depression risk. Consequently, this may influence the associations reported in the literature. This is even more likely for sensitive topics like this, as someone's personal experiences and beliefs may, intentionally or not, drive a biased interpretation of the results. Finally, the current state of knowledge is primarily based on findings from observational research. However, examination of experimental data is crucial for determining whether there is a causal relationship between hormonal contraceptive use and depression. Therefore, more research that transparently reports on the risk of hormonal contraceptives for specific subgroups, preferable based on experimental data, is needed.

The role of researchers in science journalism

Science is an integral part of society. Ideally, the public has access to accurate information on the latest insights in research. Science journalism is important in this regard as it is the main link between researchers and the public. This link has shown its importance in today's world where fake news is every day's business. For years, researchers have been hesitant to be involved in public communication, but researchers began to rediscover the value of public visibility in the late 20th century. Nowadays, motivations for researchers to be involved in public communication include the hopes that public visibility increases the positive public attitude toward research and that it increases the visibility for funding bodies.^{130, 131}

Despite the increasing interest of researchers to be involved in public communication, this has not resulted in more accurate science coverage. Unfortunately, medical news messages often contain exaggerated claims that are not supported by the research itself. For example, the results of an observational study that were published in 2016, made headlines worldwide.¹²⁴ "Are YOU on the Pill? You're more likely to be depressed: Women who use contraception are up to 70 per cent more likely to be on antidepressants," said one news article.¹³² Not was this only incorrect (there was a 29% increased risk), the news article also failed to mention that this was an increase in relative risk, not in absolute risk. In fact, for every 100 women not using hormonal contraceptives, 1.7 each year were given anti-depressants, whereas for every 100 women using hormonal contraceptives, this risk was slightly increased to 2.2. This means that for every 200 women using hormonal contraceptives, 1 additional woman uses an antidepressant compared to women not using hormonal contraceptives. Such incorrect and exaggerated statements are problematic as people may delay or avoid hormonal contraceptive use based on such claims, with unwanted pregnancies as a consequence.

Importantly, journalists are not the only ones to blame for the wrong transmission of medical news. Journalists frequently base their news articles on press releases that are approved by researchers.¹³³ However, exaggerations and caveats found in medical news message are often already present in the press release.¹³⁴ This illustrates that researchers should be more aware about their role in the dissemination of research findings.

Aims of this thesis

The overall purpose of the work described in this thesis was to define more precisely how sex steroids relate to depression by using well-phenotyped studies and reliable sex steroid measurements. Three questions will guide this thesis. Firstly, are endogenous sex steroids associated with depression, and if so, to what extent? This question will be studied in **Chapters 2 through 8**, which together form Part I of this thesis. Secondly, do hormonal contraceptives cause depression? This question will be addressed in Part II of this thesis and consists of work of **Chapters 9 through 12**. Thirdly, what can researchers do to improve science journalism? This question will be discussed in Part III of this thesis by the use of a case study described in **Chapter 13**. Finally, the findings from this thesis are summarized and discussed in **Chapter 14**. In the following paragraphs, the chapters are further introduced and the cohort studies that are used to reach the goals are described in *Table 1.1*.

Part I: Endogenous sex steroids and depression

Chapter 2 is the starting point of this thesis with a methodological paper on different sampling procedures for testosterone. The standard way is to measure testosterone levels in blood. However, if levels could also be obtained reliably from saliva, this would provide a less stressful alternative. To increase knowledge on the validity of these assessments, the association between testosterone levels derived from saliva and blood sampling procedures is examined using data from NESDA.

In **Chapter 3 and 4** the association between androgen levels, such as testosterone, and the presence and course of MDD in adult women and men is examined. In **Chapter 5**, the association between testosterone and metabolic syndrome is determined in older men and women, and the impact of MDD on this association is studied. In the final chapter on androgens, **Chapter 6**, the association between androgen levels and suicidal thoughts and suicidal attempts is examined. For Chapters 3, 4, and 6, data from NESDA are used, while for Chapter 5, data are derived from NESDO.

In the next two studies that follow, the association between variability of estradiol and progesterone levels and mood in perimenopausal women is examined. In **Chapter 7**, the association between variability in estradiol and progesterone, and depressive symptoms is studied, while in **Chapter 8** the relationship of the same sex steroids with irritability symptoms is examined. For both chapters, data from PERI is used.

Part II: Hormonal contraceptive use and depression

In the second part of this thesis, the role of hormonal contraceptives in depression is examined. **Chapter 9** describes the association between OC use, the most often used hormonal contraceptive, and depressive symptoms in young women. **Chapter 10** further expands upon the results of Chapter 9 by examining whether there is also an association between adolescent OC use and risk of MDD in early adulthood. As previous findings on this association have been conflicting, all justifiable models were analyzed to minimize the researcher's degrees of freedom. For both these chapters, data from TRAILS is used.

To examine whether the associations found in Chapter 9 and 10 are causal, in the next two Chapters data is used from randomized clinical trials on the effects of hormonal contraceptives on depressive symptoms. In **Chapter 11**, a network meta-analysis is used to define the effect of hormonal contraceptive use on depressive symptoms compared to a placebo in 'healthy' women. With the same analysis technique, the effect of OCs on depressive symptoms in women with PMS is determined in **Chapter 12**.

Part III: Science journalism

The last piece of this thesis, **Chapter 13**, is the odd one out. Instead of another study on the association between sex steroids and depression, it is a case study of my personal experiences as a researcher with the substantial media attention following the publication of findings of Chapter 9. Tips are shared that will provide other researchers with tools that increase the likelihood that the media correctly disseminate research findings.

Table 1.1. Characteristics of the data used in this thesis.

Study	Age	Sample size	Sex	Chapter with aim
NESDA ¹³⁵	18-65	n = 2048	♀♂	Chapter 2: To examine the association between salivary and plasma testosterone levels.
		n = 1659	♀♂	Chapter 3: To examine 1) whether women with current MDD have higher plasma androgen levels compared to women who never had MDD, and 2) whether androgen levels are associated with onset and remission of MDD.
		n = 823	♀♂	Chapter 4: To examine 1) whether men with lower androgen levels are more likely to have MDD or to develop MDD, or that 2) androgen levels only associate with some overlapping symptoms between MDD and hypogonadism.
		n = 2861	♀♂	Chapter 5: To examine 1) the associations of plasma androgens with concurrent and future suicidality, and if so 2) to determine whether these associations were mediated by a behavioral trait like reactive aggression.
NESDO ¹³⁶	60-93	n = 478	♀♂	Chapter 6: To determine 1) whether plasma testosterone is associated with metabolic syndrome, and 2) whether this association is different for depressed and non-depressed older men and women.
PERI	35-56	n = 50	♀♂	Chapter 7: To examine what (non-)hormonal factors predispose perimenopausal women to experience depressive symptomatology.
		n = 50	♀♂	Chapter 8: To examine what (non-)hormonal factors predispose perimenopausal women to experience irritability symptoms.
TRAILS ¹³⁷	13-25	n = 1010	♀♂	Chapter 9: To determine 1) the association between OC use and depressive symptoms in young women, and 2) whether this association is stronger in adolescents.
		n = 725	♀♂	Chapter 10: To examine the association between adolescent OC use and possible future MDD in young women analyzing all theoretically justifiable models.
Randomized clinical trials	16-32	n = 5833	♀♂	Chapter 11: To determine the effects of hormonal contraceptive use on depressive symptoms in women.
Randomized clinical trials	24-36	n = 1205	♀♂	Chapter 12: To determine the efficacy of OC use on depressive symptoms in women suffering from PMS.
Personal experiences	-	-	-	Chapter 13: To share my personal experience with media attention and to provide tips for researchers to increase the likelihood that the media will not exaggerate the research findings.

Abbreviations: MDD, major depressive disorder; NESDA, Netherlands Study on Depression and Anxiety; NESDO, Netherlands Study on Depression in Older persons; OC, oral contraceptive pill; PMS, premenstrual syndrome; TRAILS, TRacking Adolescents' Individual Lives Survey.

Part I

Endogenous sex steroids and depression

Chapter 2

Testosterone in human studies: modest associations between plasma and salivary measurements

Anouk E de Wit
Fokko J Bosker
Erik J Giltay
Carien S de Kloet
Karin Roelofs
Johannes van Pelt
Brenda WJH Penninx
Robert A Schoevers

Adapted version from the article published in *Andrologia*, March 2017.

Abstract

Aim: Testosterone is involved in many processes like aggression and mood disorders. As it may easily diffuse from blood into saliva, salivary testosterone is thought to reflect the plasma free testosterone level. If so, it would provide a welcome non-invasive and less stressful alternative to blood sampling. Past research did not reveal consensus regarding the strength of the association, but sample sizes were small. This study aimed to analyze the association in a large cohort.

Methods: In total, 2048 participants (age range 18–65 years; 696 men and 1352 women) were included and saliva (using cotton Salivettes®) and plasma were collected for testosterone measurements. Levels were determined by enzyme-linked immunosorbent assay and radioimmunoassay respectively. Free testosterone was calculated by the Vermeulen algorithm. Associations were determined using linear regression analyses.

Results: Plasma total and free testosterone showed a significant but modest association with salivary testosterone in men (adjusted standardized beta-coefficient [β] = 0.09, p = 0.01; and β = 0.15, p < 0.001, respectively) and in women (adjusted β = 0.08, p = 0.004; and β = 0.09, p = 0.002, respectively).

Conclusion: The modest associations indicate that there are many influencing factors on these associations of both technical and biological origin.

Introduction

Testosterone is secreted by the Leydig cells, under regulation of luteinizing hormone (as part of the hypothalamic-pituitary-gonadal-axis [HPG-axis]). It is known for its key role in the sexual differentiation and sexual functioning, but plays an important role in social behavior, such as in the maintenance of social status and aggression as well.^{138, 139} Moreover, as it affects the HPG-axis, it seems to be involved in various psychiatric disorders including mood disorders. This is further supported by the sex differences found in the prevalence of the latter.¹³⁹

Testosterone is commonly measured in blood, but first reports on testosterone measurement in saliva appeared in the late sixties of the last century. The reason for collecting saliva is the less invasive, stressful and painful procedure compared with vena puncture, which is an important benefit especially in epidemiological studies. There are several arguments for the assumption that salivary testosterone is a good reflection of the plasma testosterone. First, testosterone is a lipophilic molecule that can easily pass the barrier between blood and salivary glands. Second, the salivary glands do not produce this hormone, implicating that all testosterone in saliva originates from the bloodstream largely through passive diffusion.

Testosterone in blood is bound to different proteins. In adult men, only 2-3% of the circulating testosterone constitutes the free fraction. The remaining testosterone is strongly bound to sex hormone binding globulin (SHBG) (44%), and more loosely to albumin (50%) and cortisol binding globulin (3-5%). Bioavailable testosterone is the free circulating hormone plus the (easily dissociated) albumin bound fraction.¹⁴⁰ Several algorithms to calculate the free testosterone fraction have been proposed in the past, including the one used by Vermeulen et al.¹⁴¹ However their validity may largely depend on the accuracy of the measurements and the abundance of other hormones such as estradiol in the matrix.¹⁴⁰ Other complicating factors are age, sex, body mass index (BMI), somatic disease processes, and medication that modify testosterone levels.

Since the seventies of the last century, several studies have investigated the relation between blood and salivary testosterone, but results have varied. In men, marked associations between salivary testosterone and both free and total serum testosterone have been reported, with correlation coefficients ranging from $r = 0.74$ to almost 1.0.¹⁴²⁻¹⁴⁶ Recently more modest strengths of associations have also been reported.^{147, 148} In women, much poorer associations were observed, with

correlation coefficients ranging from $r = 0.23$ to $r = 0.79$.^{145, 146, 149-153} This might relate to the generally much lower testosterone values and the accuracy of blood measurements within this range, but a more complex relation in women cannot be ruled out either. Limiting factors in most previously performed studies are the small to moderate sample sizes (ranging from $n = 12$ to $n = 218$) to analyze such relationships with precision and not taking confounders into account.^{144, 154}

Subjects from the present study stem from a large prospective cohort study with more than 2000 participants. The present study therefore has enough power to address the following questions: 1) Is salivary testosterone a valid marker for plasma testosterone in large scale cohort studies, and 2) do sociodemographic, sampling and lifestyle factors confound the values of salivary or plasma testosterone and therefore confound the association?

Methods

Data was collected from the Netherlands Study on Depression and Anxiety (NESDA), which is an multicenter study which investigates the aetiology and development of depressive and anxiety disorders. Participants who were not fluent in Dutch or suffered from one of the following diseases, were excluded: psychotic disorder, obsessive compulsive disorder, bipolar disorder, or severe addiction disorder.¹³⁵ The study design was approved by the Ethical Review Board of the VU University Medical Centre and thereafter by the local ethical committees of the participating centers. All participants gave verbal and written informed consent.

In the present analysis a subset of the NESDA sample was used. We excluded 923 participants who did not collect saliva samples ($n = 826$ [27.7%]) and/or had missing values on plasma hormone data ($n = 96$ [3.3%]) and those with suspected marked abnormalities in plasma protein composition, as in chronic kidney failure or liver diseases/cirrhosis, or during pregnancy in order to calculate reliable free testosterone values.¹⁴¹ In total, 2048 participants (age range 18–65 years; 696 men and 1352 women) were included.

Assessment of testosterone in plasma and saliva

Vena puncture took place in the morning after an overnight fast. The mean time of waking up was 06:57 AM (2.5 and 97.5th percentile 05:30-08:00) and the mean time of sampling was 08:36 AM (2.5 and 97.5th percentile 08:00-09:35).

The blood samples were processed and stored at -80°C . Total testosterone was measured in duplicate after performing a diethylether extraction with an in-house radioimmunoassay using a polyclonal antitestosterone-antibody. Inter-assay coefficient of variation was estimated $<10\%$ for SHBG and between 14% and 19% for total testosterone and the analytical sensitivity was 0.5 nmol/L . Free testosterone was calculated from total testosterone and SHBG using the algorithm by Vermeulen.¹⁴¹ Because albumin concentration was not measured in the NESDA sample a constant albumin concentration of 43 g/L (0.62 mmol/L) was assumed. According to Vermeulen, variations in albumin concentration smaller than 25% will not significantly influence estimated free testosterone values. To verify whether the albumin concentration in our sample was within 25% , albumin concentrations of 30 randomly selected participants were measured in plasma. Variation in this sample was 19% (mean 45.2 , standard deviation [SD] 8.6), well within the recommended borders.

Salivary collections took place in the morning at home. Participants were instructed verbally as well as through detailed written information accompanying the Salivettes®. Saliva was collected immediately after awakening ('the moment you can open your eyes') and 30, 45 and 60 minutes later. The mean time of sampling for the first timepoint was 07.26 AM (2.5 and 97.5^{th} percentile 06.20 - 10.10). The mean time of sampling for the fourth timepoint was 08.27 AM (2.5 and 97.5^{th} percentile 06.30 - 11.15). All subjects were instructed not to eat, smoke, drink and brush teeth within 15 minutes of sampling. The Salivettes® were stored at -20°C and returned by regular mail. After receipt, Salivettes® were thawed and centrifuged at 2000g for 10 min in order to remove the mucus and obtain clear liquids which could be pipetted. The saliva were aliquoted and stored at -80°C . To smooth the episodic secretion, $75\text{ }\mu\text{L}$ of each of the four samples were mixed to yield one sample. If one of the samples was missing, a corresponding volume of the other sample(s) was taken. Biochemical analysis of testosterone in saliva was measured in duplo by the testosterone in saliva assay from Diagnostic Biochem Canada (EiAsy Testosterone Saliva, DBC: CAN-TE-300) using $2\times 100\text{ }\mu\text{L}$ material. The sensitivity of the kit is 1.0 pg/ml , while displaying low cross-reactivities with other steroids. In every assay a standard control was used, with a mean of 26.9 pg/ml (standard deviation [SD] 2.1) that was reproducible with a coefficient of variation of 7.8% . The intra assay precision was 7.1 , 3.4 , and 6.7% at the concentrations of resp. 14 , 38 and 123 pg/ml ($n = 10$). The mean coefficient of variation difference

over all testosterone measurements was 10.2%. In most cases during the data collection, there was a median time interval of eight days (inter quartile range 4 to 22) between the blood and the saliva collection.

Covariates

Sociodemographic, sampling and lifestyle variables were analyzed in order to determine whether they significantly impacted testosterone levels. Research performed in the same cohort revealed that age was strongly and inversely associated with morning salivary testosterone levels in both men and women.¹⁵⁵ To analyze which variables significantly covaried plasma testosterone levels, we included the following baseline factors; age, North European ancestry, level of education, BMI, time of awakening, time of blood sampling, fasting, smoking status, alcohol use, chronic diseases, and the physical activity. Level of education was specified in years of education. Time of awaking was demarcated as the moment you can open your eyes. Fasting was defined as an overnight fast. BMI was calculated as the weight (kg) divided by the squared height (m). Smoking status was dichotomized into non- or current smoker. Alcohol use was defined as daily alcohol intake (units). Chronic disease was prevalent when suffering from diabetes mellitus, stroke or other cardiovascular diseases or cancer. Physical activity was assessed using the International Physical Activity Questionnaire¹⁵⁶ and expressed per 1000 Metabolic Equivalent of Task /min a week. A Metabolic Equivalent of Task /min is defined as the metabolic equivalent of the number of calories consumed by a person (of 60 kg) per minute in an activity relative to basal metabolic rate. In addition, menstrual status, defined as pre- or post-menopausal with or without the use of oral contraceptive pills (OCs) was taken into account only in women.

Statistical analysis

Baseline sociodemographic and lifestyle characteristics and hormone levels were described using number of cases (percentages), mean $\pm$ standard deviation (SD) or median + inter quartile range (IQR), when appropriate in men and women. Values of salivary testosterone were log transformed in order to provide normal distribution.

Univariate analyses were carried out using univariate linear regression in subjects free of psychopathology to evaluate the effects of twelve

sociodemographic, sampling and lifestyle factors on plasma testosterone levels for men and women separately. We performed multiple linear regression analyses to assess independent determinants of plasma testosterone. A variable was defined as a confounder when the p-value was < 0.10 both in the uni- and multivariate analyses for both plasma and salivary testosterone. Bonferroni correction ($\alpha = 0.10/12 = 0.008$) was performed to adjust the results of multiple comparisons.

To determine the association between salivary and plasma testosterone, we calculated standardized beta-coefficients (β s), using linear regression, separately for strata of men and women. Multivariate analyses were performed adjusting for the determined confounders. To estimate the impact of different confounders, we calculated the percent change in β s (i.e., $[\beta^b - \beta^a]/\beta^a * 100\%$ of more than 10%, i.e., $\% \Delta > 10$).

Finally, we determined whether the association between salivary and plasma testosterone differed between participants with a depression or anxiety disorder and healthy controls and between participants with a different time sampling frame. Regression analyses were repeated in the subgroup of participants free of psychopathology and in the subgroup of participants with a time interval of maximum two weeks afterwards for one day for plasma and saliva sampling. In a last sensitivity analyses, we excluded seven outliers in men (plasma testosterone levels < 5.0 nmol/L) and 8 outliers in women (plasma testosterone levels > 7.0 nmol/L). Multivariate analyses were performed adjusted for determined confounders. A two-sided p value < 0.05 was considered statistically significant. Data was analysed using IBM SPSS Statistics 22 (SPSS Inc., Chicago, Ill., USA).

Results

In *Table 2.1* the socio-demographic and lifestyle characteristics of 2048 participants are depicted according to sex. The mean age was 44.9 ± 12.6 years in men and 42.7 ± 13.1 years in women. In total, 76.3% of the participants suffered from a current anxiety and/or depressive disorder.

Table 2.1. Characteristics of 696 men and 1352 women.

	Men n = 696	Women n = 1352
Age , mean (SD)	44.9 (12.6)	42.7 (13.1)
North-European ancestry , no. (%)	669 (96.1)	1286 (95.1)
Education , mean (SD)	12.3 (3.3)	12.4 (3.3)
Time of awakening plasma measurement , mean (SD)	06:56 (00:53)	06:58 (00:40)
Time of blood sampling , mean (SD)	08:36 (01:18)	08:35 (01:19)
Fasting , no. (%)	660 (95.0)	1307 (97.0)
Time of awakening salivary measurement , mean (SD)	07:23 (01:14)	07:29 (01:15)
Time of blood sampling point 1 , mean (SD)	07:23 (01:13)	07:28 (01:14)
Time of blood sampling point 4 , mean (SD)	08:27 (01:14)	08:29 (01:10)
Comorbid anxiety/depressive disorder , no. (%)	509 (73.1)	1053 (77.9)
BMI , mean (SD)	26.2 (4.5)	25.2 (5.3)
Currently smoking , no. (%)	261 (37.5)	420 (31.1)
Daily alcohol intake , mean (SD)	1.5 (1.8)	0.8 (1.1)
No. chronic diseases , no. (%)	145 (20.8)	187 (13.8)
Physical activity , mean (SD)	3.7 (3.3)	3.7 (2.9)
Menstrual status		
Post-menopausal, no. (%)	-	509 (38.0)
Pre-menopausal with oral contraceptive pill use, no. (%)	-	338 (25.2)
Pre-menopausal without oral contraceptive pill use, no. (%)	-	492 (36.7)
Salivary testosterone (pg/ml), median (IQR)	16.14 (2.79-26.6)	16.11 (10.30-26.77)
Plasma total testosterone (nmol/L), median (IQR)	19.10 (15.5-23.6)	2.80 (2.20-3.40)
Plasma free testosterone (nmol/L), median (IQR)	0.40 (0.32-0.48)	0.03 (0.02-0.04)

Abbreviations: BMI, body mass index; IQR, interquartile range; SD, standard deviation. Time is given in h:min, alcohol intake in units, and physical activity in 1000 Metabolic Equivalent of Task min/week.

The associations between sociodemographic, sampling and lifestyle variables and plasma total and free testosterone levels were analyzed in subjects free of psychopathology as they may be considered as confounders in the subsequent association analyses with salivary testosterone. Previous research performed in the same cohort revealed that age was a potential confounder, as it was inversely correlated with salivary testosterone.¹⁵⁵ Also, current smoking was associated with higher, and OC use, with lower salivary testosterone levels. With regard to plasma levels, age was inversely associated with plasma total testosterone in men and women, which was statistically significant (after Bonferroni correction) in both univariate analyses ($p < 0.001$ for men and $p = 0.001$ for women), but not in multivariate analyses ($p = 0.01$ for men and $p = 0.07$ for women). See *Table 2.2*.

Table 2.2. Associations of sociodemographic, sampling and lifestyle variables with plasma total and free testosterone levels in 187 men and 299 women free of lifetime psychopathology.

	Plasma total testosterone				Plasma free testosterone			
	Univariate		Multivariate		Univariate		Multivariate	
	β	p	β	p	β	p	β	p
<i>Men</i>								
Age	-0.31	<.001	-0.21	.01	-0.56	<.001	-0.52	<.001
North-European ancestry	-0.17	.02	-0.19	.01	-0.01	.89	-0.05	.44
Education	0.08	.28	0.06	.44	0.00	.95	0.00	1.00
Time awakening	0.08	.26	0.04	.55	0.08	.29	0.03	.62
Time blood sampling	0.12	.11	0.11	.10	0.12	.11	0.12	.05
Fasting	-0.01	.94	-0.00	.99	-0.09	.23	-0.07	.27
BMI	-0.37	<.001	-0.26	.001	-0.28	<.001	-0.04	.62
Currently smoking	0.10	.16	0.10	.15	0.04	.57	0.05	.46
Alcohol intake	0.01	.94	0.06	.38	-0.04	.63	-0.02	.80
Chronic diseases	-0.19	.01	-0.03	.71	-0.28	<.001	-0.06	.42
Physical activity	0.03	.64	0.09	.19	-0.04	.61	-0.01	.93
<i>Women</i>								
Age	-0.20	.001	-0.15	.07	-0.02	.79	-0.02	.81
North-European ancestry	0.07	.20	0.06	.30	0.02	.71	0.02	.74
Education	0.06	.31	0.05	.42	0.03	.56	0.04	.52
Time awakening	0.03	.59	-0.03	.65	-0.04	.45	-0.06	.34
Time blood sampling	-0.04	.46	-0.06	.35	-0.01	.93	-0.04	.50
Fasting	0.00	.98	-0.00	.95	0.07	.22	0.04	.54
BMI	-0.13	.03	-0.08	.22	0.06	.29	0.07	.29
Currently smoking	0.06	.30	0.04	.47	0.03	.58	0.04	.55
Alcohol intake	-0.03	.67	-0.01	.87	0.01	.84	0.04	.58
Chronic diseases	-0.04	.47	-0.00	.97	0.02	.72	0.02	.75
Physical activity	0.04	.45	0.05	.41	0.02	.68	0.03	.69
Menstrual status	-0.13	.02	-0.01	.94	-0.06	.30	-0.08	.31

Abbreviations: β , standardized beta coefficient; BMI, body mass index. Data are standardized β s using univariate and multivariate linear regression. Univariate analyses were linear regression with a single potential confounders, whereas multivariate analyses included all potential confounders.

Table 2.3 shows the associations of modest strength between plasma and salivary testosterone levels in 2048 participants according to sex. Age was included in the adjusted regression, as only this variable was associated with both plasma and salivary testosterone in men (total and free testosterone) and women (total testosterone) and therefore might confound the association between plasma and

salivary testosterone. Plasma total testosterone showed a significant association with salivary testosterone in men (adjusted $\beta = 0.09$, $p = 0.01$) and in women (adjusted $\beta = 0.08$, $p = 0.004$). Plasma free testosterone showed a significant association with salivary testosterone both in men (adjusted $\beta = 0.15$, $p < 0.001$) and in women (adjusted $\beta = 0.09$, $p = 0.002$). *Figure 2.1* depicts the association between plasma free testosterone and salivary testosterone (with testosterone levels categorized for illustrative purposes only). The size of the confounding effect of age on the association between salivary and plasma total and free testosterone was $\%\Delta\beta = -25.0\%$ and $\%\Delta\beta = -20.0\%$ in men, and $\%\Delta\beta = -27.3\%$ for free testosterone in women respectively.

Table 2.3. Associations between plasma and salivary testosterone levels in men and women.

	Salivary testosterone in men			Salivary testosterone in women		
	β	p	$\%\Delta$	β	p	$\%\Delta$
<i>All subjects</i>	n = 696			n = 1352		
Total testosterone	0.12	.002		0.11	<.001	
Total testosterone^a	0.09	.01	-25.0%	0.08	.004	-27.3%
Free testosterone	0.18	<.001		0.09	.002	
Free testosterone^a	0.15	<.001	-20.0%	-	-	-
<i>Sampling within 2 weeks</i>	n = 396			n = 623		
Total testosterone	0.10	.048		0.07	.05	
Total testosterone^a	0.07	.14	-30.0%	0.06	.12	-14.3%
Free testosterone	0.16	.002		0.06	.08	
Free testosterone^a	0.12	.04	-25.0%	-	-	-
<i>Sampling on same day</i>	n = 23			n = 41		
Total testosterone	-0.05	.81		0.08	.62	
Total testosterone^a	-0.07	.77	-40.0%	0.14	.42	+75.0%
Free testosterone	-0.06	.86		0.04	.80	
Free testosterone^a	-0.04	.86	+33.3%	-	-	-
<i>Free of psychopathology</i>	n = 187			n = 299		
Total testosterone	0.14	.05		0.09	.19	
Total testosterone^a	0.12	.11	-14.2%	0.06	.28	-33.3%
Free testosterone	0.17	.02		0.13	.02	
Free testosterone^a	0.15	.09	-11.7%	-	-	-

Abbreviation: β , standardized beta-coefficient. Data are standardized β by linear regression analysis. ^a Adjusted for age.

Figure 2.1. Relationship between plasma free testosterone and salivary testosterone levels in 699 men and 1352 women.

Plasma free testosterone levels were categorized for visualization purposes only, as all analyses were done with continuous data. Salivary testosterone levels were $\log_e$ transformed before analysis, and back-transformed geometric mean levels are presented on logarithmic scales. Error bars represent standard errors. The sizes of the boxes are proportional to the number of participants.

In a sensitivity analysis performed in a subgroup of participants who had a ≤ 2 weeks interval between both samplings and on the same day, the strengths of the associations were slightly attenuated. See *Table 2.3*. A further sensitivity analysis performed in the subgroup of subjects free of psychopathology (men $n = 187$ and women $n = 299$) showed that the strengths of the associations between salivary testosterone and plasma total testosterone remained the same. The adjusted β was 0.12 ($p = 0.11$) in men and 0.06 ($p = 0.28$) in women for the association between salivary testosterone and plasma total testosterone, and 0.15 ($p = 0.09$) in men for the association between salivary testosterone and plasma free testosterone. The exclusion of seven outliers in men (total testosterone adjusted $\beta = 0.10$, $p = 0.001$ and free testosterone adjusted $\beta = 0.16$, $p < 0.001$, $n = 689$) or after exclusion of eight outliers in women (total testosterone adjusted $\beta = 0.12$, $p < 0.001$, $n = 1344$) did neither remarkably change the results.

Discussion

The aim of the present study was to investigate the relation between salivary testosterone and plasma total and free testosterone in a large sample. A second objective was to investigate whether sociodemographic, sampling and lifestyle

factors covaried the values of salivary or plasma testosterone and therefore could be stated as confounders. Although the associations were modest, they were consistent across analyses in subjects with or without psychopathology and participants who had a less than two weeks interval between both samplings. Only age significantly confounded the association between plasma and salivary testosterone, however not between plasma free testosterone and salivary testosterone in women.

To find an explanation for these modest associations relative to the generally much stronger associations that have been reported previously,¹⁴²⁻¹⁴⁵ both methodological and technical issues should be discussed.

Strengths of this study are that several precautions were taken in the NESDA study including not brushing teeth, smoking or eating within 15 minutes of saliva collection to prevent contamination of the saliva samples by blood. Salivary samples are extremely sensitive for contamination by blood,¹⁵⁷ as testosterone levels in blood are higher than in saliva. Importantly, not all previous studies^{144-146, 154} have taken such precautions and this might have influenced the correlations. In addition, we included a large sample ($n = 2048$) of participants, which was much larger than prior studies. Consequently, this study was sufficiently powered to examine confounding effects of sociodemographic, sampling and lifestyle factors.

However, there are shortcomings of our study which might have influenced the correlation. First, the use of cotton Salivettes® instead of a passive drool approach has been linked with measurement errors and inflated values and therefore deemed to be less reliable.^{149, 158} However, control studies have shown that cotton Salivettes® increase testosterone values in a stable and predictable manner. As this absolute increase in measured salivary testosterone is similar, this may explain the larger relative increases that we saw in women compared to men, as well as a probably larger measurement error in women than in men. Nevertheless, their rank order stability was partly preserved, which makes Salivettes® still applicable for testosterone measurements within large cohort studies.^{155, 159} Second, differences in the sampling times of plasma and saliva could have been an important factor. The time span between these collections often exceeded two weeks, leading to assessment within different phases of the menstrual cycle. However, subgroup analyses in sampling within two weeks or even on the same day did not reveal better correlations, pointing towards other factors that played a more important role. Besides the average time interval of half an hour with the morning saliva and plasma collections could have diluted associations in

the present study^{160, 161} as testosterone has a diurnal rhythm with concentration drops to 50% from morning to evening. However, most other studies did not collect both media at the exact same time either, with timespans of one to two hours.^{142, 143, 147, 148, 151, 153} Importantly the diffusion rate from testosterone between the blood and saliva barrier is unknown and therefore, the optimal time sampling frame between blood and saliva collection neither. Third, the majority of participants had psychopathology which is thought to influence hormone levels.^{88, 155, 162} However, in a subsample of participants free of mental health problems, correlations stayed the same. Fourth, immunoassays for testosterone are considered to be less reliable than liquid chromatography–tandem mass spectrometry (LC-MS/MS) and gas chromatography–mass spectrometry analyses in particular at the lower concentrations generally measured in women (<5 nmol/L).^{163, 164} Problems noted with this method are inadequate specificity, interferences with hetero and auto-antibodies and most importantly cross-reactivity with other steroid hormones. Measurement of salivary testosterone with LC-MS/MS could hopefully broaden its implication for diagnostics in endocrine abnormalities as well. Determination of the andrological status also requires measurement of other steroid hormones, e.g., estradiol and 5 α -dihydrotestosterone (5 α -DHT), which assessment may be insufficiently reliable when using immunoassays, in part due to their lower values.¹⁶⁵ Consequently, immunoassays are more and more replaced by LC-MS/MS, which is a more specific and sensitive analysis technique¹⁶⁶ but more expensive and time consuming as well.⁸⁹ Hsing et al. compared the measurement of testosterone in serum between radioimmunoassay and gas chromatography–mass spectrometry analyses. Although the intraclass correlation coefficients were generally higher within the batch measured by gas chromatography–mass spectrometry analyses, the correlation coefficients were similar. Moreover, the Pearson's and Spearman's correlation coefficients between both methods were high (>0.9) and all statistically significant.¹⁶⁵ These high correlation coefficients were confirmed in a later studies.^{167, 168} To the best of our knowledge, less is known about the direct comparability between salivary testosterone levels measured by Enzyme-Linked Immuno Sorbent Assay versus mass spectrometry. Probably, same methodological issues when using radioimmunoassay on saliva samples were of importance, but were not elaborated further in the article, making it difficult to interpret their results.¹⁴⁹ Besides, most previous studies have used immunoassays for the detection of testosterone.^{142-148, 150-152, 154, 157} Fifth, different immunoassays for the saliva and plasma testosterone measurements in the NESDA study have

been used. Arguably, the relation between saliva and blood measurements will suffer when cross-reactivity profiles of the immunoassays are markedly different.

We would advise that future research should use passive drool samples or synthetic rather than cotton Salivettes®. Arguably, measurement of testosterone in hair using LC-MS/MS¹⁶⁹ might be an alternative and unique chance to estimate testosterone production over longer time periods. Future research should reveal the association between testosterone measured in hair, saliva and blood. Clearly, such studies must be performed with LC-MS/MS to reduce analytical errors in the lower concentration range. Moreover, saliva and blood collection should ideally be performed at the same day, ideally within hours. Conversely this would also challenge the world-wide clinical practice of measuring testosterone level at a single time point and advocate the measurement of testosterone in hair. Finally, age should be taken into account as a confounder for the association between salivary and plasma testosterone. Unknown yet but perhaps important, the matrix might be more complex for plasma than saliva, indicating that specificity issues might play a more prominent role with the plasma measurements.

In summary, although statistically significant associations were found between plasma total and free testosterone, and salivary testosterone in the present study, the associations were much more modest as compared to previous studies.¹⁴²⁻¹⁴⁵ Only age significantly confounded the association between plasma and salivary testosterone, however not between plasma free testosterone and salivary testosterone in women. These results indicate that the association is not a simple one and likely influenced by many factors of both technical (e.g., the use of Salivettes® instead of passive drools of saliva and other issues of sample collection protocols) and biological origin. Until we know the impact of these factors, the determination of testosterone in blood continues to be the method of choice, especially in clinical settings.

Acknowledgements

The infrastructure for the NESDA study (www.nesda.nl) has been funded through the Geestkracht program of the Netherlands Organisation for Health Research and Development (Zon-Mw, grant number 10-000-1002) and participating universities (VU University Medical Center, Leiden University Medical Center, University Medical Center Groningen). The authors wish to acknowledge Anneke Muller Kobold for her contribution to the plasma testosterone determinations.

Chapter 3

Plasma androgens and the presence and course of depression in a large cohort of women

Anouk E de Wit
Erik J Giltay
Marrit K de Boer
Fokko J Bosker
Aviva Y Cohn
Willem A Nolen
Ursula B Kaiser
Hadine Joffe
Brenda WJH Penninx
Robert A Schoevers

Adapted version from the article published in Translational Psychiatry, February 2021.

Abstract

Aim: Major depressive disorder (MDD) has a higher prevalence in women with supraphysiologic androgen levels. Whether there is also an association between MDD and androgen levels in the physiological range, is unknown. This study examined if women with current MDD have higher androgen levels compared to women who have never had MDD, and if androgen levels are associated with onset and remission of MDD.

Methods: In 1659 women (513 current MDD, 754 remitted MDD and 392 never MDD), baseline plasma levels of total testosterone, 5 α -dihydrotestosterone, and androstenedione were determined with liquid chromatography–tandem mass spectrometry, and dehydroepiandrosterone-sulphate and sex hormone binding globulin (SHBG) with radioimmunoassays. Free testosterone was calculated. MDD status was assessed at baseline, and at 2- and 4-years follow-up.

Results: Women were aged between 18 and 65 years (mean age 41) with total testosterone levels in the physiological range (geometric mean 0.72 nmol/L with 95% confidence interval 0.27-1.93). After adjusting for covariates and multiple testing, women with current MDD had a higher mean free testosterone than women who never had MDD (adjusted geometric mean 8.50 vs. 7.55 pmol/L, $p = 0.0005$), but this difference was not large enough to be considered clinically meaningful as it was consistent with statistical equivalence. Levels of other androgens and SHBG did not differ and were also statistically equivalent between the groups. None of the androgens or SHBG levels predicted onset or remission of MDD.

Conclusion: Our findings support the idea that plasma androgens within the physiological range have no or only limited effects on depressive disorders in women.

Introduction

Ever since the recognition of the greater preponderance of major depressive disorder (MDD) in women compared to men,¹³ sex steroids, like estrogens but also androgens, have been proposed to play a role in mood. Amongst women, the prevalence of MDD is further increased in patients with medical conditions that are associated with elevated androgen levels, such as polycystic ovarian syndrome and congenital adrenal hyperplasia.^{85, 90} This suggests that circulating androgens in women may have detrimental effects on mood. Such effects may result from androgens binding to intracellular sex steroid receptors in structures of the limbic system.⁴⁷ However, systemic androgen levels may also be affected by MDD, rather than being a causal agent. Hyperactivity of the hypothalamic-pituitary-adrenal-axis (HPA-axis) in patients with MDD could stimulate the adrenal cortex to not only enhance the release of cortisol, but of adrenal androgens as well.^{170, 171}

Epidemiological studies that have examined androgens (e.g. total testosterone) in relation to MDD prevalence, which is the product of incidence and episode duration,¹⁷² have shown equivocal results. Three cross-sectional studies found that women with current MDD had higher levels of total testosterone compared to non-depressed controls,¹⁷³⁻¹⁷⁵ whereas another study reported the opposite,¹⁷⁶ and two other studies reported no differences.^{177, 178} Four large prospective studies examined whether total testosterone levels predicted the incidence of MDD up to 10 years after baseline in adult women from the general population.^{173, 179-181} The majority reported no link,¹⁷⁹⁻¹⁸¹ whereas one study showed that an annual increase in total testosterone was associated with increased risk of MDD in women during perimenopause.¹⁷³ Only one study examined whether testosterone levels predicted the duration of MDD in 297 older women; those with lower plasma levels of total and free testosterone tended to have a higher severity of depression during 2 years follow-up that approached statistical significance.¹⁸²

The inconsistencies in these findings highlight the need to re-address the question about the role of androgens in MDD in adult women. This research should overcome previous limitations that include 1) the use of screening tools for diagnosing MDD rather than structured diagnostic interviews, 2) failure to adjust for oral contraceptive pill (OC) use and hormonal replacement therapy use, and 3) the determination of androgens by immunoassays rather than liquid chromatography-tandem mass spectrometry (LC-MS/MS).⁸⁹ Moreover, analyses should include other androgens besides free and total testosterone, as many additional endogenous androgenic compounds exist.

Here we sought to examine the role of plasma androgens in MDD in women. We hypothesized that: 1) androgen levels are higher in women with current MDD compared to women who never had MDD, 2) in women without current MDD, there is an elevated risk for a new episode with higher androgen levels, and 3) in women with current MDD, higher levels of androgens are associated with a longer time to recovery. Looking at a cohort of women from the clinically well-phenotyped Netherlands Study on Depression and Anxiety (NESDA), multiple androgens were measured by high performance LC-MS/MS. These androgen levels were examined in relation to the presence and course of MDD over four years, while taking into account important confounders such as age and OC use.

Methods

Study population

Data were derived from NESDA; a large multi-center study designed to investigate the long-term course and consequences of depressive and anxiety disorders. Between 2004 and 2007, 2981 currently and former depressed and anxious, and euthymic men and women (aged 18-65 years) were recruited. Exclusion criteria included not speaking Dutch fluently, or being diagnosed with one of the following psychiatric diseases according to Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV criteria: psychotic disorder, obsessive compulsive disorder, bipolar disorder, or substance dependence.¹³⁵ The ethical committees of the participating centers approved the study design and all participants gave verbal and written informed consent.

For this study, 1979 women were included. Exclusion criteria included women who had missing plasma androgen measurements ($n = 40$), extreme androgen values (>8 standard deviation [SD] below or above the mean, $n = 5$), were pregnant ($n = 17$) or transgender ($n = 1$). Women who used androgens (World Health Organization Anatomical Therapeutic Chemical [ATC code]¹⁸³ G03B and E), gonadotrophins (ATC code G03G), antiandrogens (ATC code G03H), gonadotrophin-releasing hormone analogues (ATC code L02AE), antiandrogens (ATC code L02BB), or aromatase inhibitors (L02BG) during the blood draw at baseline ($n = 13$) were also excluded. Finally, women who were diagnosed (currently or in the past) with anxiety disorder but not with MDD ($n = 320$) were excluded as well. The final baseline sample comprised 1659 women (83.8% of all women). Of this baseline

sample, 1432 women (85.3%) had participated at least once in a NESDA follow-up assessment after two or four years. Those without follow-up assessments ($n = 227$, 12.2%) had received less education (mean $\pm$ SD years of education 11.3 ± 3.1 vs 12.4 ± 3.2 , $p < 0.001$), and more often suffered from MDD (68.5% vs. 44.4%, $p < 0.001$) and/or an anxiety disorder (47.8% vs. 25.2%; *Supplement 1*).

Hormone assays

Androgens were measured in plasma. Blood draws took place in the morning (mean time 08:48h SD 22min) after an overnight fast (success rate 96.1%). Total testosterone, 5 α -dihydrotestosterone (5 α -DHT), and androstenedione were determined by LC-MS/MS at the Clinical Chemistry department of the University Medical Centre Groningen. The lower limit of quantitation was 0.04 nmol/L for total testosterone and androstenedione, and 0.12 nmol/L for 5 α -DHT. Women who had values below the detection limit for 5 α -DHT ($n = 153$, 9.2%) were given values 0.01 nmol/L less than the lower limit of quantitation. dehydroepiandrosterone-sulphate (DHEAS) and sex hormone binding globulin (SHBG) were determined using ARCHITECT (ABBOTT, Wiesbaden, Germany), a one and two-step radioimmunoassay with chemiflex assay protocols. The lower limit of detection of DHEAS was ≤ 0.3 μ mol/L with a calibration range of 0.00-40.71 μ mol/L, and ≤ 0.1 nmol/L for SHBG with a calibration range of 0.0-250 nmol/L. The inter-assay coefficient of variation were estimated to be $\leq 10\%$ for both. Free testosterone was calculated based on SHBG and total testosterone levels, and an assumed albumin level of 43 g/L with the formula of Vermeulen.¹⁴¹ As a measure of HPA-axis activation, salivary cortisol levels were determined. These measures were used to calculate the area under the curve (AUC) with respect to the ground (AUC_g) or increase (AUC_i) of salivary cortisol. For more details on the determination of cortisol and the calculation of the AUC we refer to previous work.¹⁸⁴ In brief, saliva samples were collected at baseline using Salivettes® (Sarstedt AG and Co, Nümbrecht, Germany) at awakening, and 30, 45 and 60 minutes later. After the determination of cortisol levels using competitive electrochemiluminescence immunoassay (E170; Roche, Basel, Switzerland), the AUC_g and AUC_i was calculated. Missing values on this covariate ($n = 655$, 39.5%) were not imputed, as substitution of this large amount of data would cause bias.

Presence, remission, and onset of major depressive disorder

At baseline, women were categorized into three groups according to their psychopathology status; never MDD (no current or past psychopathology, $n = 392$), remitted MDD (history of MDD, but currently not depressed, $n = 754$), current MDD (currently depressed, $n = 513$). The presence of MDD, currently or in the past, was ascertained using the lifetime version of the Composite International Diagnostic Interview (CIDI), version 2.1.¹⁸⁵ A current MDD was ascertained when women fulfilled the DSM-IV-based CIDI criteria within the past month, and remitted MDD was determined when participants had MDD earlier in life, but not within in the last month. The CIDI is a validated instrument with high interrater reliability (any depressive disorder kappa = 0.95) and high validity for depressive and anxiety disorders.^{185, 186} The severity of depressive symptoms were assessed with the 30-item self-report Inventory of Depressive Symptoms.¹⁸⁷ Time to remission was determined as the number of years until a participant no longer fulfilled the criteria for MDD in the subset of depressed participants at baseline on the CIDI. First or recurrent onset of MDD on the other hand, was ascertained when women without MDD at baseline fulfilled the criteria of MDD during follow-up. For both outcomes, participants who dropped out before the first follow-up assessment ($n = 108$ for time to remission, and $n = 119$ for time to onset) were excluded, and dropouts after that point were censored. This resulted in samples of 405 and 1027 participants who were used for calculating time to remission, and time to onset of MDD, respectively.

Statistics

At baseline, group characteristics of never MDD, current MDD, and remitted MDD were compared using one-way analysis of variance for independent samples or χ^2 tests, depending on whether the variable was continuous or categorical. Due to right-skewed distributions, all androgen levels were $\log_e$ -transformed, and to ease comparability of effect sizes, standardized before the analyses (into z-values).

For the first hypothesis, we compared group means of androgen levels of women with never MDD to those with remitted and current MDD, using two-way analyses of covariance. Analyses were conducted with and without adjustment for the following time-independent covariates measured at baseline; age (in years), education (in years), smoking status (dichotomously: smoker/no smoker), alcohol use (dichotomously: $\leq$ or >7 units a week), measured body mass index

(BMI, continuously), number of treated chronic somatic disorders (ordinal), OC use (dichotomously: use/ no use), hormonal replacement therapy use (dichotomously: use/ no use), and menopausal status (dichotomously: postmenopausal / not postmenopausal). The definition of the covariates and the motivation for their use, are described in *Supplement 2*. Missing values on menopausal status were imputed with postmenopausal when women were aged ≥ 51 ($n = 21$; 1.2%). Missing values on the covariate alcohol use ($n = 24$; 1.4%) were imputed with the mean. Other covariates had no missing values. Results were back-transformed to the original scale to get the geometric mean with 95% confidence interval (CI) for visualization purposes. Using linear regression analysis, we additionally examined whether the potential association between androgen levels and MDD is driven by the severity on a continuous scale rather than the dichotomous diagnostic categorization. As a sensitivity analysis, we examined whether the association between androgen levels and severity of depressive symptoms was different between pre- and postmenopausal women by adding the interaction “androgen level*menopausal status” to the analyses. As a second sensitivity analyses, we examined whether the associations were better described by a non-linear relationship by adding quadratic terms of androgen levels to the analyses.

For the second and third hypothesis, we analyzed the association of androgen levels to both time to onset and time to remission of MDD using Cox proportional hazard models. We checked whether the analyses met the Cox proportional hazards assumption with appropriate log-minus-log curves. Again, models were adjusted for the covariates measured previously, and additionally for lifetime MDD or anxiety (time to first or recurrent MDD), or for lifetime anxiety disorder (time to remission of MDD). In sensitivity analyses, non-linear associations were examined by adding squared androgen levels to the models.

Finally, we explored whether altered levels of androgens in women with MDD - if present - might be mediated by salivary cortisol (AUC_g and AUC_i) using the indirect method by Preacher and Hayes.¹⁸⁸ This method estimates the total, direct, and indirect effects of the independent variable on the dependent variable through the mediator variable.

Data were analyzed using IBM SPSS Statistics (IBM Corp) version 25, using two-sided tests. As multiple tests were performed, we calculated an adjusted False Discovery Rate p cut-off value to avoid the inflation of false-positive findings. A p-value of < 0.0075 was considered statistically significant.

Results

The baseline characteristics of the 1659 women are shown in *Table 3.1*. Women who suffered from current MDD, followed successively by women with remitted MDD and never MDD, presented with the highest mean BMI scores (26.0, 25.1, and 24.7, $p < 0.001$), the highest mean number of chronic diseases (0.6, 0.5, and 0.4, $p < 0.001$), and the highest percentage of current smokers (45.4%, 39.0%, and 24.0%, $p < 0.001$). Levels of total testosterone, free testosterone, 5 α -DHT, androstenedione and DHEAS were lower, while SHBG was elevated in OC users compared to non-users. See *Supplement 3*.

Table 3.1. Characteristics of 1659 women at baseline.

	Never MDD n = 392 ^a	Remitted MDD n = 754 ^a	Current MDD n = 513 ^a	p
Age , mean (SD)	40.7 $\pm$ 14.4	41.4 $\pm$ 12.7	40.6 $\pm$ 12.2	.49
Education , mean (SD)	12.7 $\pm$ 3.2	12.5 $\pm$ 3.2	11.4 $\pm$ 3.2	<.0001
Body Mass Index , mean (SD)	24.7 $\pm$ 4.7	25.1 $\pm$ 5.1	26.0 $\pm$ 5.9	<.0001
Smoking , no. (%)	94 (24.0)	294 (39.0)	233 (45.4)	<.0001
>1 alcohol unit/day , no. (%)	52 (13.3)	101 (13.4)	76 (14.8)	.73
Treated chronic diseases , geo. mean (95%CI)	0.4 (0.0-1.9)	0.5 (0.0-2.3)	0.6 (0.0-2.7)	<.0001
Oral contraceptive use , no. (%)	123 (31.4)	192 (25.5)	137 (26.7)	.10
Hormonal replacement therapy use , no. (%)	7 (1.8)	16 (2.1)	5 (1.0)	.02
Menopausal , no. (%)	141 (36.0)	266 (35.3)	145 (27.7)	.02
Typical cycle length , mean (SD) ^b	30 (12)	28 (4)	28 (4)	.16
Days since last day of menstruation , mean (SD) ^b	20 (40)	24 (68)	30 (77)	.40
Antidepressant use , no. (%)	5 (1.3)	212 (28.1)	225 (43.9)	<.0001
Current anxiety disorder , no. (%)	0 (0.0)	273 (36.2)	350 (68.3)	<.0001
Severity of depressive symptoms , geo. mean (95%CI)	8.3 (0.0-26.1)	18.2 (2.8-46.5)	34.8 (17.1-62.2)	<.0001

Abbreviations: CI, confidence interval; MDD, major depressive disorder; SD, standard deviation. ^a Based on one-way analysis of variance for independent samples or χ^2 tests. ^b In subsample of premenopausal women not using oral contraceptive pills (n = 670).

Cross-sectional associations

The (adjusted) mean plasma levels of androgens per group are depicted in *Figure 3.1* (see detailed two-way analyses of covariance results in *Supplement 4*, and raw data points as well as data distribution in *Supplement 5*). After adjusting for covariates, only free testosterone levels in women with current MDD were significantly higher

(by 12.6%) compared to those with never MDD (adjusted geometric mean never and current MDD: 7.55 and 8.50). Levels of SHBG, 5 α -DHT, androstenedione and DHEAS did not differ between the groups. All covariates except education, were associated with at least one of the androgen or SHBG levels, but OC use and age explained most of the variance. See the table in *Supplement 2* for the F-statistics and p-values of the association of each covariate with each of the androgen and SHBG levels.

As shown in *Figure 3.2* (data in *Supplement 6*), none of the androgens or SHBG were significantly associated with depressive symptom severity. There was also no difference in the strength of the associations between pre- and postmenopausal women (adjusted p-value for the interaction of menopausal status with total testosterone p = 0.80, SHBG, p = 0.89, free testosterone p = 0.54, 5 α -DHT p = 0.47, androstenedione p = 0.34, and DHEAS p = 0.81). Non-linear associations (through quadratic terms) were also all non-significant.

Post-hoc analyses

Given the generally negative findings, we used a two one-sided tests procedure to test for equivalence and reject the presence of a smallest effect size of interest. This procedure examines whether the hypothesis that there are effects extreme enough to be considered meaningful can be rejected.¹⁸⁹ As there is no smallest effect size of interest known for androgens in women with regards to depression, we used 0.5 * SD of the baseline sample as a proxy.¹⁹⁰ The analyses showed that all mean differences between women with current or remitted MDD, and women who never had a MDD, were statistically equivalent. See also *Supplement 7*.

Mediation analysis

Free testosterone levels were positively associated with salivary cortisol levels, though only significant for AUC_g (p = 0.0006) but not for AUC_i (p = 0.09). However, neither AUC_i nor AUC_g mediated the association between free testosterone and MDD.

Figure 3.1. Mean androgen and SHBG levels in women according to their diagnosis at baseline.

Abbreviations: 5 α -DHT, 5 α -dihydrotestosterone; DHEAS, dehydroepiandrosterone-sulphate. Data are geometric means with 95% confidence intervals. P-values are based on adjusted mean analyses by two-way analyses of covariance models. Models were adjusted for age, education, body mass index, smoking, alcohol use, number of treated chronic diseases, menopausal status, oral contraceptive pill use, and hormonal replacement therapy use. To convert nmol/L to ng/dL for total testosterone, 5 α -DHT, and androstenedione multiply by 28.84, 29.07, and 28.64, respectively. To convert pmol/L to pg/ml for free testosterone divide by 3.47. To convert μ mol/L to μ g/dl for DHEAS, multiply by 36.85. To convert to nmol/L to mg/ml for SHBG, multiply by 0.095.

Figure 3.2. Androgen and SHBG levels in women according to severity of depression.

Abbreviations: 5 α -DHT, 5 α -dihydrotestosterone; DHEAS, dehydroepiandrosterone-sulphate; SHBG, sex hormone binding globulin. Figure shows adjusted standardized beta-coefficients for the association between androgen levels and the depressive symptom severity scores, examined with linear regression analysis. Analyses were adjusted for: age, education, body mass index, smoking, alcohol use, number of treated chronic diseases, menopausal status, oral contraceptive pill use, and hormonal replacement therapy use. Depressive symptom severity score as measured with the self-report Inventory of Depressive Symptoms. To convert nmol/L to ng/dL for total testosterone, 5 α -DHT, and androstenedione multiply by 28.84, 29.07, and 28.64, respectively. To convert pmol/L to pg/ml for free testosterone divide by 3.47. To convert μ mol/L to μ g/dl for DHEAS, multiply by 36.85. To convert to nmol/L to mg/ml for SHBG, multiply by 0.095.

Prospective associations

Table 3.2 shows the prospective association of androgen levels with incidence and remission of MDD, respectively. During four years of follow-up, 360 women out of 1027 women with no MDD at baseline (35.1%) experienced a first or recurrent episode of MDD. None of the androgens predicted time to incident MDD. Non-linear associations (through quadratic terms) were also non-significant. Conversely, MDD episodes remitted during follow-up in 254 women out of 405 women with MDD at baseline (62.7%). Again, none of the androgens predicted time to remission of MDD and non-linear associations (through quadratic terms) were also non-significant.

Table 3.2. Time to new/recurrent episode of MDD and time to remission of MDD in women according to androgen and SHBG levels during up to 4 years of follow-up.

	First or recurrent MDD				Remission MDD			
	HR (95%CI)	p	HR (95%CI)	p ^a	HR (95%CI)	p	HR (95%CI)	p ^a
<i>Events</i>	360 / 1027 (35.1%)				254 / 405 (62.7%)			
Total	1.14	.02	1.08	.17	1.05	.48	1.05	.49
testosterone	(1.02-1.26)		(0.97-1.21)		(0.93-1.18)		(0.92-1.20)	
SHBG	0.98	.75	1.04	.55	1.01	.92	1.00	.98
	(0.89-1.09)		(0.91-1.12)		(0.89-1.14)		(0.86-1.16)	
Free testosterone	1.12	.03	1.05	.40	1.03	.65	1.04	.59
	(1.01-1.25)		(0.93-1.19)		(0.91-1.16)		(0.90-1.21)	
5α-DHT	1.05	.38	1.04	.55	1.05	.43	1.04	.58
	(0.94-1.17)		(0.92-1.17)		(0.94-1.17)		(0.91-1.17)	
Androstenedione	1.15	.01	1.11	.11	1.07	.31	1.06	.41
	(1.03-1.29)		(0.98-1.25)		(0.94-1.20)		(0.92-1.23)	
DHEAS	1.06	.31	1.06	.37	1.09	.14	1.10	.20
	(0.95-1.18)		(0.93-1.20)		(0.97-1.22)		(0.95-1.26)	

Abbreviations: CI, confidence interval; MDD, major depressive disorder; 5 α -DHT, 5 α -dihydrotestosterone; DHEAS, dehydroepiandrosterone-sulphate; HR, hazard ratio; SHBG, sex hormone binding globulin. Data are HR's for 1 standard deviation change in each biomarker based on cox regression analyses. ^a Adjusted for age, education, body mass index, smoking, alcohol use, number of treated chronic diseases, oral contraceptive pill use, hormonal replacement therapy use, and menopausal status. The models for time to first or recurrent MDD were also adjusted for lifetime MDD or anxiety, and models for time to remission of MDD were additionally adjusted for lifetime anxiety disorder

Discussion

In this study, we examined the cross-sectional and longitudinal association between androgens and the presence and course of MDD in pre- and postmenopausal women aged 18 to 65. In contrast to our first hypothesis, only plasma free testosterone levels were significantly higher in women with current MDD compared

to never MDD, but this difference was not large enough to be considered clinically meaningful (statistically equivalent). Still, this group difference was independent of important confounders such as age and OC use, and was not explained by hyperactivity of the HPA-axis. All other differences in plasma androgen and SHBG levels between women with current or remitted MDD and never MDD were not statistically significant and also statistically equivalent. Moreover, none of the androgens predicted the time to a new MDD during 4 years of follow-up. Hence, the general negative findings of this study, suggest that plasma androgens do not play an important role in the pathogenesis of depressive disorders in women.

Our finding that women with current MDD had higher levels of free testosterone than women with never MDD differs from results from previous studies which found similar levels for depressed and non-depressed women.^{173, 178, 191} However, previous studies used radioimmunoassays rather than LC-MS/MS,^{173, 178} or lifetime rather than current MDD for group definition,¹⁸⁰ and hence may have lacked sensitivity to detect subtle differences. Yet, in contrast to free testosterone, total testosterone levels rendered significance after correction for multiple testing and covariates. This finding is in line with two studies in 215 and 634 women showing no significant differences in total testosterone levels between women with and without MDD.^{177, 178} Still, a large previous study in 3302 women revealed that higher testosterone levels were slightly more common among depressed women compared to non-depressed controls.^{173, 174} However, this finding would not have remained statistically significant if a multiple testing correction would have been performed. Therefore, the overall non-significant differences in androgen levels between women with and without MDD after the adjustment for covariates, underscore the importance of the execution of such sensitivity analyses.

Interestingly, free testosterone but none of the other androgens were associated with MDD. This may be explained by the fact that other androgens, in contrast to free testosterone, have a higher binding affinity to SHBG and/or albumin.¹⁹² Hence, free testosterone is the only androgen that can freely pass through the blood-brain barrier. In support of this hypothesis, 5 α -DHT has a stronger affinity to binding proteins than total testosterone and androstenedione, and indeed the associations between the latter two and MDD were stronger than for 5 α -DHT. Once free testosterone has entered the brain, there are a myriad of mechanisms by which it may affect mood. It can for example bind to androgen receptors in brain regions involved in affect regulation.^{48, 50} Free testosterone

can also interact with other important regulatory systems implicated in mental illnesses, such as neurotransmitter and immune systems.^{36, 48, 77, 80} On the other hand, free testosterone might also be altered as a consequence of MDD, for example, due to a chronic stress response. Although cortisol and free testosterone were positively associated, salivary cortisol did not mediate the association between free testosterone and MDD. If anything, one should bear in mind that although free testosterone levels were on average higher in women with current MDD compared to women with never MDD, this difference was found to be statistically equivalent. This suggests that, although present, the difference was of no or limited clinical relevance.

In contrast with our second hypothesis, we did not find an association between androgens and incident MDD during four years of follow-up. This is in concordance with previous prospective studies which followed 980, 1711, and 3840 adult women from the general population for up to 10 years and showed no association of total testosterone with incidence of MDD.¹⁷⁹⁻¹⁸¹ In contrast, one previous study showed that an annual increase in total testosterone was associated with increased odds of depressive symptoms in perimenopausal women using the self-report CES-D, however, in this study, testosterone was measured with immunoassays, and analyses were not adjusted for lifetime MDD.¹⁹³ This is important because lifetime MDD presence is the main predictor for a future episode.¹⁹⁴

Additionally, in contrast to our third hypothesis, we did not find an association between baseline androgens levels and time to remission. This suggests that though androgens might be associated with a current diagnosis of MDD, it has no or only limited predictive power. Altogether, there is little evidence for a temporal association between single androgen levels within the physiological range and the development of MDD in women, but the current data cannot reject the possibility that changes in androgen levels are of importance to the development of MDD in women as we only measured androgens at baseline. Studying testosterone during windows of development in which it may also cause permanent effects on cerebral functioning, such as antenatally or during puberty, may be an intriguing alternative direction of research.⁴⁹

Strengths of this study include the examination of a well-phenotyped sample where psychiatric diagnoses were based on structured diagnostic interviews (CIDI) and the sample size was principally capable of detecting clinically significant differences between groups. Importantly, we used LC-MS/MS for the determination

of total testosterone, androstenedione, and 5 α -DHT, which is imperative given the fact that detection of physiologic levels of androgens that are seen in women (e.g., testosterone <5.0 nmol/L) are less reliable when using immunoassays levels due to cross reactivity with other hormones.⁸⁹ Furthermore, we were able to adjust for a wide range of confounders. As confirmed by the differences in androgen levels between OC users and non-users in our study, OC use was associated with lower levels of androgens, especially total testosterone, likely through inhibiting ovarian and adrenal androgen synthesis and by increasing levels of SHBG.^{195, 196}

Some limitations of our study merit further discussion. First, because we used observational data, we cannot conclude whether elevated free testosterone levels in MDD are a cause or a consequence of the disease. Second, peripherally measured androgens might lack the sensitivity to reflect androgen levels in the brain. Though all androgens can pass the blood-brain-barrier, binding proteins limit the extent. Third, the lack of prospective associations might be because changes in free testosterone were a consequence of MDD rather than a cause, and because single measurements may be insufficient to predict psychopathology over longer time periods, as changes in levels may be driving a possible relationship. Additionally, androgens have diurnal variation, vary by the phase of the menstrual cycle, and decline with age.^{197, 198} Nonetheless, all blood draws were done in the morning with little variance, and there was no statistical difference in our findings related to menopausal status or days since the last menstrual cycle of those women not using OCs and not being menopausal between the psychopathological groups. Moreover, we have adjusted for important covariates. Therefore, strong associations would resist this limitation, suggesting that, if present at all, prospective associations would have been weak. Finally, although the sociodemographic characteristics of the sample with and without follow-up were comparable and the attrition rate was low (13.7%), the sample without follow-up had more depressive symptoms than the sample with follow-up at baseline. This may point to a selection bias for the longitudinal analyses.

In summary, this study showed no differences in most androgen levels between adult women with and without MDD. Higher plasma levels of free testosterone in women with MDD compared to women without a history of MDD were found, independent of important confounders like age and OC use, but these levels did not predict future MDD. Therefore, this study suggests that the role of plasma androgen levels in women's vulnerability for MDD is limited.

Acknowledgements

The infrastructure for the NESDA study (www.nesda.nl) is funded through the Geestkracht program of the Netherlands Organisation for Health Research and Development (ZonMw, grant number 10-000-1002) and financial contributions by participating universities and mental health care organizations (Amsterdam University Medical Centers (location VUmc), GGZ inGeest, Leiden University Medical Center, Leiden University, GGZ Rivierduinen, University Medical Center Groningen, University of Groningen, Lentis, GGZ Friesland, GGZ Drenthe, Rob Giel Onderzoekscentrum). H Joffe's research program is supported by grants from V Foundation, Merck, Pfizer, Que-Oncology, and NeRRe/KaNDy. H Joffe is also a consultant for NeRRe/KaNDy, Merck, Sojournix, Eisai, and Jazz Pharmaceutical. BWJH Penninx has received research grants from Jansen Research and Boehringer Ingelheim. Other authors have nothing to disclose. The funding source had no role in the design and conduct of the study, collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Chapter 4

Plasma androgens and the presence and course of depression in a large cohort of men

Anouk E de Wit
Erik J Giltay
Marrit K de Boer
Willem A Nolen
Fokko J Bosker
Brenda WJH Penninx
Robert A Schoevers

Adapted version from the article published in Psychoneuroendocrinology, May 2021.

Abstract

Aim: Hypoandrogenic men showed a higher prevalence of major depressive disorder (MDD), which could be ascribed to overlapping symptoms such as sexual dysfunction, or additionally to core emotional symptoms such as sadness and anhedonia. We examined whether androgen levels 1) differ between men with and without MDD cross-sectionally, 2) are associated with an elevated risk for onset of MDD prospectively, and 3) associate with all individual MDD symptoms, or only with hypogonadism overlapping symptoms.

Methods: In 823 men (mean age 43.5 years), baseline plasma levels of total testosterone, 5 α -dihydrotestosterone (5 α -DHT), and androstenedione were determined with liquid chromatography–tandem mass spectrometry, and dehydroepiandrosterone-sulphate (DHEAS) and sex hormone binding globulin with radioimmunoassays, whereas free testosterone was calculated. MDD status was assessed at baseline and after two years using structured interviews and individual MDD symptoms were self-rated at baseline, and after one and two years.

Results: None of the androgen levels were associated with current or onset (incidence or recurrence) of MDD. Free testosterone was only inversely associated with interest in sex. Also, androstenedione and DHEAS were positively associated with some individual MDD symptoms, and 5 α -DHT levels showed non-linear associations (both with low and high levels) with MDD symptom severity and several individual MDD symptoms.

Conclusion: These results support the idea that circulating androgens synthesized by the testes are of limited clinical relevance to MDD in adult men, but levels of androstenedione, DHEAS and 5 α -DHT may be associated with some individual MDD symptoms.

Introduction

The idea that adequate testosterone levels serve a protective function for depression in men, was originally rooted in the observation that hypogonadism is associated with more depressive symptoms.¹⁹⁹ Such a relationship may be a consequence of sex steroid receptor bindings of testosterone that result in serotonin neurotransmitter function alterations or other genomic and non-genomic changes in the limbic system.^{47, 200} However, such a relationship may also be the indirect result of shared symptoms of hypogonadism and major depressive disorder (MDD), such as sexual dysfunction, apathy and fatigue,²⁰¹⁻²⁰³ that may additionally also have an effect on mood. To date, the association between testosterone and MDD remains disputed because studies examining the association between testosterone and MDD have yielded conflicting results.

The majority of cross-sectional studies that used structured diagnostic assessment procedures to identify MDD symptomatology and diagnoses reported no association between total or free testosterone and MDD^{177, 204, 205} or severity of MDD symptoms.²⁰⁶⁻²⁰⁸ These findings were confirmed by prospective studies showing that neither free nor total testosterone levels, nor a change in these levels, predicted the severity of MDD symptoms at two- or three-years follow-up.^{182, 208} However, some other studies have shown that men with low or lower levels of free or total testosterone are more often depressed,^{182, 209} have increased severity of MDD symptoms,^{206, 210} or have an increased risk for first onset of MDD after two years.^{211, 212}

Besides testosterone, there are also other endogenous androgenic compounds circulating in the blood. An example includes 5 α -dihydrotestosterone (5 α -DHT) which is substantially more potent than testosterone as an agonist of the androgen receptor.²¹³ However, studies examining 5 α -DHT in relation to depression are limited,²¹⁴ mainly because the previous gold-standard method for measuring androgens (radioimmunoassays) was less reliable for this compound.²¹⁵ However, the currently widely available liquid chromatography–tandem mass spectrometry (LC-MS/MS) method is capable of reliably measuring 5 α -DHT as well as other androgens, including the lower levels that are present in hypogonadal men.⁸⁹

To clarify the possible role of androgens in depression among men, future research should examine which specific symptoms drive the association between androgens and depression. Knowing if, and if so, which specific MDD symptoms are related to altered androgen levels in men, may provide insight into the

pathophysiological mechanisms underlying this type of depression in men.²¹⁶ If anything, future research should also overcome previous (methodological) limitations of studies that hampered the ability to make definitive conclusions about the association of androgens and depression. These limitations include scarcity of results reported on the full range of testosterone levels as many examined hypogonadal levels and the focus on free and total testosterone while many more androgenic compounds circulate in the blood. Additionally, limited number of studies have had prospective designs, precluding the assessment of temporal relationships. Finally, future research should also examine non-linear effects as both low and high levels of androgens, may be related to MDD symptoms.^{201, 217}

In this study, we examined whether androgens levels 1) differ between men with and without MDD cross-sectionally; 2) are associated with an elevated risk for onset of MDD prospectively; and 3) are associated with all individual MDD symptoms or only with subset of symptoms. We used data from a large cohort of men from the Netherlands who were representative of men with depressive disorders in different health care and population settings and stages of the developmental history and in whom total testosterone, androstenedione, and 5 α -DHT were determined with LC-MS/MS and dehydroepiandrosterone-sulphate (DHEAS) with radioimmunoassay. Given the predominantly negative findings in previous research, we a priori hypothesized that lower androgen levels are: 1) equally present in men with and without MDD; 2) not associated with an elevated risk for new onset of MDD; and 3) only associated with some overlapping symptoms of MDD and hypogonadism such as less sexual interest and more fatigue.

Methods

Study population

Netherlands Study on Depression and Anxiety (NESDA) is a large multicenter study designed to investigate the long-term course and consequences of depressive and anxiety disorders. Between 2004 and 2007, men and women (aged 18-65 years) were recruited from different health care settings (community, primary care and specialized mental health care) and in different stages of the developmental history of disorders (healthy controls, high familial risk, subthreshold disorders, first and recurrent episodes) in order to be representative of those with depressive and anxiety disorders in different health care settings and stages of the developmental history. Exclusion criteria included not speaking Dutch fluently, or being diagnosed

with one of the following psychiatric diseases: psychotic, obsessive compulsive, bipolar, or substance use disorder.¹³⁵ For the baseline assessment, 26393 men and women were contacted of whom 2981 (11.3%) participated. The ethical committees of the participating centers approved (VU Medical Center, University Medical Center Groningen, Leiden University Medical Center) the study design and all participants gave verbal and written informed consent.

For this study, 1002 men were considered for inclusion. Exclusion criteria included men who had missing plasma androgen measurements ($n = 24$), extreme androgen values (> 8 standard deviation [SD] below or above the mean, $n = 8$), who were transgender ($n = 2$), or in whom a measurement or sampling error was likely ($n = 1$). Men who used androgens World Health Organization Anatomical Therapeutic Chemical (ATC) G03B and E, gonadotrophins (ATC G03G), antiandrogens (ATC G03H), gonadotrophin-releasing hormone analogues (ATC L02AE), or antiandrogens (ATC L02BB) at the time of the blood draw were also excluded ($n = 4$). Finally, men who were diagnosed (lifetime) with an anxiety disorder but not with MDD ($n = 140$) were excluded as well. Exclusion of these men following the above-mentioned criteria resulted in a baseline sample of 823 participants (82.1% of all men) to test the first hypothesis. This sample was categorized into three groups according to their psychopathology status; without lifetime MDD (neither currently nor in the past depressed, $n = 246$), remitted MDD (lifetime MDD, but currently not depressed, $n = 318$), and current MDD (currently depressed, $n = 259$). Of the sample of men with remitted MDD or without lifetime MDD ($n = 564$), 504 men (89.4% of the non-depressed baseline sample) did participate in the follow-up assessment after two years and were used to test the second hypothesis. Compared to men with follow-up, men without follow-up ($n = 60$, 10.6%) more often had lifetime MDD (76.7% vs. 54.0%) and low free testosterone levels (18.3% vs. 8.9%; *Supplement 8*). Finally, at baseline and after one and two years of follow-up, the severity of individual MDD symptoms was assessed. Out of the baseline sample, 751 men (91.3%) had responded to at least two out of three of the items during these assessments and were therefore included in the analyses to test the third hypothesis. Men without follow-up ($n = 72$, 8.7%) were comparable to men without follow-up regarding psychopathology, but smoked more often (62.5% vs. 39.5%; *Supplement 8*). A flow diagram of men included in the analyses to test each of the three hypotheses is shown in *Supplement 9*.

Assessment of androgens and SHBG

Blood plasma collection took place in the morning (mean time 08:46h SD 32min) after an overnight fast (success rate 94.4%). Total testosterone, 5 α -DHT, and androstenedione were determined by LC-MS/MS at the Clinical Chemistry department of the University Medical Centre Groningen. The lower limit of detection was 0.04 nmol/L for total testosterone and androstenedione, and 0.12 nmol/L for 5 α -DHT. The inter-assay coefficients of variation for all androgens were low (total testosterone: 3.1%, 5 α -DHT: 7.8%, androstenedione: 2.2%). DHEAS and sex hormone binding globulin (SHBG) were determined using ARCHITECT (ABBOTT, Wiesbaden, Germany), a one and two-step radioimmunoassay with chemiflex assay protocols. The lower limit of detection of DHEAS was 0.3 μ mol/L with a calibration range of 0.00-40.71 μ mol/L, and 0.1 nmol/l for SHBG with a calibration range of 0.0-250 nmol/L. Free testosterone was calculated based on SHBG and total testosterone levels, and an assumed albumin level of 43 g/L with the formula of Vermeulen.¹⁴¹

Assessment of depressive disorders

The presence of MDD, currently or in the past, was ascertained using the lifetime version of the Composite International Diagnostic Interview (CIDI), version 2.1, a validated instrument with high interrater reliability (any depressive disorder kappa = 0.95) and high validity for depressive and anxiety disorders.^{185, 186} Current MDD was ascertained according to Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV-based CIDI criteria within the past month, remitted MDD was determined when participants had MDD earlier in life, but not within in the last month. First or recurrent onset of MDD between baseline and 2 years follow-up, was ascertained when men without MDD at baseline did fulfil the criteria of MDD during follow-up. Finally, at baseline and after one and two years of follow-up, the severity of individual MDD symptoms was assessed with the self-report Inventory of Depressive Symptoms.¹⁸⁷ This questionnaire consists of 30 items that assess the symptom severity over the last seven days. Besides the core MDD symptoms, this questionnaire also assesses the severity of symptoms that overlap with hypogonadism such as interest in sex (measured as no or little desire, derived pleasure or interest in sex) and fatigue. Each item is scored with a value between 0 and 4, with higher values indicating increased severity. This questionnaire had excellent internal consistency in previous research¹⁸⁷ and at our baseline assessment (Cronbach's α = 0.92).

Covariates

The following covariates measured at baseline were taken into account in the models described below the next header: age (continuously), waist circumference (cm), number of treated chronic somatic disorders (continuously), smoking status (dichotomously), alcohol use (dichotomously), physical activity (continuously), and the presence of a current anxiety disorder (dichotomously). Age was measured in years and waist circumference in cm. Smoking status was dichotomized in smoker and no smoker, and alcohol use was dichotomized in $\leq$ or >7 units a week. Physical activity was measured as the total sum of MET-minutes for vigorous, moderate and walking activities as measured with the International Physical activity Questionnaire. Finally, presence of current anxiety disorders was, like for MDD, measured with the lifetime version of the CIDI, version 2.1. Missing values on alcohol use ($n = 13$; 2.9%), waist circumference ($n = 2$, 0.2%), and physical activity ($n = 44$, 5.3%) were imputed with the mean. Other covariates had no missing values.

Statistical analyses

At baseline, group characteristics of without lifetime MDD, current MDD, and remitted MDD were compared using one-way analyses of variance for independent samples or Kruskal Wallis in case of continuous variables with normal or non-normal distribution, respectively, and with χ^2 tests in case of categorical variables.

To test our first hypothesis, we ran two-way analyses of covariance to test whether androgen levels varied by group status, and used two one-sided tests procedures to reject the presence of a smallest effect size of interest.¹⁸⁹ The latter was done since non-significant p-values obtained via classical testing would not declare the absence of meaningful effects. We considered 0.20 Cohen's d as the minimally clinically meaningful important difference.²¹⁸ To test our second hypothesis, we ran logistic regression analyses with baseline androgen levels as predictors and the presence or absence of first or recurrent onset of MDD between baseline and two years as the outcome measure in men without MDD at baseline. To test our third hypothesis, we ran 186 linear mixed models considering baseline androgen and SHBG levels as fixed factors, and total and individual MDD symptom severity scores as outcome measures. A random intercept and random slope were added to the model as they lowered the Akaike information criterion. A random intercept accounted for between-participant variability of severity of

MDD symptoms and the random slope (time) allowed for individual slopes for MDD symptoms over time. An interaction between time and androgens was not added as it decreased the model fit. Mixed models with unstructured covariance structures were fitted using restricted likelihood estimation, but the final model was estimated with maximum likelihood estimation. All models for the testing hypotheses 1, 2, and 3, were adjusted for the above-described covariates. The models for examining hypothesis 2 were additionally adjusted for presence of remitted MDD.

In sensitivity analyses, we examined whether the results of the first, second and third hypothesis testing changed when free and total testosterone levels were dichotomized into low and normal. In line with previous research, cut-off points of <11 nmol/L for total testosterone and <220 pmol/L for free testosterone were used.²⁰¹ We used logistic regression analyses to test the first hypothesis with low free and total testosterone presence as the outcome measure, presence of current, remitted or without lifetime MDD (reference group) as predictor, and adjusted for the same covariates as mentioned above. The second and third hypothesis were tested in the same way as described above, except that continuous androgen levels were replaced with the presence of low free or total testosterone.

As a second sensitivity analyses, we examined whether the associations between androgens and individual MDD symptoms (hypothesis 3) were better described by a non-linear relationship by adding quadratic terms of androgen levels to the regression models.

Data were analyzed using R (version 1.2.1335). In all models, the predictor and outcome (if not dichotomous) were standardized (into z-values) to ease comparability of effect sizes. In order to control for type-I-errors, we used the false discovery rate using the use the Benjamini-Hochberg procedure to adjust the p-values generated by the analyses examining hypothesis 1 and 2.²¹⁹ For the analyses conducted for hypothesis 3, we considered a p-value of < 0.05 being statistically significant given the more explorative character of this hypothesis.

Results

The characteristics of the 823 men are shown in *Table 4.1*. Men with current MDD presented with the highest mean waist circumference and the highest percentage of current smokers, followed successively by men with remitted MDD and without lifetime MDD.

Table 4.1. Characteristics of 823 men at baseline.

	Without lifetime MDD n = 246	Remitted MDD n = 318	Current MDD n = 259	p ^a
Age , mean (SD)	42.3 (15.1)	44.4 (12.0)	43.7 (11.3)	.15
Education (years) , mean (SD)	13.0 (3.2)	12.0 (3.3)	11.2 (3.0)	<.001
Body Mass Index , mean (SD)	25.6 (4.3)	26.4 (4.6)	26.7 (4.7)	.02
Waist circumference , mean (SD)	94.5 (12.6)	96.8 (12.7)	97.0 (13.3)	.04
Current smoking , no. (%)	78 (31.7)	143 (45.0)	121 (46.7)	<.001
>1 alcohol unit/day , no. (%)	81 (32.9)	91 (28.6)	75 (29.0)	.49
Physical activity , mean (SD)	3420 (1691-5062)	3052 (1386-5486)	2598 (1095-4418)	.02
≥1 treated chronic diseases , no. (%)	163 (33.7)	137 (43.1)	129 (48.8)	.08
Current antidepressant use , no. (%)	4 (1.6)	91 (28.6)	110 (42.5)	<.001
Current anxiety disorder , no. (%)	0 (0.0)	107 (33.6)	153 (59.1)	<.001
Severity of MDD symptoms , median (IQR)	5 (2-10)	19 (11-27)	36 (28-44)	<.001

Abbreviations: IQR, interquartile range; MDD, major depressive disorder; SD, standard deviation. ^a Based on one-way analyses of variance for independent samples or χ^2 tests for normally distributed variables and by Kruskal Wallis testing for non-normally distributed variables (severity of MDD symptoms).

Major depressive disorder

Mean total and free testosterone, SHBG, androstenedione, and DHEAS levels were not statistically different between men without lifetime, and men with remitted or current MDD (*Figure 4.1*), and equivalence testing showed that differences examined, were equal to zero (*Supplement 11*). Only the results on 5 α -DHT were inconclusive, as levels were not statistically different between the groups, but also not statistically equivalent to zero.

During two years of follow-up, 99 (19.6%) out of 504 men without MDD at baseline experienced a first or recurrent onset of MDD. However, none of the androgens statistically significantly predicted this event (*Table 4.2*).

Figure 4.1. Adjusted mean androgen and SHBG levels in 823 men according to their diagnosis at baseline.

Abbreviations: 5α-DHT, 5α dihydrotestosterone; DHEAS, dehydroepiandrosterone-sulphate. Data are estimated means with 95% confidence intervals based two-way analyses of covariance models adjusted for age, waist circumference, number of treated chronic somatic disorders, smoking status, alcohol use, physical activity, the presence of a current anxiety disorder. To convert nmol/L to ng/dL for total testosterone, 5α-DHT, and androstenedione multiply by 28.84, 29.07, and 28.64, respectively. To convert pmol/L to pg/ml for free testosterone divide by 3.47. To convert μmol/L to μg/dl for DHEAS, multiply by 36.85. To convert to nmol/L to mg/ml for SHBG, multiply by 0.095. The data for this figure are given in *Supplement 10*.

Individual depressive symptoms

None of the androgens examined was associated with the total MDD symptom severity score over two years. Lower free testosterone levels were associated with less interest in sex, but other individual MDD symptoms over two years were not associated with free or total testosterone, nor with SHBG or 5α-DHT levels (*Figure 4.2*). Higher androstenedione levels however, were associated with: pessimism, suicidality, anxiety or tension, and less increase in weight. Also, higher DHEAS levels were positively associated with lower self-esteem, suicidality, decrease in weight and pessimism.

Table 4.2. Odds ratio's for first or recurrent episode of MDD in 504 men according to plasma androgen and SHBG levels during up to two years of follow-up.

	First or recurrent MDD			
	OR (95%CI)	p	OR (95%CI) ^a	p
<i>Events</i>	99 / 504 (19.6%)			
Total testosterone	1.03 (0.81-1.29)	.82	1.01 (0.78-1.30)	.93
SHBG	0.98 (0.77-1.23)	.88	0.98 (0.74-1.27)	.89
Free testosterone	1.01 (0.80-1.27)	.94	0.96 (0.72-1.28)	.80
5α-DHT	0.95 (0.76-1.19)	.68	0.94 (0.73-1.19)	.61
Androstenedione	1.03 (0.82-1.28)	.82	0.99 (0.77-1.27)	.92
DHEAS	1.00 (0.78-1.25)	.97	0.93 (0.69-1.24)	.65
Low total testosterone (<11 nmol/L)	0.91 (0.45-1.76)	.80	0.94 (0.44-1.87)	.86
Low free testosterone (<220 pmol/L)	0.59 (0.21-1.38)	.26	0.55 (0.19-1.37)	.23

Abbreviations: 5 α -DHT, 5 α -dihydrotestosterone; CI, confidence interval; DHEAS, dehydroepiandrosterone-sulphate; MDD, major depressive disorder; OR, odds ratio; SHBG, sex hormone binding globulin. Data are OR's for 1 standard deviation change in each biomarker based on logistic regression analyses. ^a Adjusted for age, waist circumference, number of treated chronic somatic disorders, smoking status, alcohol use, physical activity, the presence of a current anxiety disorder, and the presence of remitted MDD.

Sensitivity analyses: low (hypogonadal) free and total testosterone

The sensitivity analyses confirmed the robustness of the results described above. Compared to men without lifetime MDD, men with current or remitted MDD were not more likely to have low free (adjusted odds ratio [OR] with 95% Confidence Interval [CI] current MDD 1.27 [0.57; 2.47], $p = 0.64$, and remitted MDD 1.06 [0.55; 2.05], $p = 0.87$) or total testosterone levels (current MDD 1.41 [0.81; 2.50], $p = 0.11$, and remitted MDD 1.67 [0.90; 3.12], $p = 0.23$). Moreover, baseline presence of low free or total testosterone did not predict the onset of MDD during two years of follow-up (*Table 4.2*). Furthermore, low free or total testosterone levels were not associated with increased total MDD symptom severity over two years in the adjusted analyses (*Supplement 15*). Only low free testosterone levels were associated with hypersomnia, and low total testosterone levels were associated with difficulty staying asleep and diurnal mood variation (*Figure 4.3*). Although low free testosterone levels were not significantly associated with less interest in sex, the β was in the same direction and of similar strength as the association of lower free testosterone levels with less interest in sex.

Sensitivity analyses: non-linear analyses individual major depressive disorder symptoms

Non-linear associations between androgens and individual MDD symptom severity scores over two years were also mostly non-significant, except for the associations with 5 α -DHT (*Supplement 16*). In that specific case, the quadratic term was statistically significant with mainly positive effect estimates, indicating that both lower and higher 5 α -DHT levels tended to be associated with increased total MDD symptom severity score, and with difficulty staying asleep, waking up early, low responsiveness of mood, pessimism, low energy level, low capacity for pleasure, psychomotor retardation, sympathetic arousal, and leaden paralysis.

◀Figure 4.2. Baseline plasma androgen and SHBG levels and their association with individual MDD symptoms over two years in 751 men.

Abbreviations: 5 α -DHT, 5 α -dihydrotestosterone; DHEAS, dehydroepiandrosterone-sulphate; MDD, major depressive disorder. Figure shows adjusted standardized betas with 95% Confidence Interval for the association between androgen levels and the individual MDD symptom scores over two years examined with a linear mixed model. The blue dots and lines represent significant associations, whereas the non-significant associations are printed in grey. All models were adjusted for age, waist circumference, number of treated chronic somatic disorders, smoking status, alcohol use, physical activity, the presence of a current anxiety disorder. The data for these figures are given in *Supplements 12 to 14*.

Discussion

In this study on the relationship between androgens and depression in men, we showed that none of the levels of different androgens, including those of free testosterone, were associated with current or new onset MDD, which was in line with our first and second hypothesis. Also, in line with our third hypothesis, lower free testosterone levels were only associated with less interest in sex. However, higher levels of androstenedione and DHEAS, and both lower and higher levels of 5 α -DHT were associated with several individual MDD symptoms. These results may suggest that, in contrast to what is often thought, testosterone is not of clinical relevance for MDD in men.

The finding that lower testosterone levels associate with some overlapping symptoms between MDD and hypogonadism but do not beget a higher risk of MDD, is in line with studies reporting that total and free testosterone levels did not differ between men with and without MDD¹⁷⁷ and had no predictive value for the development of MDD symptoms.^{182, 208} As similar results were found testing low (hypogonadal) levels of testosterone, our results add to the evidence of studies showing that sexual complaints, such as decreased libido, decreased morning erections or erectile dysfunction, are the most prominent symptoms in middle-aged and elderly men with low total or free testosterone levels.^{201, 220}

We observed for the first time a non-linear association between 5 α -DHT level and (several individual) MDD symptoms. This suggests that both lower and higher levels rather than solely lower levels, tended to be associated. Importantly, 5 α -DHT is the most potent androgen as an agonist of the androgen receptor, so also more potent than testosterone.²¹³ These results therefore highlight the importance of examining non-linear relationships when studying the association between androgens and depression, and the importance of measuring 5 α -DHT in future studies.

◀Figure 4.3. Presence of low free and total testosterone levels at baseline and their association with individual MDD symptoms over two years in 751 men.

Abbreviation: MDD, major depressive disorder. Figure shows adjusted standardized beta coefficients with 95% confidence interval for the association between androgen levels and the individual MDD symptom scores over two years examined with a linear mixed model. The blue dots and lines represent significant associations, whereas the non-significant associations are printed in grey. The analyses were adjusted for age, waist circumference, number of treated chronic somatic disorders, smoking status, alcohol use, physical activity, the presence of a current anxiety disorder. The data for these figures are given in *Supplement 15*.

In contrast to the testosterone findings, androstenedione and DHEAS levels were associated with some individual MDD symptoms. The fact that both androstenedione and DHEAS are precursors of total testosterone and have little androgenic activity itself, may suggest that these levels were elevated as a consequence of (processes related to) MDD, rather than that they directly caused the symptoms. For example, we propose that an altered hypothalamic-pituitary-adrenal-axis (HPA-axis) with a chronic stress response could underlie this phenomenon. The HPA signalling pathways may be altered such that hyperactivity of the HPA-axis and impaired negative feedback responses have been described in those patients with current and remitted MDD.^{184, 221} Aside from activating cortisol release, these pathways also have a stimulatory effect on the zona fasciculata and reticularis of the adrenal cortex to produce adrenal androgens, like DHEAS, and other androgens such as androstenedione which are not solely produced by the adrenals but still for a reasonable amount.²²² This is apparent as those with central adrenal insufficiency who lack adrenocorticotrophic hormone production, have a deficiency in cortisol and dehydro-epiandrosterone. Therefore, in a chronic stress-response model of MDD, it would seem that an enhanced HPA-axis activation of the adrenal cortex not only stimulates release of cortisol, but of adrenally produced androgens as well. Nevertheless, this and other mechanisms remain to be further explored.

Strengths of this study include the examination of a well-phenotyped sample where psychiatric diagnoses were based on structured diagnostic interviews and androgens determinations with LC-MS/MS, which is imperative given the detection of lower androgens levels in hypogonadal men.⁸⁹ Furthermore, we were able to differentiate associations with individual MDD symptoms, which enabled us to determine which symptoms could drive an association between androgens and MDD. Our study has also some limitations. First, because we used observational data, we cannot conclude whether elevated androstenedione levels are a cause or an effect of MDD. Second, the number of men without MDD at baseline was small, which limited the power to detect an association between androgens and onset of MDD. Moreover, the majority of these men had a history of MDD. Hence, we studied both new onset and recurrence of MDD in relation to androgen levels. Third, although androgens can pass the blood-brain-barrier, binding proteins limit the extent which may consequently limit the sensitivity of peripherally measured androgens to reflect androgen levels in the brain. Fourth, a repeated

measurement of androgen levels after a short interval would have increased the precision of the androgen levels measured. Fifth, the single measurement of androgens might have lacked the sensitivity to predict depression over a two-year course. Future studies should measure androgens at least twice as change levels provide more insights in the within-person variation. Alternatively, measurement of androgens in hair using LC-MS/MS would offer the unique chance to estimate androgen production over longer time periods.^{169, 223} Sixth, multiple testing may have induced type-I errors. Finally, free testosterone levels were calculated but not measured.

In conclusion, in our study we did not find differences in testosterone levels between adult men with and without MDD, nor an association of testosterone levels with individual MDD symptoms. Only lower free testosterone levels were associated with less interest in sex. Therefore, this study confirms the idea that single measured circulating androgens synthesized by the testes have a limited clinical relevance to MDD in adult men, although they may affect some individual symptoms that overlap with hypogonadism.

Acknowledgements

The infrastructure for the NESDA study (www.nesda.nl) is funded through the Geestkracht program of the Netherlands Organisation for Health Research and Development (ZonMw, grant number 10-000-1002) and financial contributions by participating universities and mental health care organizations (Amsterdam University Medical Centers (location VUmc), GGZ inGeest, Leiden University Medical Center, Leiden University, GGZ Rivierduinen, University Medical Center Groningen, University of Groningen, Lentis, GGZ Friesland, GGZ Drenthe, Rob Giel Onderzoekscentrum). BWJH Penninx has received research grants from Jansen Research and Boehringer Ingelheim. Other authors have nothing to disclose. The funding source had no role in the design and conduct of the study, collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Chapter 5

Associations between testosterone and metabolic syndrome in depressed and non- depressed older men and women

Anouk E de Wit
Erik J Giltay
Marrit K de Boer
Fokko J Bosker
Roos C van der Mast
Hannie C Comijs
Richard C Oude Voshaar
Robert A Schoevers

Adapted version from the article published in International Journal of Geriatric
Psychiatry, November 2018.

Abstract

Aim: Older age and major depressive disorder (MDD) are both risk factors for the development of cardiovascular diseases. Testosterone has been associated with MDD and metabolic syndrome (MetS) in men, though associations in women are less clear. Therefore, we investigated whether testosterone is associated with MetS and whether this association is different for depressed and non-depressed older men and women.

Methods: In this prospective cohort study, 478 participants (349 patients with MDD and 129 controls) aged between 60 and 93 years from the Netherlands Study on Depression in Older persons were included. Total testosterone and sex hormone binding globulin levels were measured using a second-generation radioimmunoassay. Free testosterone was calculated based on total testosterone. MetS was defined according to the adjusted National Cholesterol Education Program Adult Treatment Panel III criteria.

Results: A higher risk for MetS was found in men with lower free and total testosterone (odds ratio [95% confidence interval] 0.67 [0.47-0.95], and 0.51 [0.34-0.75]), and in women with higher free testosterone (1.41 [1.08-1.82]). Strong associations in the same direction were found with adiposity, glucose and plasma lipid MetS components at baseline, but not with changes in these components at two-year follow-up. The associations did not significantly differ between MDD patients and controls.

Conclusion: Independently of having MDD, lower testosterone levels in men, and in contrast, higher testosterone levels in women were significantly associated with MetS and its components.

Introduction

The older population shows the highest prevalence of metabolic syndrome (MetS)⁹³ which puts this group at increased risk for cardiovascular diseases and cardiovascular related mortality.²²⁴ Among the proposed mechanisms to explain the association between age and MetS is the influence of testosterone. Testosterone levels decline during aging, especially in men, who have 10-times higher levels than women in young adulthood.²²⁵

Low testosterone levels in men, and high levels in women are thought to alter adipose tissue function, with effects on lipid metabolism, insulin resistance and fat mass expansion which may induce metabolic disadvantages.⁹⁴ However, this association is complex, as testosterone levels have both detrimental and beneficial effects depending on the type of tissue and sex.⁹⁴ Earlier research in men has shown that lower testosterone levels may contribute to the development of MetS but that the magnitude of this effect is dependent on several other factors and that the relation may also be bidirectional.^{95, 96} In contrast to the findings in men, higher testosterone levels in women have been associated with MetS.⁹⁷ That said, prospective analyses are lacking and not all studies have found this association.²²⁶

In spite of a range of studies on this issue, many aspects of the complex relationship between testosterone and MetS remain unclear. This warrants further investigation into possible underlying pathophysiological processes, especially in the older population who are at the highest cardiovascular diseases mortality risk.²²⁷ Influencing factors may include major depressive disorder (MDD), which has been associated with 80-90% elevated risk for cardiovascular diseases onset.²²⁸ Additionally, the presence of MetS in depressed older adults has been associated with greater symptom severity of depression.²²⁹ Dysregulations in the immune response and the hypothalamic-pituitary-adrenal-axis (HPA-axis) seen in MDD⁹⁸ closely respond to alterations in the hypothalamic-pituitary-gonadal-axis (HPG-axis). This not only alters testosterone levels but also impacts its function.^{36, 94, 99} Indeed, low testosterone levels have been associated with MDD in men,^{86, 209, 212, 230} though associations in women are inconsistent.^{182, 209, 226, 231}

Given these associations, we hypothesized that MDD or the use of antidepressants increases the vulnerability for testosterone level alterations. Therefore, we aimed to determine the associations between testosterone and MetS in older adults and test whether this association is stronger in depressed or antidepressant-using patients.

Methods

Study population

Data for this study were obtained from Netherlands Study on Depression in Older persons (NESDO). For methodological details, we refer to previously published work.^{136, 232} In short, this multicenter prospective cohort aims to study the course, predictors and consequences of late-life depression. In total, 510 participants, aged 60 to 93 years, were recruited from 2007 to 2010. A group of 378 older persons with MDD, dysthymia or minor depression was recruited from both general practitioners and mental health care facilities to cover a wide range of severities. The 132 controls without a history of depression were recruited via general practitioners. After baseline assessment, participants were invited for a second face-to-face interview with repeated measures of psychological, social and physical functioning after two years.²³³ During this period, 28 people died. Of the 482 participants who were still available for the study, 401 (83.4%) took part in the follow-up assessment. All participants gave their written and oral consent. The study protocol was approved by the Ethical Review Board of the VU Medical Center in Amsterdam, and subsequently by the local ethical review boards.

The presence or absence of depression was assessed by the Composite International Diagnostic Interview (CIDI) version 2.1¹⁸⁵ at baseline and after two years of follow-up and was administered in face-to-face interviews. Depression diagnosis was present when participants met the criteria with a six months recency.

For the current study, we analyzed MDD patients (n = 359) and healthy controls (n = 132). Thirteen participants were excluded due to missing data on plasma testosterone or sex hormone binding globulin (SHBG). In total, 349 MDD-patients (119 men and 230 women) and 129 controls (50 men and 79 women) were available for the baseline analyses. Ninety-nine participants (36 men and 63 women) dropped out between the baseline and the two-year follow-up assessment (20.7% of baseline sample). Female non-responders suffered more often from MDD and male non-responders were older compared to responders. MetS rates did not differ between responders and non-responders. In total, 266 MDD patients (91 men and 175 women) and 113 control participants (42 men and 71 women) were available for longitudinal analysis.

Measurements

Predictors included testosterone, free testosterone, and SHBG. Blood was taken after an overnight fast at baseline (mean time of sampling 9:04 AM; range 8:10 to 10:45). EDTA plasma was stored at -80°C until determination of the androgenic markers. Total testosterone was measured using a second generation testosterone assay without extraction (Architect i2000 analyzer; Abbott Diagnostics, Abbott Park, IL). In contrast to earlier radioimmunoassays, second generation testosterone assays suffer less from cross-reactivity with other steroid hormones and are less influenced by SHBG levels.^{234, 235} Strong correlation coefficients are reported between our radioimmunoassay and the gold standard liquid chromatography–tandem mass spectrometry (LC-MS/MS) measurement in women ($r = 0.95$).²³⁵ Intra- and inter-assay variations were between 4% and 10% at concentrations of 0.02, 2, and 20 nmol/L. The lower limit of quantification was 0.45 nmol/L. In 23 of 318 women (7.2%), the total testosterone level was below the threshold and therefore imputed as 0.44 nmol/L as done before.¹⁸² SHBG was measured using an automated immunometric assay (Architect i2000 analyzer). Intra- and inter-assay variations were between 4% and 6% over the whole range. The lower limit of detection for SHBG was 2 nmol/L. Free testosterone was calculated from total testosterone, SHBG and albumin levels using the formula published by Vermeulen.¹⁴¹

Outcome measures were MetS (as a dichotomous variable) and its individual components (as continuous variables) at baseline and after two years of follow-up. MetS was defined according to the adjusted National Cholesterol Education Program Adult Treatment Panel III criteria.²²⁴ To fulfil the diagnosis of MetS, at least three of the following criteria had to be met: blood pressure >130 mmHg systolic or >85 mmHg diastolic or antihypertensive medication; waist circumference >102 cm for men and >88 cm for women; fasting glucose levels >5.6 mmol/L (100 mg/dL) or anti-diabetic medication; high-density-lipoprotein (HDL) cholesterol levels <1.0 mmol/L (40 mg/dL) for men and <1.3 mmol/L (50 mg/dL) for women or medication for reduced HDL cholesterol; and fasting triglycerides levels >1.7 mmol/L (150 mg/dL) or medication for hypertriglyceridemia. Waist circumference was determined by measuring tape, measuring between the lower rib margin and the iliac crest following normal expiration, over light clothing. HDL cholesterol, triglycerides and glucose levels were measured by standard laboratory methods after overnight fasting. Nineteen patients were set as missing value because they did not fast before the assessment. Blood pressure was measured in a supine position on

the right arm by an electronic Omron sphygmomanometer and was averaged over two readings. HDL cholesterol levels, triglyceride levels, and fasting glucose levels were adjusted for medication effects.²³⁶⁻²³⁸ When participants were using antihypertensive medication 10mmHg was added to the systolic blood pressure and 5mmHg was added to the diastolic blood pressure.²³⁹ When participants were using fibrates, 0.10 mmol/L [3.8 mg/dL] was subtracted from HDL cholesterol and 0.67 mmol/L [60 mg/dL] was added to triglycerides.^{236, 237} When participants were using anti-diabetic medication, when the glucose level was below 7 mmol/L [126 mg/dl] a value of 7mmol/L was given.²³⁸ Body mass index (BMI) is not a MetS component, however, is often used as a weight measurement. Therefore, we included this variable as an outcome measure as well, to see whether both assessments are comparable. In general, waist circumference and BMI are highly correlated, independent of sex ($r = 0.93$, range 0.87–0.95).²⁴⁰

Statistical analyses

All analyses were performed in strata separately for men and women. Data were analysed using IBM SPSS Statistics 24. A two-tailed p-value < 0.05 was considered statistically significant.

Characteristics were determined at baseline, using percentages, by chi-squared tests or means with standard deviation (SD) by analysis of variance, dependent on the type of variable.

The first aim was to test the association of total and free testosterone with MetS diagnosis (dichotomous) and its individual components (continuous) at baseline and after two-year follow-up. In the cross-sectional analyses, these associations were examined using logistic and linear multivariable regression analyses, when appropriate. All analyses were adjusted for age, and for education, smoking status, alcohol use, number of chronic diseases, and testosterone-affecting medication in a second adjustment. Level of education was specified in years of education. Smoking status was dichotomized into non- and current smoker. Alcohol use was defined as daily alcohol intake in units. The number of chronic diseases was counted (including hypertension, range 0 through 8). Use of medication was considered as present when using World Health Organization Anatomical Therapeutic Chemical (ATC) codes at least 50% of the days.¹⁸³ Out of the 478 participants, 33 were taking medications that could influence their testosterone level; seven participants were taking sex hormones (ATC code G03),

six participants were taking selective estrogen receptor modifiers or aromatase inhibitors (ATC code L02B) and 21 participants were taking glucocorticoids (ATC code H02AB). All users of sex hormones, selective estrogen receptor modifiers or aromatase inhibitors were female. As none of them was using (anti-) androgens and their testosterone levels were within the reference intervals, we decided not to exclude these women. However, the main analyses were repeated in a sensitivity analysis with exclusion of this group. Non-linear associations were analyzed by adding squared total and free testosterone respectively to the model. Free testosterone was logarithmically transformed to reduce their positively skewed distributions. For interpretation purposes only, total and free testosterone were standardized (calculated z-scores) before entering the regression model calculating odds ratios (ORs) for MetS. In the prospective analyses, associations between standardized total and free testosterone with the change in MetS components over two years were examined using linear regression analyses. Two-year follow-up values of MetS components were subtracted from baseline values. Therefore, a positive delta score means a lowering of MetS components over two years. For waist circumference, triglycerides, blood pressure, glucose and BMI, this would, in most cases, be associated with an improved metabolic profile. In contrast, a positive delta score for HDL would in most cases mean a worse metabolic profile. These delta scores were entered as outcomes in the regression model. Again, all analyses were adjusted for age first, and for the other covariates afterwards.

To test whether the association between testosterone levels and MetS was similar for persons with and without MDD, interaction terms (i.e., testosterone*MDD) were entered in the cross-sectional models. Next, the interaction term testosterone*antidepressant use was added to the regression models. When interactions were significant ($p < 0.05$), tests were run separately for MDD subgroups. All analyses were adjusted according to the previously mentioned models.

Results

Findings in older men

The characteristics of the 169 men are shown in *Table 5.1*. Of our sample of men, 70% was currently depressed and 46% was diagnosed with MetS. The mean plasma total and free testosterone levels were $16.7 \pm$ (standard deviation [SD]) 6.3 nmol/L and 307.1 ± 100.0 pmol/L.

Associations between total and free testosterone with MetS (components) and BMI are shown in *Table 5.2* and *Figures 5.1* and *5.2*. Overall, both lower total and free testosterone levels were associated with a higher risk for MetS (standardized OR [95% confidence interval (CI)] 0.51 [0.34-0.75] and 0.67 [0.47-0.95]). Lower testosterone levels were linearly associated with adverse MetS components except for blood pressure, with the strongest associations displayed for total testosterone levels. After adjustment for all covariates, a 1SD (i.e. 6.3 nmol/L) lower total testosterone level was associated with a mean 3.5 cm higher waist circumference ($p < 0.001$), a mean 0.18 mmol/L higher triglycerides ($p = 0.006$), a mean 0.04 mmol/L lower HDL ($p = 0.01$), a mean 0.23 mmol/L higher fasting glucose ($p = 0.03$), and a mean 0.96 kg/m² higher BMI ($p < 0.001$). SHBG showed ORs and standardized beta-coefficients (β s) in same direction as for total and free testosterone (data not shown).

Table 5.2. Baseline characteristics of 169 men and 309 women.

	Men n = 169	Women n = 309
Age , mean (SD)	69.6 (7.4)	70.9 (7.3)
European ancestry , no. (%)	159 (94.1)	292 (94.5)
Education , mean (SD)	11.9 (3.8)	10.5 (3.2)
Body mass index , mean (SD)	26.5 (3.7)	26.4 (4.5)
Smoking , no. (%)	33 (19.5)	72 (23.2)
Alcohol use , mean (SD)	0.9 (1.2)	0.6 (0.8)
Chronic disease , mean (SD)	2.1 (1.5)	2.5 (1.6)
Testosterone affecting medication , no. (%)	7 (4.1)	20 (6.5)
Total testosterone (nmol/L), mean (SD)	16.7 (6.3)	0.96 (0.40)
Free testosterone (pmol/L), mean (SD)	307.1 (100.0)	14.1 (7.62)
Metabolic syndrome , no. (%)	77 (45.6)	134 (43.4)
Waist circumference , mean (SD)	99.6 (11.4)	91.8 (12.8)
Triglycerides (mmol/L), mean (SD),	1.5 (0.8)	1.4 (0.7)
High-density lipoprotein cholesterol (mmol/L), mean (SD)	1.3 (0.3)	1.7 (0.4)
Systolic blood pressure (mmHg), mean (SD)	155.8 (20.8)	151.3 (22.4)
Diastolic blood pressure (mmHg), mean (SD)	86.3 (10.9)	82.8 (11.3)
Fasting glucose (mmol/L), mean (SD)	6.1 (1.3)	5.8 (1.3)
Major depressive disorder , no. (%)	119 (70.4)	230 (74.4)
Antidepressant users , no. (%)	84 (49.7)	170 (55.0)

Abbreviations: SD, standard deviation. To convert the values of triglycerides to milligrams per deciliter, divide by 0.01129. To convert the values for HDL cholesterol to milligrams per deciliter, divide by 0.02586. To convert the values of glucose to milligrams per deciliter, divide by 0.05551

Figure 5.1. Relationship between plasma total and free testosterone and metabolic syndrome in 169 men and 309 women.

Abbreviation: MDD, major depressive disorder. Adjusted odds ratios for MetS are shown for all participants and separately for depressed and controls on a logarithmic scale. Interaction terms for free testosterone and total testosterone with MDD are shown as well. Total testosterone and free testosterone levels were standardized before analysis. The sizes of the diamonds are proportional to the number of participants. Error bars represent standard errors.

Associations between total and free testosterone with two-year changes in MetS components and BMI are shown in *Table 5.3*. In contrast to the cross-sectional analyses, neither baseline total nor free testosterone levels were associated with changes in MetS components over two years, except for fasting glucose. A higher free testosterone level was associated with lower fasting glucose levels after two years ($p = 0.02$).

Although associations were slightly stronger in the MDD group than in controls (*Figure 5.1*), they were not statistically significantly different. For the MetS components there was only one exception. The adverse effects of lower total testosterone levels on systolic blood pressure were stronger in depressed patients compared to controls (β for interaction -0.46 ($p = 0.04$), for MDD patients -0.20 ($p = 0.02$) and for controls 0.14 ($p = 0.43$).

Findings in older women

The characteristics of the 309 women are shown in *Table 5.1*. Of our female sample, 72% was currently depressed and 43% was diagnosed with MetS. The

mean plasma total and free testosterone levels were 0.96 ± 0.40 nmol/L and 14.1 ± 7.62 pmol/L (mean $\pm$ SD).

Association between total and free testosterone with MetS (components) and BMI are shown in *Table 5.2*, and *Figure 5.1* and *5.2*. Overall, higher free testosterone levels were associated with a higher risk for MetS (standardized OR [95%CI] 1.41 [1.08-1.82]; $p = 0.01$). After adjustment for all covariates, higher free but not total testosterone levels were linearly associated with MetS as well as adverse MetS components, except for systolic and diastolic blood pressure. In contrast to men, a 1-SD (7.62 pmol/L) higher free testosterone level was associated with a mean 2.4 cm higher waist circumference ($p < 0.001$), a mean 0.18 mmol/L higher triglycerides ($p < 0.001$), a mean 0.05 mmol/L higher HDL ($p = 0.02$), a mean 0.18 mmol/L higher fasting glucose ($p = 0.02$) and a mean 1.0 kg/m² higher BMI ($p < 0.001$). SHBG showed β s in opposite direction as for free testosterone. The exclusion of 13 women using aromatase inhibitors, selective estrogen receptor modifiers, estrogens or progestogens did not remarkably change the results (standardized OR [95%CI] 1.40 [1.07-1.82], $p = 0.01$).

Associations between total and free testosterone with a two-year change in MetS components and BMI are shown in *Table 5.3*. As in men, neither total nor free testosterone were associated with changes in MetS components over two years.

Similar to older men, associations were slightly stronger in the MDD group than in controls (*Figure 5.1*), though not statistically significantly different. Only women with MDD and higher levels of free testosterone showed a higher mean waist circumference and BMI compared to controls (β for interaction 1.13 [$p = 0.03$] and 1.05 [$p = 0.045$]), for MDD participants 0.24 ($p < 0.001$) and 0.29 ($p < 0.001$) and for controls -0.04 ($p = 0.77$) and 0.02 ($p = 0.88$).

Table 5.2. Associations of total and free testosterone levels with metabolic syndrome diagnosis and components in 169 men and 309 women.

	MetS diagnosis		Waist		Triglycerides		HDL		Systolic BP		Diastolic BP		Glucose		BMI	
	OR (95%CI)	p	β	p	β	p	β	p	β	p	β	p	β	p	β	p
<i>Men</i>																
Total testosterone																
Minimally adjusted ^a	0.52 (0.36-0.76)	<.001	-0.34	<.001	-0.21	.005	0.23	.002	-0.13	.08	-0.09	.22	-0.18	.02	-0.29	<.001
Fully adjusted ^b	0.51 (0.34-0.75)	<.001	-0.33	<.001	-0.22	.006	0.19	.01	-0.11	.13	-0.09	.22	-0.18	.03	-0.26	<.001
Free testosterone																
Minimally adjusted ^a	0.66 (0.46-0.92)	.02	-0.31	<.001	-0.15	.06	0.12	.13	-0.03	.67	-0.04	.65	-0.16	.05	-0.29	<.001
Fully adjusted ^b	0.67 (0.47-0.95)	.02	-0.30	<.001	-0.15	.06	0.09	.21	-0.02	.79	-0.04	.58	-0.14	.09	-0.27	<.001
<i>Women</i>																
Total testosterone																
Minimally adjusted ^a	0.89 (0.70-1.12)	.33	-0.03	.66	0.02	.67	0.06	.34	0.07	.23	0.08	.18	-0.02	.77	0.02	.69
Fully adjusted ^b	0.92 (0.71-1.19)	.53	-0.01	.92	0.02	.69	0.08	.19	0.08	.19	0.10	.08	-0.01	.86	0.06	.28
Free testosterone																
Minimally adjusted ^a	1.31 (1.04-1.66)	.02	0.19	.002	0.25	<.001	-0.12	.04	0.09	.11	0.09	.11	0.16	.01	0.22	<.001
Fully adjusted ^b	1.41 (1.08-1.82)	.01	0.19	.001	0.25	<.001	-0.13	.02	0.11	.07	0.10	.08	0.14	.02	0.23	<.001

Abbreviations: β , standardized beta-coefficient; BP, blood pressure; CI, confidence interval; HDL, high-density lipoprotein; MetS, metabolic syndrome; OR, odds ratio. Data are standardized OR with 95%CI or β s determined by (log) linear regression. As the independent variables were naturally log-transformed in the logistic regression analyses, a one unit change in free- or total testosterone will not be a constant effect in the odds but will vary with free- and total testosterone itself. The dependent and independent variables were therefore standardized, in order for the effect sizes of the ORs to become comparable among the different models. The standardized ORs and β s refer to how many standard deviations (SD) the dependent variable will change per SD increase in the independent variable. ^a Adjusted for age. ^b Additionally adjusted for education, smoking status, alcohol use, number of chronic diseases and testosterone affecting medication.

Table 5.3. Associations of total testosterone and free testosterone levels with delta changes in metabolic syndrome components in 133 men and 246 women over two years

	Δ Waist		Δ Triglycerides		Δ HDL		Δ Systolic BP		Δ Diastolic BP		Δ Glucose		Δ BMI	
	β	p	β	p	β	p	β	p	β	p	β	p	β	p
<i>Men</i>														
Total testosterone														
Minimally adjusted ^a	-0.09	.31	0.06	.53	0.03	.76	0.06	.53	0.00	.99	0.15	.11	-0.08	.36
Fully adjusted ^b	-0.10	.26	0.07	.47	-0.01	.90	0.05	.63	0.01	.94	0.15	.10	-0.12	.20
Free testosterone														
Minimally adjusted ^a	-0.05	.57	0.12	.21	-0.03	.74	0.07	.46	0.04	.70	0.22	.02	-0.09	.30
Fully adjusted ^b	-0.07	.48	0.13	.17	-0.06	.49	0.07	.44	0.05	.62	0.22	.02	-0.14	.14
<i>Women</i>														
Total testosterone														
Minimally adjusted ^a	-0.05	.50	0.01	.86	0.11	.14	-0.05	.43	-0.00	.97	-0.01	.84	-0.03	.61
Fully adjusted ^b	-0.05	.47	0.02	.81	0.10	.14	-0.05	.47	0.00	.99	-0.02	.81	-0.03	.70
Free testosterone														
Minimally adjusted ^a	0.01	.84	0.12	.09	0.07	.36	-0.03	.72	0.05	.61	0.00	.96	-0.06	.35
Fully adjusted ^b	-0.02	.76	0.12	.10	0.05	.48	-0.03	.68	0.04	.61	-0.01	.88	-0.05	.48

Abbreviations: β , standardized beta-coefficient; BP, blood pressure; HDL, high-density lipoprotein. Data are bs determined by linear regression. As the independent variables were naturally log-transformed in the logistic regression analyses, a one unit change in free- or total testosterone will not be a constant effect in the odds but will vary with free- and total testosterone itself. The dependent and independent variables were therefore standardized, in order for the effect sizes of the b to become comparable among the different models. The bs refer to how many standard deviations (SD) the dependent variable will change per SD increase in the independent variable. ^a Adjusted for age. ^b Additionally adjusted for education, smoking status, alcohol use, number of chronic diseases and testosterone affecting medication

Figure 5.2. Relationship between plasma free testosterone (pmol/L) and metabolic syndrome components in 169 men and 309 women.

Abbreviations: BMI, body mass index; HDL, high density lipoprotein. Free testosterone levels were categorized for visualization purposes only, as all analyses were carried out with continuous data. Free testosterone levels were log-transformed before analysis, and back-transformed geometric mean levels are presented on logarithmic scales. Error bars represent standard errors. The sizes of the boxes are proportional to the number of participants.

Discussion

In this study we aimed to investigate whether plasma testosterone is associated with MetS and whether this association is different for depressed and non-depressed older men and women. Strong associations between lower levels of total and free testosterone in men and, in contrast, higher levels of free testosterone in women with MetS and its components were found. However, the association in depressed older people did not differ significantly from controls.

Our findings in men of an inverse association between testosterone and MetS are in line with earlier findings in men aged older than 60^{226, 241, 242} as were the relationships with MetS components in one study,²⁴¹ but not in the other study.²²⁶ In contrast, relationships of largely the opposite direction were found in our sample of older women. Higher levels of free testosterone were associated with MetS and its components, which is also in line with previous studies of postmenopausal women.^{97, 243-245} Similarly, in younger women with polycystic ovarian syndrome, high free and total testosterone levels have also been linked to MetS,⁹⁹ whereas such data are largely lacking in postmenopausal women with polycystic ovarian syndrome.⁹⁹ Caution should be taken when interpreting findings with calculated free testosterone in epidemiological research, as the formula is dependent on age.²⁴⁶

SHBG may explain why women, in contrast to men, are particularly sensitive to free and not to total testosterone levels. In our study, the associations for SHBG with MetS were in the same direction for both sexes, which is remarkable given the contrasting results for free and total testosterone. SHBG binds to androgens and determines for a large part the amount of free testosterone. However, it also has metabolic effects, and helps to protect women against exposure to excessively high testosterone levels. The occupancy of SHBG steroid-binding sites by androgens is sex-dependent with an occupation rate of only 18% in women and up to 56% in men. This possibly leads to a more swift response to increased or decreased androgen levels in women,²⁴⁷ which might explain our strong findings with free but not with total testosterone.

In line with other studies, we found no longitudinal associations between total and free testosterone and the incidence of MetS in older people.²⁴⁸⁻²⁵⁰ Any changes in (components of) the MetS may have already taken place years before our baseline assessment, as a high or low androgen status may have been present since puberty - exerting its metabolic effects over decades. Moreover, a single

measurement of testosterone could have been insufficient to detect associations. However, as suggested before,²⁵⁰ the case of reverse causation might be true as well supposing between total and free testosterone can be considered as a marker of MetS, but any potential causal role should be investigated during longer follow-up periods and/or in experimental designs. In supplementation studies, testosterone may help to improve lipid profiles, body composition, and insulin sensitivity in older and hypogonadal men, whereas anti-androgenic treatment had similar effects in women with polycystic ovarian syndrome.⁹⁴ Unfortunately, experimental studies in older women are scarce.²⁵¹

Despite being the subject of many studies, the underlying mechanisms of the relationship between MDD and MetS remain largely unclear. Several mechanisms, like unhealthy lifestyle habits and immune-inflammatory dysregulation, may be involved, which all influence androgen levels.^{94, 99} We could not confirm, however, that the relationship between androgenic markers and MetS is moderated by MDD or antidepressant use in older people, since associations were similar in depressed and non-depressed older people. Given the complexity of the relationship between androgenic markers and MetS and the heterogeneity of MDD, it is possible that the interaction terms that were tested were too crude to be informative. Regarding antidepressants, effects on immune-inflammatory dysregulation but also obesity might be medication-type specific.²⁵² Moreover, certain subtypes of MDD might be more vulnerable for testosterone alterations. In addition, the HPA-axis and HPG-axis are closely linked, as the end-products steroid hormones of both systems have the same backbone, being derived from cholesterol. Furthermore, these two systems crosstalk, with stimulating and inhibiting effects from both sides.²⁵³

Some limitations of our study need to be discussed. Total testosterone was measured at one time point, which limits its merit to reflect the androgenic status of the participants. However, total testosterone gradually declines with age and may be too small to be measurable and of clinical importance over a period of two years for both men and women.^{251, 254} Previous studies generally also measured the androgenic status at one time point. In addition, a radioimmunoassay was used to determine testosterone levels, which is less reliable than LC-MS/MS analyses, in particular at the lower concentrations generally measured in women (<5 nmol/L).^{163, 164} However, as stated in the method section, good correlations have been reported between this radioimmunoassay we used and the LC-MS/MS measurement of testosterone.²³⁵ A strength of our study is its design with

measurements of MetS components at two time points, which enabled us to study testosterone not only as a marker but also as a predictor for changes in MetS components. Finally, we studied the association in a well characterized sample of older subjects with standardized assessment of psychiatric disorders.

In conclusion, our study showed that lower total and free testosterone levels in older men and higher levels of free testosterone in older women were associated with MetS. Remarkably, testosterone had opposite adverse effects on the metabolic phenotype among the sexes. However, this association was not moderated by MDD or antidepressant use. Therefore, between total and free testosterone could not explain the previously reported relationship between MDD and MetS in older people.

Acknowledgements

The infrastructure for the NESDO study (<http://nesdo.amstad.nl>) is funded through the Fonds NutsOhra (project 0701-065), Stichting tot Steun VCVGZ, NARSAD The Brain and Behaviour Research Fund (grant ID 41080), and the participating universities and mental health care organizations (VU University Medical Center, Leiden University Medical Center, University Medical Center Groningen, UMC St Radboud, and GGZ inGeest, GG Net, GGZ Nijmegen, GGZ Rivierduinen, Lentis, and Parnassia). The authors have nothing to disclose. The funding source had no role in the design and conduct of the study, collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Chapter 6

Associations of plasma androgens with suicidality among men and women: a 9-year longitudinal cohort study

Anouk E de Wit
Marrit K de Boer
Fokko J Bosker
Willem AJ van der Does
Louis JG Gooren
Willem A Nolen
Brenda WJH Penninx
Robert A Schoevers
Erik J Giltay

Adapted version from the article published in Journal of Affective Disorders, May 2020.

Abstract

Aim: Testosterone has been implicated in suicidality in cross-sectional studies. Stress that coincides with a suicide attempt may alter androgen levels, so prospective studies are needed to exclude reverse causation. We aimed to examine the associations of plasma androgens with concurrent and future suicidality, and if so, whether these associations were mediated by a behavioral trait like reactive aggression.

Methods: Baseline plasma levels of total testosterone, 5 α -dihydrotestosterone, and androstenedione were determined with liquid chromatography–tandem mass spectrometry, and dehydroepiandrosterone-sulphate with a radioimmunoassay. Suicidality was assessed using the Suicidal Ideation Scale at baseline and after two-, four-, six-, and nine-year follow-up. Men and women were analyzed separately, and potential confounders were considered.

Results: Participants (n = 2861; 66.3% women) had a mean age of 42.0 years (range 18 to 65) and almost half (46.9%) fulfilled criteria for a major depressive or anxiety disorder. At baseline 13.2% of men and 11.2% of women reported current suicidal ideation. In participants who were non-suicidal at baseline, slightly more men than women reported suicidal ideation during follow-up (14.7% vs. 12.5%), whereas the reverse pattern was observed for suicide attempts (3.6% vs. 4.2%). None of the associations between androgens and current and future suicidality were significant.

Conclusion: The lack of associations between plasma levels of androgens determined by ‘gold-standard’ laboratory methods with suicidality do not support previous cross-sectional and smaller studies in adult men and women with values within the physiological range.

Introduction

Suicidal ideation and suicide attempts (together also known as suicidality), are a common mental health concern.²⁵⁵ They do not only cause a heavy emotional burden on individuals, but also on relatives as well as clinicians.^{256, 257} Approximately twice as many men as women commit suicide.²⁵⁶ Though reasons for this sex difference remain largely unknown, one could postulate that differences in the circulating levels of androgens between the sexes, are implicated in this phenomenon.²⁵⁸ They are known to alter and to be altered by the hypothalamic-pituitary-adrenal-axis (HPA-axis), and serotonergic system, systems that are likely dysregulated in suicidality.^{38, 259, 260} Furthermore, these androgens can induce direct genomic and non-genomic effects by binding to sex steroid receptors in the brain.^{36, 37}

Support for a link between androgens and suicidality has come from cross-sectional studies, mostly conducted in men. One study revealed that plasma and cerebral spinal fluid testosterone levels were higher in nine male suicide attempters measured two weeks after the attempt relative to the levels in 12 healthy controls.¹⁰⁰ Opposite results were reported in two other studies showing lower plasma testosterone levels in 15 and 80 suicide attempters compared to an about equal number of controls.^{101, 102} However, three other studies with 31 to 112 suicide attempters found no association.¹⁰³⁻¹⁰⁵ Studies in women are more scarce, but results are similarly inconsistent. A study in 51 women with bipolar disorder reported that high plasma testosterone was weakly associated with more suicide attempts, but only in an adjusted analysis that also included 16 men.¹⁰⁶ Other studies in 17 and 344 women that measured cerebral spinal fluid and/or plasma testosterone found no significant associations with suicide attempts.^{100, 103} No study showed a statistically significant link between testosterone and suicidal ideation, neither in men nor in women.^{103, 105-107}

The above-described heterogeneity in findings may be related to the acute stress that accompanies an attempt. Due to the tight coupling between the hypothalamic-pituitary-gonadal (HPG)- and the HPA-axis, testosterone levels will be altered after such an event.³⁸ Hence, we cannot conclude from cross-sectional studies whether the association suggests a mechanism for suicidality or whether it is a result of experienced stress, thus reflecting reverse causation. Two studies acknowledged this, and examined whether testosterone levels predicted suicide attempts during a follow-up of 2.5 and 21 years.^{100, 107} Higher baseline plasma testosterone levels predicted for suicide attempts in one, but not in the other

study.^{100, 107} However, the sample sizes were limited to 51 and 47 participants, and the analyses were not adjusted for previous attempts nor stratified for sex.

Reactive aggression refers to aggressive behavior evoked by a certain situation. It has been hypothesized that besides the suggested direct effect of testosterone on suicidality, reactive aggression mediates this association.²⁵⁸ This follows from studies that showed increased rates of suicidality among those with aggressive traits,^{261, 262} and from studies that showed -though more consistently in other primates than humans- associations between androgen levels and aggression.²⁶³⁻²⁶⁵ The revised Leiden Index of Depression Sensitivity (LEIDS-R) measures cognitive changes that may occur during depressed mood (for the purpose of this paper, reactive aggression will refer to aggressive cognitive reactivity).²⁶⁶ As suicidality is especially prevalent among depressed individuals,²⁶⁷ this questionnaire provides the unique opportunity to study whether a phenotype with more aggressive behavioral traits or cognitive styles mediates the association between androgens and suicidality.

The reported inconsistencies and the paucity of large and prospective studies prompted us to examine the role of androgens in suicidality. To accomplish this, androgens were determined by LC-MS/MS in a large study (n = 2861) of men and women with depressive and anxiety disorders and healthy controls who were followed over nine years. Using this data, we sought to examine the association between androgen levels and a) concurrent suicidal ideation, and b) suicidal ideation and suicide attempts during nine years of follow-up, and c) whether these potential associations are mediated by reactive aggression.

Methods

Study sample

Data were derived from the Netherlands Study on Depression and Anxiety (NESDA), an ongoing longitudinal observational cohort study that started in 2004.¹³⁵ Current and remitted depressed participants and healthy controls aged 18 to 65 years were recruited from the Dutch general population (19%), primary health care (54%) and specialized mental health care (27%). Ethics committees of participating universities approved the research design. After giving written informed consent, 2981 participants were included. Exclusion criteria were a poor comprehension of the Dutch language, and having a primary clinical diagnosis of psychotic disorder, obsessive-compulsive disorder, bipolar disorder or severe addiction disorder.

Presence of mood disorders (within the past month) were ascertained using the Composite International Diagnostic Interview (CIDI) version 2.1, administered by trained researcher assistants.¹⁸⁵ Participants were reassessed after two (wave 3 [W3]), four (W4), six (W5), and nine years (W6). All follow-up assessments were included in this study and had high retention rates (the proportion of the baseline sample participating) of 87.1% (W3), 80.6% (W4), 75.7% (W5), and 69.4% (W6).

The present study included participants of whom plasma androgens levels were available ($n = 2917$). Seventeen participants were excluded because they were using sex hormones (World Health Organization Anatomical Therapeutic Chemical [ATC] G03) or hormone antagonists (ATC L02).¹⁸³ Another 39 participants were excluded, as they were pregnant, transgender, had severely aberrant testosterone levels probably due to a measurement error, or had missing values on the outcome measure. Altogether, this resulted in a final sample of 2861 participants at baseline (96.0% of the total sample). Excluded participants ($n = 120$) drank less alcohol (percentage that drinks more than one unit a day 19.8% vs. 12.5%, $p = 0.046$), and had more often attempted suicide than included participants (18.8% vs. 15.8%, $p = 0.04$; see *Supplement 16*).

Androgens

Blood was drawn in the morning (mean 08:48 AM $\pm$ 0:25h) after an overnight fast (success rate 95.6%). Total testosterone, androstenedione, and 5 α -dihydrotestosterone (5 α -DHT) were determined with LC-MS/MS at the Clinical Chemistry department of the University Medical Centre Groningen. In case insufficient biomaterial was available to obtain 200 μ L (15%), samples were pooled. The lower limit of quantitation was 0.04 nmol/L for total testosterone and androstenedione, and 0.12 nmol/L for 5 α -DHT. Participants with values below the detection limit ($n = 1$ [0.03%] for total testosterone and $n = 178$ [6.2%] for 5 α -DHT) were given values of 0.01 below the lower limit of quantitation. Dehydroepiandrosterone-sulphate (DHEAS) was determined using ARCHITECT (ABBOTT, Wiesbaden, Germany) a radioimmunoassay with chemiflex assay protocol. The detection rate of DHEAS was ≤ 0.3 μ mol/L with a calibration range of 0.00-40.7 μ mol/L. The inter-assay coefficient of variation was estimated as $\leq 10\%$.

Primary outcomes: suicidal ideation and attempts

Suicidal ideation and suicidal attempts at baseline and follow-up were assessed with the Suicidal Ideation Scale.^{268, 269} Suicidal ideation during the last week was determined with the first 5 items. With these items, participants' attitudes toward living and dying were examined with questions like 'Did you want to live and how strong was this wish?' and 'Did you feel the desire to harm or poison yourself?'. Suicidal ideation was considered present (i.e., a dichotomous outcome) when any form of ideation, ranging from death wish to suicide plans or threats, was expressed. Time to suicidal ideation was determined as the number of years until a participant reported suicidal ideation using the herein mentioned questionnaire. Suicidal attempts were assessed with item 6 'Have you ever seriously attempted to commit suicide?'. Time to suicidal attempt was ascertained as the number of years until a participant reported a suicide attempt during one of the follow-up assessments. When present, the date was specified and the number of years between baseline and the suicide attempt could be calculated.

Secondary outcomes: reactive aggression

Reactive aggression to sad mood was measured with the reactive aggression subscale of the LEIDS-R.²⁶⁶ This questionnaire intends to measure cognitive changes that may occur during depressed mood. Participants filling out the LEIDS-R were instructed to recall a situation when they felt sad and then rate how well they were able to imagine this situation (i.e. well, somewhat, or not at all). The reactive aggression subscale contains 6 items like "When I feel bad, I feel more like breaking things" and "When I feel down, I lose my temper more easily". Participants rated the extent to which each statement applied to them on a 5-point Likert scale (0, not at all; 1, a bit; 2, moderately; 3, strongly; 4, very strongly). The sum score of these items ranged from 0 to 24 with a higher score indicating stronger reactive aggression. The LEIDS-R reactive aggression subscale has good psychometric properties with a Cronbach α of 0.83.²⁶⁶

Covariates

Age (in years), education (in years), married or with partner (yes/no), currently smoking (yes/no), alcohol use ($\leq$ or $>$ 7 units a week), measured body mass index (BMI), and number of treated chronic somatic disorders were considered as covariates. In women, the models were also accounted for the use of oral contraceptive pills (OCs, yes/no).

Statistical analyses

All analyses were done in strata, separately for men and women. Baseline characteristics of participants were presented by means $\pm$ standard deviation [SD], geometric mean with 95% confidence interval (CI), and frequencies with percentages, dependent on the type and distribution of the variable. Missing values on the covariate alcohol use ($n = 39$; 1.4%) and BMI ($n = 2$; 0.0007%) were imputed with the mean. Due to right skewedness, all androgen levels in women, and androstenedione and DHEAS levels in men, were log transformed. To ease comparability of effect sizes, levels of androgens were also standardized before the analyses (z-values).

In the primary analysis, we examined linear and non-linear associations of androgen levels with concurrent suicidal ideation using logistic regressions. Next, we analyzed linear and non-linear association between androgen levels and future suicidal ideation, and suicide attempts using Cox proportional hazard models. Participants with suicidal ideation ($n = 340$) or suicidal attempts ($n = 328$) at baseline were excluded. Participants who dropped out before the first follow-up assessment ($n = 221$) were also excluded for the analyses on suicidal ideation. Dropouts after 2 years of follow-up were censored. The Cox proportional hazards assumption was checked with log-minus-log curves.

In the secondary analysis, we examined the association between androgens and reactive aggression. Subsequently, we tested whether the significant associations in the primary analysis – if any – were mediated by reactive aggression, using the indirect method by Preacher and Hayes.¹⁸⁸ This method estimates the total, direct, and indirect effects of the independent variable on the dependent variable through the mediator variable.

All analyses were adjusted for the time-independent covariates as mentioned below the header ‘covariates’. Data were analyzed using IBM SPSS Statistics (IBM Corp) version 25, using two-sided tests. As multiple tests were performed, we calculated an adjusted False Discovery Rate p cut-off value to avoid the inflation of false-positive findings. A p-value of < 0.0005 was considered statistically significant.

Results

Data of 965 men and 1896 women with an age range from 18 to 65 years were included in the analyses. Both men and women were slightly overweight on average (mean BMI 26.2 ± 4.5 and 25.3 ± 5.2), and 41.8% and 37.0% were smokers, respectively. Almost half of the sample suffered from either a major depressive or

an anxiety disorder in the preceding month (46.9%), and 99 men (10.3%) and 229 women (12.1%) had a preceding suicide attempt. Men had almost 20 times higher total testosterone levels than women (geometric mean total testosterone 15.6; 95%CI, 14.6 to 16.6, vs., 0.79; 95%CI, 0.04 to 1.80; respectively). See also *Table 6.1*.

Table 6.1. Characteristics at baseline of 2861 men and women.

	Men n = 965	Women n = 1896
Age , mean (SD)	43.6 (12.9)	41.1 (13.1)
Education , mean (SD)	12.1 (3.3)	12.2 (3.3)
Married or with partner , no. (%)	675 (69.9)	1306 (68.9)
Body mass index , mean (SD)	26.2 (4.5)	25.3 (5.2)
Smoking , no. (%)	403 (41.8)	702 (37.0)
Alcohol intake >1 unit a day , no. (%)	307 (31.8)	262 (13.8)
Treated chronic diseases , geo. mean (95%CI)	0.5 (0.0-1.5)	0.4 (0.0-1.4)
Oral contraceptive use , no. (%)	-	508 (26.8)
Antidepressant use , no. (%)	238 (24.7)	498 (26.3)
MDD or anxiety disorder past month , no. (%)	453 (46.9)	889 (46.9)
Suicide ideation in preceding week , no. (%)	127 (13.1)	213 (11.2)
Suicide attempt in past , no. (%)	99 (10.3)	229 (12.1)
Reactive aggression , geo. mean (95%CI)	3.8 (2.8-4.8)	3.4 (2.4-4.4)
<i>Predictors - Androgenic features</i>		
Total testosterone (nmol/L), geo. mean (95%CI)	15.6 (14.6-16.6)	0.79 (0.04-1.80)
5α-DHT (nmol/L), geo. mean (95%CI)	1.5 (0.5-2.5)	0.36 (0.12-1.36)
Androstenedione (nmol/L), geo. mean (95%CI)	3.7 (2.7-4.7)	3.3 (2.3-4.3)
DHEAS (nmol/L), geo. mean (95%CI)	7.2 (6.2-8.2)	4.7 (3.7-5.7)

Abbreviations: 5 α -DHT, 5 α -dihydrotestosterone; CI, confidence interval; DHEAS, dehydroepiandrosterone-sulphate; MDD, major depressive disorder; SD, standard deviation. To convert nmol/L to ng/dl for total testosterone, 5 α -DHT, and androstenedione multiply by 28.84, 29.07, and 28.64, respectively. To convert pmol/L to pg/ml for free testosterone divide by 3.47. To convert μ mol/L to μ g/dl for DHEAS, multiply by 36.85.

Suicidal ideation and suicide attempts

As shown in *Table 6.2* and depicted in the *Figure 6.1*, about similar rates of men and women experienced suicidal ideation (13.2% and 11.2%, respectively) at baseline. Higher levels of androstenedione were associated with increased odds of suicidal ideation in men (crude standardized odds ratio (OR) [95%CI] 1.42 [1.17-1.72]; $p = 0.0004$), but not after consideration of covariates (1.31 [1.06-1.62]; $p = 0.01$). None of the other androgens was associated with suicidal ideation in men or women at baseline. Slightly more men than women reported suicidal ideation during follow-

up (113 / 769 [14.7%] compared to 192 / 1531 [12.5%]), whereas the reverse was observed for suicidal attempts (29 / 866 [3.3%] compared to 64 / 1667 [3.8%]). As shown in *Table 6.3*, androgen levels were neither linear nor non-linearly associated with future suicidal ideation or suicide attempts in men nor in women.

Table 6.2. Associations between baseline plasma androgen levels and concurrent suicidal ideation in men and women.

	Current suicidal ideation			
	OR (95% CI)	p	OR (95% CI)	p ^a
<i>Men</i>				
No. events (%)	127 / 965 (13.2)			
Total testosterone	1.01 (0.84-1.22)	.88	0.95 (0.76-1.17)	.60
5α-DHT	0.97 (0.81-1.18)	.78	0.92 (0.75-1.13)	.42
Androstenedione	1.42 (1.17-1.72)	<.0004	1.31 (1.06-1.62)	.01
DHEAS	1.07 (0.88-1.29)	.50	1.11 (0.87-1.41)	.39
Total testosterone	0.91 (0.74-1.11)	.33	0.81 (0.64-1.01)	.06
Total testosterone^{2,b}	1.21 (1.02-1.44)	.03	1.29 (1.08-1.54)	.01
5α-DHT	0.93 (0.75-1.15)	.49	0.86 (0.68-1.08)	.20
5α-DHT^{2,b}	1.04 (0.95-1.14)	.35	1.06 (0.96-1.16)	.23
Androstenedione	1.44 (1.17-1.78)	.001	1.34 (1.07-1.69)	.01
Androstenedione^{2,b}	0.97 (0.85-1.11)	.66	0.95 (0.83-1.09)	.45
DHEAS	1.01 (0.81-1.26)	.92	1.01 (0.81-1.26)	.92
DHEAS^{2,b}	0.88 (0.75-1.03)	.10	0.88 (0.75-1.03)	.10
<i>Women</i>				
No. events (%)	213 / 1896 (11.2)			
Total testosterone	1.15 (0.99-1.33)	.07	1.07 (0.92-1.25)	.40
5α-DHT	1.13 (0.98-1.30)	.11	1.05 (0.89-1.24)	.54
Androstenedione	1.22 (1.06-1.42)	.01	1.14 (0.95-1.36)	.16
DHEAS	1.08 (0.93-1.25)	.33	0.96 (0.81-1.15)	.67
Total testosterone	1.18 (1.02-1.35)	.02	1.10 (0.94-1.28)	.23
Total testosterone^{2,b}	1.13 (1.00-1.28)	.05	1.11 (0.98-1.26)	.10
5α-DHT	1.12 (0.98-1.28)	.09	1.04 (0.89-1.22)	.59
5α-DHT^{2,b}	1.19 (1.07-1.34)	.002	1.19 (1.06-1.34)	.003
Androstenedione	1.25 (1.08-1.44)	.002	1.17 (0.98-1.39)	.09
Androstenedione^{2,b}	1.08 (1.00-1.16)	.06	1.06 (0.97-1.14)	.20
DHEAS	1.13 (0.97-1.31)	.13	0.99 (0.82-1.20)	.93
DHEAS^{2,b}	1.06 (0.98-1.15)	.14	1.03 (0.95-1.12)	.45

Abbreviations: 5 α -DHT, 5 α -dihydrotestosterone; CI, Confidence Interval; DHEAS, dehydroepiandrosterone-sulphate; MDD, major depressive disorder; OR, odds ratio. Data are ORs with 95%CI based on logistic regression. To convert nmol/L to ng/dl for total testosterone, 5 α -DHT, and androstenedione multiply by 28.84, 29.07, and 28.64, respectively. To convert pmol/L to pg/ml for free testosterone divide by 3.47. To convert μ mol/L to μ g/dl for DHEAS, multiply by 36.85. ^a Adjusted for age, education, married or with partner, body mass index, smoking, alcohol use, and no. of treated chronic diseases, (and oral contraceptive pill use in women only). ^b Androgen level is squared to examine non-linear relationship with current suicidal ideation

Figure 6.1. Percentage of participants with suicidal ideation in the last week according to septiles with their mean levels of plasma androgens in 965 men and 1896 women.

Abbreviations: 5α-DHT, 5α-dihydrotestosterone; DHEAS, dehydroepiandrosterone-sulphate. To convert nmol/L to ng/dL for total testosterone, 5α-DHT, and androstenedione multiply by 28.84, 29.07, and 28.64, respectively. To convert pmol/L to pg/ml for free testosterone divide by 3.47. To convert μmol/L to μg/dl for DHEAS, multiply by 36.85. Vertical error bars represent standard errors of the mean, and horizontal error bars represent 95% confidence intervals of the mean of each septile. A second order regression line was fitted.

Table 6.3. Associations between baseline plasma androgen levels and future suicidality over nine years.

	Future suicidal ideation				Future suicide attempt			
	No. events (%)	HR (95%CI)	p	HR (95%CI)	p ^a	HR (95%CI)	p	HR (95%CI)
<i>Men</i>								
	113 / 769 (14.7)						29 / 866 (3.3)	
Total testosterone	.17	0.87 (0.71-1.06)	.17	0.85 (0.68-1.06)	.15	1.01 (0.69-1.46)	.97	1.02 (0.67-1.55)
5α-DHT	.16	0.87 (0.71-1.06)	.16	0.86 (0.69-1.06)	.16	0.88 (0.60-1.30)	.53	0.85 (0.55-1.30)
Androstenedione	.65	1.04 (0.87-1.26)	.65	1.07 (0.89-1.30)	.47	1.13 (0.79-1.62)	.52	1.20 (0.82-1.75)
DHEAS	.79	0.98 (0.82-1.17)	.79	1.13 (0.90-1.41)	.28	0.91 (0.64-1.29)	.60	1.08 (0.70-1.66)
Total testosterone	.13	0.85 (0.69-1.05)	.13	0.82 (0.66-1.03)	.09	1.02 (0.67-1.57)	.91	1.02 (0.64-1.62)
Total testosterone^{2,b}	.58	1.08 (0.83-1.40)	.58	1.12 (0.785-1.47)	.43	0.96 (0.55-1.66)	.87	1.00 (0.57-1.76)
5α-DHT	.05	0.82 (0.66-1.00)	.05	0.80 (0.64-0.99)	.04	0.96 (0.59-1.57)	.87	0.92 (0.55-1.55)
5α-DHT^{2,b}	.11	1.07 (0.99-1.16)	.11	1.08 (0.99-1.17)	.08	0.78 (0.50-1.21)	.27	0.78 (0.50-1.22)
Androstenedione	.55	1.06 (0.89-1.26)	.55	1.08 (0.89-1.31)	.45	1.12 (0.79-1.59)	.52	1.20 (0.82-1.75)
Androstenedione^{2,b}	.38	1.05 (0.95-1.15)	.38	1.01 (0.92-1.12)	.81	1.03 (0.84-1.26)	.77	1.00 (0.82-1.22)
DHEAS	.75	1.03 (0.86-1.25)	.75	1.23 (0.97-1.55)	.09	0.85 (0.56-1.31)	.46	0.99 (0.58-1.66)
DHEAS^{2,b}	.18	1.07 (0.97-1.18)	.18	1.08 (0.98-1.19)	.12	0.92 (0.70-1.20)	.52	0.91 (0.70-1.19)
<i>Women</i>								
	192 / 1531 (12.5)						64 / 1667 (3.8)	
Total testosterone	.27	0.92 (0.80-1.06)	.27	0.96 (0.83-1.11)	.57	1.09 (0.84-1.40)	.52	0.99 (0.75-1.29)
5α-DHT	.30	0.93 (0.80-1.07)	.30	0.99 (0.84-1.17)	.94	1.29 (1.00-1.66)	.05	1.24 (0.93-1.64)
Androstenedione	.67	0.97 (0.84-1.12)	.67	1.02 (0.87-1.20)	.82	1.14 (0.88-1.47)	.33	1.00 (0.74-1.35)
DHEAS	.005	0.83 (0.73-0.95)	.005	0.87 (0.75-1.02)	.09	0.99 (0.77-1.27)	.92	0.89 (0.66-1.20)
Total testosterone	.42	0.94 (0.81-1.09)	.42	0.96 (0.82-1.12)	.61	1.08 (0.84-1.41)	.55	0.98 (0.74-1.30)
Total testosterone^{2,b}	.51	1.04 (0.93-1.17)	.51	1.00 (0.89-1.12)	1.00	0.96 (0.69-1.34)	.82	0.94 (0.67-1.32)
5α-DHT	.31	0.93 (0.80-1.07)	.31	0.99 (0.84-1.17)	.89	1.27 (0.99-1.63)	.06	1.22 (0.92-1.62)
5α-DHT^{2,b}	1.00	1.00 (0.88-1.14)	1.00	0.97 (0.85-1.11)	.68	1.05 (0.86-1.30)	.62	1.04 (0.85-1.27)
Androstenedione	.69	1.03 (0.89-1.19)	.69	1.07 (0.90-1.28)	.45	1.17 (0.93-1.49)	.18	1.04 (0.77-1.39)
Androstenedione^{2,b}	.02	1.07 (1.01-1.14)	.02	1.04 (0.98-1.11)	.21	1.10 (0.98-1.23)	.11	1.08 (0.95-1.22)
DHEAS	.10	0.88 (0.75-1.03)	.10	0.89 (0.74-1.08)	.25	1.04 (0.80-1.35)	.76	0.94 (0.68-1.28)
DHEAS^{2,b}	.28	1.04 (0.97-1.12)	.28	1.01 (0.94-1.09)	.72	1.07 (0.93-1.22)	.35	1.06 (0.92-1.22)

Abbreviations: 5 α -DHT, 5 α -dihydrotestosterone; CI, Confidence Interval; DHEAS, dehydroepiandrosterone-sulphate; HR, hazard ratio. Data are HR with 95%CI based on cox regression analysis. To convert nmol/L to ng/dl for total testosterone, 5 α -DHT, and androstenedione multiply by 28.84, 29.07, and 28.64, respectively. To convert pmol/L to pg/ml for free testosterone divide by 3.47. To convert μ mol/L to μ g/dl for DHEAS, multiply by 36.85. ^a Adjusted for age, education, married or with partner, BMI, smoking, alcohol use, and no. of treated chronic diseases, (and oral contraceptive use pill in women only). ^b Androgen level is squared to examine non-linear relationship with suicidality.

Reactive aggression

We examined the association of androgens with reactive aggression. Only higher levels of androstenedione in women were associated more reactive aggression (crude standardized beta coefficient [β] [95%CI] 0.362 [0.166-0.559]; $p = 0.0003$), but not after the adjustment for covariates (0.012 [-0.219-0.243]; $p = 0.92$). See *Supplement 17*. As none of the associations in the primary analyses was significant, no mediation analyses with reactive aggression were performed.

Discussion

In this large study among 965 men and 1896 women whose plasma androgen levels were measured with the best available laboratory methods, we could not demonstrate significant associations between plasma levels of androgens with current as well as suicidal ideation or suicide attempts during follow-up. Therefore, our results challenge findings from previous studies that plasma androgen levels might be a viable biomarker of suicidality.

Our analyses focused on associations of plasma androgen levels at baseline with future suicidal ideation and suicide attempts during follow-up, as our study was founded on the notion that acute stress that is experienced by the suicide attempter may increase androgen levels whereas chronic stress may decrease androgen levels. Hence, this might explain why we cannot confirm the either positive or negative associations between androgens and suicidality that were demonstrated in previous smaller studies that examined androgen levels directly after a suicide attempt.^{100-106, 270} These studies also rarely considered confounders in their analyses and used less reliable methods than LC-MS/MS to determine androgen levels. As a consequence, this might have led to overestimation of true effect sizes.^{100-106, 270} However, an alternative explanation for the discrepancies between our results and that of others might be lying in differences in study population regarding psychiatric diagnoses. We examined participants with depression and anxiety, and healthy controls, while previous significant positive associations were derived from patients with schizophrenia,^{101, 102} and significant negative associations from patients with bipolar disorder.^{106, 107} Nevertheless, since this study included the largest sample of men and women so far, the absence of significant associations suggests that these are –if present at all– of weak strength and likely not clinically meaningful in a population of depressed and anxious patients. Finally, the positive findings in smaller previous studies may

be false positive findings or the effect of publication bias. Studying testosterone during windows of development in which it may also cause permanent effects on cerebral functioning like antenatal or during puberty, may be an intriguing alternative direction.²⁵⁸ Indirect measures of intra-uterine androgen exposure like second-to-fourth finger length, birth weight and maternal androgen levels, and thus masculinization of the brain, have been associated with increased risk of suicide completion during adulthood.²⁵⁸ Future studies could also look into androgen receptor CAG trinucleotide repeats rather than plasma levels, as those better reflect androgen sensitivity.²⁷¹

We also observed no association between levels of androgens and reactive aggression after the consideration of important covariates. Higher levels of aggressiveness have mostly been reported as a side effect of anabolic androgenic steroid abuse in men and women,²⁷²⁻²⁷⁴ but those concerned supraphysiological 5 α -DHT levels and might therefore not be comparable to the physiological levels in our cohort of men and natural cycling women. Otherwise, a recent randomized controlled trial revealed that testosterone could rapidly potentiate aggressive behavior, but only among men with dominant or impulsive personality styles.²⁷⁵ Also, men who received testosterone were more likely to pay money to punish proposers who made unfair offers as well as to reward proposers who made fair offers in the ultimatum game.²⁷⁶ Apparently, testosterone may evoke both prosocial and antisocial (e.g. aggressive) aspects depending on the context. Yet, these time- and situation depending differences were not considered in this study which might explain the lack of associations.

Strengths of this study include the examination of the largest sample of men and women up to date who were followed over prolonged periods of time with rigorously assessment of their phenotype with well-validated questionnaires.^{266, 269} Previous studies were smaller, and included up to 490 participants in cross-sectional analyses¹⁰³ and 51 participants in prospective analyses.¹⁰⁷ Also, we present the first study that reliably measured testosterone in women as androgens were determined by the 'golden-standard' LC-MS/MS method. Immunoassays were less reliable when measuring the low testosterone (<5.0 nmol/L) levels that are typically present in women due to cross reactivity with other hormones.⁸⁹ Finally, we were able to adjust for a wide variety of confounders that are important in the potential association between testosterone and suicidality. Nevertheless, this study has some limitations. Suicidality was assessed with a 5-item scale and may wax and wane over time. Therefore, it may have been undetected due to

the relatively large time intervals between follow-up assessments in our study. However, if strong associations would exist, those would resist the long timespan that were present in the longitudinal analyses. Also, single measurements may have been insufficient to predict suicidality over longer periods of time. Yet, levels of testosterone, 5 α -DHT and androstenedione tend to only slightly decrease over time in women, and be even stable after menopause.²⁷⁷⁻²⁸⁰ This decrease is most pronounced in women in their third decade,²⁷⁷ but this age category was only presented by 23% of our cohort. In middle aged men, a slow aging-related decline in total testosterone levels of 0.3-1.8% annually, is seen.²⁸¹ Moreover, 5 α -DHT levels tend to be stable in men until the age of 65, like the age range that was included this study.²⁸² Likewise, reported intraclass correlation coefficients for testosterone, 5 α -DHT and DHEAS are good, ranging from 0.74 to 0.97, depending on the time between measurements (4 weeks to 5 years) and method used (radioimmunoassay or LC/MS-MS).²⁸³⁻²⁸⁵ Altogether, this suggests that a single measure may be reliable for ranking men and women according to their androgen level for epidemiologic research. Furthermore, we were not able to account for recent perceived stress levels, which may be interesting since the HPA-axis and HPG-axis are known to be coupled.³⁸ Finally, we were not able to study completed attempts. Hence, our results are limited to the notion of the role of androgens in suicidal ideation and non-lethal attempts.

In conclusion, this study, that followed a large sample of participants over nine years, showed the absence of significant associations between plasma androgens and suicidality. Thus, our results do not support findings from previous cross-sectional and smaller studies on the role of plasma androgens in suicidality in adult men and women with values within the physiological range.

Acknowledgements

The infrastructure for the NESDA study (www.nesda.nl) is funded through the Geestkracht program of the Netherlands Organisation for Health Research and Development (ZonMw, grant number 10-000-1002) and financial contributions by participating universities and mental health care organizations (Amsterdam University Medical Centers (location VUmc), GGZ inGeest, Leiden University Medical Center, Leiden University, GGZ Rivierduinen, University Medical Center Groningen, University of Groningen, Lentis, GGZ Friesland, GGZ Drenthe, Rob Giel Onderzoekscentrum). The authors have nothing to disclose. The funding source

had no role in the design and conduct of the study, collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Chapter 7

Impact of estradiol variability and progesterone on mood in perimenopausal women with depressive symptoms

Hadine Joffe
Anouk E de Wit
Jamie Coborn
Sybil Crawford
Marlene Freeman
Aleta Wiley
Geena Athappilly
Semmie Kim
Kathryn A Sullivan
Lee S Cohen
Janet E Hall

Adapted version from the article published in Journal of Clinical Endocrinology & Metabolism, March 2020.

Abstract

Aim: Women are at increased risk for depressive symptoms during the menopause transition. Changes in estradiol secretion and presence of vasomotor symptoms (VMS) contribute to perimenopausal depressive symptoms, but links with progesterone have not been investigated. Here we aim to determine whether estradiol variability, ovulatory levels of progesterone, and VMS burden are independently associated with perimenopausal depressive symptomatology.

Methods: Fifty unmedicated perimenopausal women with mild-to-moderate depressive symptoms (measured with the Montgomery-Åsberg Depression Rating Scale) were included in an eight-week observational study. Depressive symptoms, serum levels of estradiol and progesterone, and VMS frequency were assessed weekly. Association of mood with estradiol variability, ovulatory levels of progesterone, and VMS frequency were estimated using generalized estimating equation models.

Results: During the study, 90.0% of participants had varying estradiol levels, 51.1% had ovulatory progesterone levels, and 89.8% had VMS. Greater estradiol variability and absence of progesterone levels consistent with ovulation, but not VMS frequency, are associated with higher levels of depressive symptoms (beta-coefficient [95% confidence interval] 0.11 [0.04 to 0.18]; $p = 0.001$, and $-2.62 [-4.52 \text{ to } -0.71]$; $p = 0.007$, respectively), after accounting for higher body mass index, lifetime history of major depressive disorder, and stressful life events.

Conclusion: Increasing dysregulation of ovarian hormones, but not VMS, associates with more depressive symptom burden during perimenopause. These results suggest that perimenopausal mood instability is driven by the underlying hormonal dysregulation of the menopause transition involving changes in both estradiol and progesterone.

Introduction

Women are at increased risk for mood disturbance across the perimenopause, with 45% to 68% developing depressive symptoms during this reproductive transition.¹¹⁰ While mood disturbances may be severe and constitute a full-blown major depressive episode, subsyndromal depressive symptoms are more common and remain important as they cause significant distress, impair quality of life, and increase health care utilization.^{286, 287} Identifying factors that drive vulnerability for depressive symptoms in the perimenopause helps women and their clinicians understand the potential course of depressive symptoms across the menopause transition and optimize therapeutic approaches.

Menopause-specific risk factors for the development of perimenopausal depressive symptoms include greater variability in serum levels of estradiol, higher and more variable levels of follicle stimulation hormone (FSH) and luteinizing hormone (LH), as well as the presence of vasomotor symptoms (VMS).^{116, 118, 288-292} These associations highlight the potential interactions of changing gonadal steroids dynamics with neural mechanisms regulating affect²⁹³⁻²⁹⁵ and the impact of VMS, particularly nocturnal VMS and associated sleep disturbance,²⁹⁰ in the emergence of depressive symptoms in the perimenopause. Gonadal steroid changes during the menopause transition involve wide excursions in estradiol and less frequent production of progesterone with the emergence of anovulation. However, the potential independent impact of progesterone on mood, possibly mediated through its neurosteroid metabolite allopregnanolone, has received little attention.²⁹⁵

In the current study, we investigated whether depressive symptom severity in perimenopausal women is linked independently with variability of estradiol, production of postovulatory levels of progesterone, as well as VMS frequency, after accounting for a priori potential confounders such as depression history and stressful life events. To achieve this goal, we conducted an eight-week clinical observational study of 50 perimenopausal women with untreated depressive symptoms who completed weekly assessments of mood, serum gonadal steroid levels, and VMS. We hypothesized that estradiol variability, progesterone levels consistent with ovulation, and VMS frequency would independently predict higher levels of depressive symptoms in perimenopausal women.

Methods

Women with mild-to-moderate depressive symptoms who were not using psychotropic or systemic hormonal medications were enrolled in an eight-week observational study. Throughout the study, participants completed assessments of depressive symptom severity, weekly gonadal steroid and gonadotropin assays, and a daily VMS diary. All subjects provided written informed consent for study procedures, which were approved by the Partners HealthCare Institutional Review Board and conducted at Massachusetts General Hospital and Brigham and Women's Hospital.

Participants

Fifty-eight women from the ages of 35 to 56 years who were perimenopausal, defined by the Stages of Reproductive Aging Workshop criteria as being in the early menopause transition (menstrual cycle length varying > 7 days in consecutive cycles) or late menopause transition (> 60 days without menses but amenorrhea < 12 months in the prior 12 months), participated in the study.²⁹⁶ Participants were recruited from the community through local advertising and enrolled if they met eligibility criteria during the period of study accrual. They were required to have current mild-to-moderate depressive symptoms, defined as a score from 10 to 25 on the clinician-rated Montgomery-Åsberg Depression Rating Scale (MADRS),²⁹⁷ and/or a score > 7.7 on the self-rated Kellner Symptom Questionnaire Anger-Hostility subscale.²⁹⁸ Exclusion criteria included menstrual irregularities attributable to other conditions (pregnancy, lactation, hyperprolactinemia, thyroid dysfunction, polycystic ovarian syndrome, hypothalamic amenorrhea), suicidal ideation, a MADRS score > 25 indicating severe depression, history of severe psychiatric illness (bipolar disorder, psychosis, psychiatric hospitalization or suicide attempts within the past five years), substance use disorder in the past year, sleep disorder diagnosis, and use of antidepressants, systemic hormones, or other centrally active medications known to alter mood or VMS. Sleep was monitored at home once using the ApneaLink sleep monitor (ResMed Corp., Poway, CA) to exclude women with obstructive sleep apnea.

Study overview

After obtaining written informed consent and confirming eligibility, including assessment of depressive symptom severity on the MADRS, blood was drawn to

measure serum estradiol, progesterone, and FSH at baseline. Participants also completed a daily VMS diary to determine a weekly average of VMS frequency. These assessments were repeated thereafter at the weekly visits and during the week between visits for the remainder of the eight-week study duration. At baseline, current major and minor depressive disorder was assessed using the Patient Health Questionnaire-9²⁹⁹ and lifetime history of major depressive disorder (MDD) was assessed by study psychiatrists using a clinical interview. In addition, recent stressful life events were self-reported using Life Experience Survey.³⁰⁰

Severity of depressive symptoms

Severity of depressive symptoms was measured using the clinician-rated 10-item MADRS (range 0–60) as described previously.²⁹⁷ The MADRS has good internal consistency (Cronbach's $\alpha = 0.82$), test-retest reliability (correlation = 0.90), and validity, and is highly sensitive to changing symptom levels.^{301, 302} Scores of 7–19 and 20–30 (out of 60) are commonly used for identifying mild and moderate depression, respectively.³⁰³

Hormonal assays

Serum levels of estradiol were measured by liquid chromatography–tandem mass spectrometry (LC-MS/MS) (Mayo Clinic, Rochester, NY), with 10 pg/ml as the lower limit of quantitation.^{304, 305} The inter-assay coefficient of variation for estradiol in this range is 8.6%.³⁰⁴ Serum progesterone was measured by chemiluminescence immunoassays (Abbott Architect ci8200 and Beckman Coulter, Fullerton, CA). The lower analytical sensitivities of these assays were 0.1 ng/ml and 0.08 ng/ml, respectively and both inter-assay coefficient of variation were $\leq 10\%$. Serum FSH was measured by a chemiluminescence immunoassay (Abbott Architect ci8200). The analytical sensitivity was 0.05 IU/L and the inter-assay coefficient of variation was estimated to be $\leq 10\%$.

Vasomotor symptoms

Subjective daily VMS were assessed using a self-report vasomotor symptoms diary, as previously described.³⁰⁶ VMS frequency was then categorized as daytime or nighttime events. These data were then averaged separately for the week between visits to obtain weekly averages of daytime, nighttime, and total VMS frequency, consistent with established procedures.³⁰⁶

Statistical analyses

This analysis was restricted to the 50 (86.2%) of 58 eligible subjects who completed a minimum of 4 consecutive weekly visits and provided complete information on the key predictors and outcome. The enrolled sample of 50 did not differ from the eight ineligible women on key demographic and clinical characteristics. As the data involved multiple within-person observations, we used generalized estimating equation (GEE) models to examine the association of depressive symptom severity with estradiol variability, progesterone peaks, and VMS frequency, accounting for within-person correlation.³⁰⁷ Visits with estradiol levels below the detection limit ($n = 83$) were given values of 9 pg/ml. Estradiol was log transformed (ln-estradiol) due to its right-skewed distribution. The coefficient of variation of ln estradiol (hereafter referred to as variability in estradiol) was calculated for each subject to integrate the extent to which serum estradiol levels varied within each subject over the eight-week period. A progesterone level > 6 ng/mL in weekly samples was used to define presumed ovulation or its absence in the analysis.³⁰⁸ VMS were analyzed as average daily frequency over the previous week. Therefore, analyses linked the time-invariant predictors variability in estradiol and presence of ovulation and the time-variant predictor VMS with weekly depressive symptom severity scores describing the same week.

We used GEE models to determine the association of severity of depressive symptoms on the MADRS with each of the 3 main predictors (variability in estradiol, occurrence of ovulatory progesterone levels, VMS frequency) and then each of the following a priori covariates: age, Caucasian (yes/no), body mass index (BMI), marital status, college graduate, lifetime history of MDD, and number of stressful life events. For each of the univariate models for which the predictor variable was statistically significant, multivariate models were developed by adjusting first for the other primary predictors and covariates examined that achieved the same significance threshold. Sensitivity analyses divided VMS according to daytime and nighttime frequency. In order to examine whether the associations were stronger for women who had experienced recent stressful life events or who had suffered from a MDD previously, the interaction terms stressful life events*estradiol variability and stressful life events*presence of presumed ovulation, and lifetime MDD*estradiol variability and lifetime MDD*presence of presumed ovulation were added to the multivariate model.

All analyses were adjusted for the number of days between visits (average one week). Data were analyzed using STATA 15 (College Station, Texas), using a p-value of 0.05 as the threshold for statistical significance.

Results

Participant characteristics at baseline

Baseline characteristics of the 50 subjects in this analysis are reported in *Table 7.1*. Briefly, the mean $\pm$ standard deviation (SD) age of study participants was 48.4 ± 3.9 years, mean BMI was 27.2 ± 6.7 kg/m², and the majority were either Caucasian (56%) or African American (36%). Menopause status was evenly divided between those in the early and late menopause transition. The median FSH and estradiol at baseline were 24.8 (interquartile range [IQR], 6.9–54.4) IU/L and 83.0 (IQR, 41.0–148.0) pg/ml, respectively. VMS were reported by 87.5% of participants. The mean baseline MADRS score was 15.5 ± 5.3 , reflecting mild to-moderate depressive symptom burden. Depressive symptoms were subthreshold for a current major depressive episode in the vast majority of participants (87.3%). One-third of participants reported previously

Table 7.1. Baseline characteristics of 50 perimenopausal women with mild-to-moderate depressive symptoms.

	Baseline	Follow up + baseline
Age, mean (SD)	48.4 (3.9)	-
Caucasian, no. (%)	28 (56.0)	
Married, no. (%)	25 (50.0)	-
College graduate, no. (%)	28 (56.0)	-
Body mass index, mean (SD)	27.2 (6.7)	-
FSH (IU/L), median (IQR)	24.8 (6.9–54.4)	22.4 (6.9–52.5)
Estradiol (pg/ml), median (IQR)	83.0 (41.0–148.0)	77.0 (30.0–144.0)
Presence of ovulatory progesterone, no. (%)	-	23 (51.1)
Vasomotor symptoms, no. (%)	42 (87.5)	44 (89.8)
Vasomotor symptom frequency		
24 hours, median (IQR)	2.5 (1.4–4.7)	2.1 (0.9–4.6)
Night, median (IQR)	1.3 (0.7–2.1)	1.2 (0.6–2.1)
MADRS scores, mean (SD)	15.5 (5.3)	10.8 (6.5)
Participants with current MDE, no. (%)	6 (12.7)	-
Participants with lifetime MDD, no. (%)	16 (32.0)	-
Stressful life events in past 6 months, no. (%)	44 (88.0)	-

Abbreviations: FSH, follicle stimulating hormone; IQR, Interquartile range; MADRS, Montgomery-Åsberg Depression Rating Scale; MDE, major depressive episode; MDD, major depressive disorder; SD, standard deviation.

experiencing an episode of MDD. The current episode of mood symptoms developed concurrently or after onset of perimenopause in 70% of participants and concurrent with or after VMS in 56% of participants.

Predictors of depressive symptom severity during follow-up

Associations of depressive symptom severity with estradiol variability, presence of presumed ovulation, and VMS frequency are shown in *Table 7.2*. Univariate models revealed that greater variability in serum estradiol levels was associated with higher levels of depressive symptoms (beta-coefficient [β] [95% confidence interval (CI)] 0.11 [0.02-0.21]; $p = 0.02$). Univariate models also revealed that presence of at least one ovulatory progesterone level was associated with lower depressive symptom levels (β [95%CI] -2.62 [-4.52 to -0.71]; $p = 0.007$). VMS frequency was not associated with depressive symptom severity (β [95%CI] 0.05 [-0.24 to 0.34]; $p = 0.75$). Results were unchanged when VMS were analyzed separately by nighttime and daytime (data not shown). Other predictors of worse depressive symptoms in univariate models were higher BMI (β [95%CI] 0.23 [0.08-0.37]; $p = 0.002$), history of MDD (3.77 [1.97-5.57]; $p < 0.001$), and more stressful life events (0.31 [0.01-0.62]; $p = 0.04$), but not the subset of stressful life events that were severe, nor age, Caucasian, marital status, or college graduate (all $p > 0.10$).

Table 7.2. Associations of estradiol variability, ovulatory progesterone, and vasomotor symptoms with severity of depressive symptoms in 50 perimenopausal women with mild-to-moderate depressive symptoms.

	Univariate		Multivariate	
	B (95%CI)	p	B (95%CI)	p
Variability in estradiol	0.11 (0.02, 0.21)	.02	0.11 (0.04, 0.18)	.002
Ovulatory progesterone levels	-2.62 (-4.52, -0.71)	.007	-2.68 (-4.10, -1.27)	<.001
Vasomotor symptom frequency	0.05 (-0.24, 0.33)	.75		
Age	0.18 (-0.08, 0.43)	.17		
Caucasian	0.25 (-1.74, 2.24)	.81		
Married	-0.46 (-2.44, 1.52)	.65		
College graduate	-1.57 (-3.54, 0.40)	.12		
Body mass index	0.23 (0.08, 0.37)	.002	0.22 (0.12, 0.32)	<.001
Recent stressful life events	0.31 (0.01, 0.62)	.04	0.30 (0.08, 0.51)	.007
Lifetime MDD	3.77 (1.97, 5.57)	<.001	3.38 (1.93, 4.84)	<.001

Abbreviations: B, beta-coefficient; MDD, major depressive disorder. Estimates are B's determined by generalized estimating equation analyses. All models were adjusted for the number of days between study visits.

A final multivariate model was built adjusting for BMI, history of MDD, and stressful life events, covariates that were significantly associated with MADRS scores in univariate models. The association of the severity of depressive symptoms with the primary predictors of interest, estradiol variability and ovulatory progesterone levels, remained significant (adjusted B [95%CI] 0.11 [0.04-0.18]; $p = 0.002$, and $-2.68 [-4.10 \text{ to } -1.27]$; $p < 0.001$, respectively, *Table 7.2* and *Figure 7.1*). This translates to an increase of 1.1 on the MADRS depression score for each 10% increase in the relative variability of estradiol and a 2.7-point decrease of the MADRS depression score for those who had at least one progesterone level consistent with ovulation relative to those who did not. In the same model, all covariates (higher BMI, history of MDD, stressful life events) remained independent predictors of more severe depressive symptoms (all $p < 0.01$). *Figure 7.1* presents the marginal means of depressive symptom severity with estradiol variability and by the presence of progesterone levels consistent with ovulation, respectively. When the adjusted analysis was re-run with the number of presumed ovulatory cycles as a 3-category variable (1 or 2+ vs 0), the difference in depressive symptom severity compared with the reference group was significant for both the group that had only one ($p < 0.001$) and for those who had two or more ovulatory cycles ($p = 0.013$). However, there was no “dose response” relationship; those with multiple ovulatory cycles did not have lower levels of depressive symptoms than those with one cycle ($p = 0.36$).

The association of estradiol variability and ovulatory progesterone levels with depressive symptom severity did not vary by the presence of recent stressful life events nor by a lifetime diagnosis of MDD ($p > 0.10$ for all interaction terms).

Figure 7.1. Depressive symptom severity by estradiol variability (left) and presence ovulatory progesterone levels (right) during the eight-week study period.

Abbreviations: B, beta-coefficient. Figures show the depressive symptom severity on the Montgomery Åsberg Depression Rating Scale (range 0–60, higher score worse) according to 10 septiles (i.e., quantiles that divided an ordered sample into ten equally numerous subsets) with a fitted regression line for the left figure and according to 'ovulation' and 'no ovulation' in the right figure. The presence of ovulation during the study period was indicated by at least one distinct serum progesterone peak > 6 ng/dl. Vertical error bars represent standard errors of the mean.

Discussion

Results of this study show that depressive symptom severity during the perimenopause is linked to a greater degree of concurrent estradiol variability and the absence of progesterone at levels suggesting recent ovulation, but not with the presence of VMS. These findings in women with mild-to-moderate depressive symptoms remain after accounting for other factors known to be associated with depressive symptom severity during the perimenopause, including a prior history of MDD, the presence of recent stressful life events, and a higher BMI. Our results therefore suggest that perimenopausal mood instability is driven in part by the underlying dysregulation of both estradiol and progesterone that is characteristic of the menopause transition.

Our analysis focused uniquely on concurrent changes in hormonal profiles in women actively experiencing depressive symptoms. To achieve this, we used a repeated-measures approach to capture weekly measures of mood, serum gonadal steroid levels, and VMS over an eight-week time period. This is in contrast to epidemiologic approaches that have examined reproductive hormone levels in women as predictors of subsequent depressive symptoms and mood disorders that emerged months or years later.^{116-118, 289} Such approaches contrast with our

underlying hypothesis that depressive symptoms are caused by concurrent exposure of neural circuitry linked with mood to highly variable perimenopausal hormone dynamics. During the menopause transition, excursions in estradiol and the occurrence of ovulation vary widely both within women over time and between women, even among those with relatively comparable menstrual patterns.^{309, 310} This marked variability may explain why depressive symptoms can ebb and flow over time during the perimenopause and why some women are susceptible to mood disturbance at some, but not all, times over the duration of this transition.

Results of our study highlight the importance and independent roles of both estradiol variability and ovulatory levels of progesterone in perimenopausal mood regulation. Numerous studies have focused on serum estradiol levels in relation to menopause-related depression^{22, 117-121} but associations with serum progesterone have not been explored. Thus, we observed for the first time that lower levels of depressive symptoms occur surrounding an ovulatory cycle. It is notable that women are at increased risk for depressive symptoms during both the early menopause transition-when ovulation is more common, as well as during the late menopause transition-when prolonged amenorrhea and anovulation occur more commonly.³¹¹ Therefore, while estradiol variability is more common in the early menopause transition and anovulatory cycles are more common in the late menopause transition, both occur in either phase of the transition, as does the increased prevalence of depressive symptoms. Our population was evenly divided between early and late menopause, which gave us the opportunity to examine associations across the menopause transition and a wide range of hormone profiles. Of the 90% of women who had at least one detectable estradiol level, half had at least one progesterone level consistent with ovulation. Therefore, regardless of whether women were in the early or late menopause transition or whether estradiol levels fluctuated more or less markedly, depressive symptoms were less severe in our population across the eight-week period of observation when progesterone levels indicated recent ovulation.

Consistent with our observation of an inverse relationship between depressive symptoms and concurrent exposure to progesterone is the recent finding of an antidepressive effect of the progesterone-derived neurosteroid allopregnanolone for treatment of postpartum depression,³¹² leading to the approval of allopregnanolone by the Food and Drug Administration for postpartum depression. As a treatment for another reproductive hormone associated mood

disturbance,³¹³ allopregnanolone's efficacy for postpartum depression challenges earlier presumptions that progestins adversely affect mood. Allopregnanolone may act as a neurosteroid to mediate the protective effect of peripheral progesterone on mood through direct inhibition of glutamatergic and γ -aminobutyric acid (GABA) receptors. However, the relationship of ovarian production of progesterone to centrally active allopregnanolone and GABA inhibition remains to be determined in women with hormonally sensitive mood disturbance. Collectively, these data illustrate a beneficial effect of progesterone and its metabolites on mood in women with hormonally linked mood disturbance.

We observed higher levels of depressive symptoms in association with more marked concurrent variability in serum estradiol. Levels of estradiol are known to be more variable during the early menopause transition than during the follicular phase of premenopausal menstrual cycles.^{20, 108} Variability in estradiol levels during the perimenopause may adversely impact mood through mediation of serotonergic and dopaminergic systems within the central nervous system since estradiol receptors are widely dispersed across these neuronal networks involved in mood-regulating neurotransmission.⁴⁸ Alternatively, because estradiol increases allopregnanolone, marked variability of estradiol during perimenopause may cause fluctuations in allopregnanolone, resulting in failure of the GABA_A receptor to regulate overall GABA-ergic tone.¹⁶² Dysregulation in GABA-ergic tone may also lead to perturbation of the hypothalamic-pituitary-adrenal-axis (HPA-axis), potentially conferring vulnerability to depression.¹⁶²

Our finding that variability in estradiol is associated with depressed mood is consistent with those of some,¹¹⁶⁻¹¹⁸ but not all,¹¹⁹⁻¹²¹ studies examining associations of serum estradiol with depressive symptoms. In contrast to our study, others did not examine estradiol changes in relation to concurrent mood, but in relation to subsequent emergence of depression.¹¹⁶⁻¹¹⁹ Differences between our and other studies also include the frequency of estradiol sampling. Our findings using weekly assessments are generally consistent with others in which estradiol was measured more frequently (weekly,¹¹⁷ monthly,^{116, 314} or every 6 months¹¹⁸) rather than less (every 8-12 months¹¹⁹⁻¹²¹) frequently. Given the changes in estradiol across the menopause transition, comparison between studies examining associations between estradiol variability with adverse mood that utilize different estradiol sampling frequencies must be interpreted cautiously. Our study population also differs from previous studies that were restricted to women with clinically

relevant mood disorders^{22, 117, 119, 314} as the vast majority of participants in the current study had less severe depressive symptoms that did not constitute a MDD. This is important because factors associated with mood disturbance during the perimenopause differ between those who have MDD and those with less severe depressive symptoms.²⁹¹

In contrast with previous studies, we did not find stronger associations of estradiol variability (or postovulatory levels of progesterone) with depressive symptoms in women who recently experienced stressful life events¹¹⁸ nor in those who had a history of MDD.¹¹⁷ Results of our analysis therefore suggest that previous exposure to these independent risk factors for depression does not sensitize women to greater depressive symptom burden when estradiol varies nor when progesterone levels are low.

VMS, especially when nocturnal, are associated with risk for subthreshold depressive symptoms,^{288, 290, 311} but not with clinically relevant depressive disorders in perimenopausal women.^{288, 311, 314} We did not observe an association between VMS and depressive symptom severity. Our findings are consistent with previous studies in perimenopausal women who had variable estradiol levels,³¹⁴ but not with other studies in which hypo-estrogenism was persistent.²⁹⁰ It is possible that our sample might be too homogeneous to detect an association with VMS as almost 90% of participants reported VMS and only 5% had sustained hypoestrogenism. Alternatively, the impact of VMS on mood may only be evident in perimenopausal women during a period of hypo-estrogenism, further contributing to the complex interrelationships among gonadal steroids, VMS and mood across the perimenopause.

Our study has important strengths, particularly the weekly, concurrent assessment of depressive symptoms, gonadal steroids, and VMS in a population whose mental health and menopause profile was well-characterized. This approach informs how co-varying menopause specific factors contribute to concurrent depressive symptom burden. To date, the majority of studies examining hormonal predictors of depressive symptoms have assessed mood, reproductive hormones, and VMS as infrequently as every 8–12 months.^{119–121} While these studies powerfully characterize the course and predictors of depressive symptoms over a prolonged period of time, they do not provide critical insight into the concurrent relationship between real-time reproductive hormone dynamics and depressive symptom levels during this dynamic phase of a woman's life, as our design achieves. Another

strength is that estradiol was assayed using the gold-standard LC/MS-MS, which has improved reliability over immunoassays in the low estradiol range.³¹⁵

Limitations in our study include inherent limitations of any observational study; namely, data interpretation is confined to associations rather than causal relationships. While frequent, our assessments were weekly, reducing our ability to capture all occurring ovulatory cycles based on progesterone levels. However, this would have biased any association toward the null, thereby limiting our ability to detect an association. Because all analyses were adjusted for time between assessments, any between- and within-participant variation due to variable time intervals was accounted for in the analysis.

In summary, we observed that a more optimal concurrent mood state is present when cycles are ovulatory and have less variability in estradiol; conversely, greater depressive symptom burden is present during anovulatory periods when progesterone is low and estradiol is especially variable. Therefore, regardless of whether women are in the early or late menopause transition, the highly variable and unpredictable reproductive hormone dynamics during the perimenopause explain, at least in part, the variability in depressive symptoms even within an individual.

Acknowledgments

This work was supported by the National Institutes of Mental Health [5R01MH082922]. H Joffe's research program is supported by grants from NIH, V Foundation, Merck, Pfizer, Que-Oncology, and NeRRe/KaNDy. H Joffe is also a consultant for NeRRe/KaNDy, Merck, Sojournix, Eisai, and Jazz Pharmaceutical. Other authors have nothing to disclose. The funding source had no role in the design and conduct of the study, collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Chapter 8

Predictors of irritability symptoms in mildly depressed perimenopausal women

Anouk E de Wit
Erik J Giltay
Marrit K de Boer
Margo Nathan
Aleta Wiley
Sybil Crawford
Hadine Joffe

Adapted version from the article published in Psychoneuroendocrinology, January 2021.

Abstract

Aim: Irritability is a highly burdensome complaint, commonly, but not universally, linked with depressive symptoms. While increased variability in estradiol has been associated with depressive symptoms during perimenopause, more insight is needed into reproductive hormone dynamics and other factors that predispose perimenopausal women to irritable mood.

Methods: Among 50 mildly depressed perimenopausal women (mean [standard deviation] age 48.4 [3.9] years), severity of irritability symptoms (on Symptom Questionnaire Hostility subscale, range 0-23) was assessed weekly for eight weeks, concurrent with potential predictors. Associations between these were examined using generalized estimating equating models.

Results: Most women (82.0%) reported having moderate to severe irritability at least once. However, the severity of irritability was highly variable from week-to-week (between-subject mean coefficient of variation 72.9% and within-subject mean coefficient of variation 63.7%). In multivariate analyses, less variable serum estradiol levels (adjusted standardized beta-coefficient [95% confidence interval] within-person coefficient of variation -0.23 [-0.32, -0.14]; $p < 0.001$), greater depressive symptom severity (0.45 [0.35, 0.56]; $p < 0.001$), younger age (-0.23, [-0.28, -0.09]; $p < 0.001$), and more frequent vasomotor symptoms (0.14 [0.05, 0.23]; $p = 0.002$) were associated with more irritability. Depressive symptom severity explained the largest portion of the variance in irritability, but still not more than 20.3%. Neither crude values, weekly change in, or variability of progesterone or follicle stimulating hormone levels were associated with irritability.

Conclusion: Irritability was highly prevalent among mildly depressed perimenopausal women. In contrast to depressive symptoms, decreased rather than increased variability in estradiol levels was associated with more irritability. This highlights that irritable mood can be disentangled from depressive symptoms in perimenopausal women and might be linked with different estradiol dynamics.

Introduction

Irritability, defined as a low threshold for experiencing frustration or anger, is prevalent during the perimenopause.^{111, 112, 316} Perimenopausal women complain 41.2% more often about irritability compared to premenopausal women.¹¹² Though often coincidental with depression,³¹⁷ irritability is a distinct dysphoric mood state.³¹⁸ Its presence decreases quality of life,³¹⁹ and is associated with greater severity of depressive symptoms.³²⁰ In addition to this personal burden, irritability, more than depressive symptomatology, may provoke excessive frustration among others.³²¹ Unfortunately, identification of mechanisms that predispose perimenopausal women to irritability, has garnered little attention. Patterns of reproductive hormones are of particular interest as the perimenopause is characterized by more marked fluctuations of serum estradiol, alternating with periods of sustained lower and unchanging levels of estradiol, resulting in less predictable ovulation.^{20, 108} Moreover, previous research has shown that greater variability in estradiol levels,^{116, 117, 322} and absence of ovulation as measured with progesterone,³²² is associated with more depressive symptomatology during the perimenopause.

Three large studies have examined risk factors for irritability, defined using a single question, in midlife women transitioning from pre- to post-menopause.^{111, 112, 323} Only one of these, the Penn Ovarian Aging Study, investigated reproductive hormone patterns.¹¹¹ This study showed that mean levels of estradiol or follicle stimulating hormone (FSH) obtained every nine to twelve months over nine years were not associated with concurrent moderate to severe irritability.¹¹¹ However, measuring hormone levels approximately annually does not capture the dynamics of the perimenopause, and other hormones of interest, such as progesterone, were not examined. The two other studies that investigated menopause-related risk factors for irritability in women across the menopausal transition originated from the Study of Women's Health Across the Nation cohort. Here it was shown that more frequent vasomotor symptoms (VMS) and difficulty sleeping were associated with higher odds of severe irritability.^{112, 323}

Together, insight in menopause-related factors and specific hormone dynamics that predispose perimenopausal women to irritability is limited, but much needed. In the current study, we sought to determine which characteristics influence vulnerability to irritability as part of a dysphoric mood presentation during the perimenopause. To accomplish this, we conducted an eight-week

observational study in perimenopausal women with untreated depressive symptoms, whereby weekly assessments of reproductive hormones and other predictors were obtained concurrent with assessment of irritability and depressive symptoms using well validated questionnaires.

Methods

Participants

Data were derived from 50 perimenopausal women who were enrolled in an eight-week observational study. Perimenopausal status was defined using the Stages of Reproductive Aging Workshop criteria.²⁹⁶ Subjects were included based on mild-to-moderate depressive symptoms defined as a score 10–25 on the Montgomery-Åsberg Depression Rating Scale (MADRS),²⁹⁷ and/or moderate irritability, defined as a score >7.7 on the Hostility scale of the Symptom Questionnaire²⁹⁸ to ensure that the sample represented the rich variety of mood symptoms that manifest during the perimenopause. Details of exclusion criteria are described in our previous work reporting the association of estradiol and progesterone variability with depressive symptoms severity.³²² All participants provided written informed consent for study procedures, which were approved by the Partners HealthCare Institutional Review Board and conducted at Massachusetts General Hospital and Brigham and Women's Hospital. The dataset is comprised of 420 visits collected in 50 women in whom information on irritability occurring during the past week was available.

Every week, women rated their level of irritability for the preceding seven days. Irritability was measured using the 23-item Hostility scale of the Symptom Questionnaire.²⁹⁸ Items were rated as yes or no, and a total score was calculated by summing up the individual item scores (range 0–23). A higher score indicates more irritability, and a score >7.7 is indicative of at least moderate irritability.²⁹⁸ The Symptom Questionnaire has been validated and found to be highly sensitive in detecting irritability in populations characterized by subclinical affective symptoms.^{298, 324} Internal consistency was good in previous studies^{298, 325} and at our baseline assessment (Cronbach's $\alpha = 0.87$). To provide insight into the variability of irritability within and between women, both within-person coefficient of variation ($[\text{personal standard deviation (SD) irritability scores}_{\text{visit 1 to 9}} / \text{personal mean irritability scores}_{\text{visit 1 to 9}}] * 100\%$) and between-person coefficient of variation

($[\text{population SD irritability scores}_{\text{visit 1 to 9}} / \text{population mean irritability scores}_{\text{visit 1 to 9}}] * 100\%$) were calculated. However, absolute irritability scores were used for the prediction model as the dependent variable, as described below.

Hormonal predictors

Serum levels of estradiol were measured weekly by liquid chromatography-tandem mass spectrometry (LC-MS/MS) (Mayo Clinic, Rochester, NY), with 10 pg/ml as the lower limit of detection and an inter-assay coefficient of variation of 8.6%. Participants with values below the detection limit (14 participants with 83 datapoints [19.7% of all data points]) were imputed with the value of 9 pg/ml. As a sensitivity analyses, measures were imputed with a different method. To this, we replaced the left-censored data (observations that are not quantified, but are known to be less than the detection limit) by a random draw from a distribution computed between zero and the lower limit of detection, imputing it using a log-transformation of the distribution (range of imputation from minus infinity to $\log_{10} \text{ pg/mL}$).³²⁶ Serum progesterone was measured by chemiluminescence immunoassays (Abbott Architect ci8200 and Beckman Coulter, Fullerton, CA). The lower level of detection of the assays were 0.10 ng/ml and 0.08 ng/ml, respectively. Serum FSH was measured by a chemiluminescence immunoassay (Abbott Architect ci8200) that had a lower level of detection of 0.05 IU/l. The inter-assay coefficient of variations of the progesterone and FSH assays were $\leq 10\%$. From these values, we investigated three different hormonal predictors. The first contained the crude hormone levels. The second comprised a measure of variability of estradiol, progesterone and FSH within an individual woman. This within-person coefficient of variation was calculated by dividing the within-subject SD of hormone level_{visit 1 to 9} by the within-subject mean hormone level_{visit 1 to 9} multiplied by 100%. The third comprised a value of weekly changes in hormone levels which was calculated by subtracting the value of hormone level_{visit x-1} from the value of hormone level_{visit x}. This way, we tested whether the magnitude of change in hormone levels over the preceding week was associated with irritability levels during that same week.

Other predictors

At baseline, age, ancestry, college graduate, measured body mass index (BMI), recent stressful events, and lifetime history of major depressive disorder (MDD)

were collected. Ancestry was collected as “Caucasian” (n = 30, 60.0%), “Black” (n = 18, 36.0%), “Native American” (n = 1, 2.0%), and “Asian” (n = 1, 2.0%). For model simplicity however, ancestry was dichotomized into Caucasian versus non-Caucasian. We therefore refer to this covariate with “Caucasian”. Whether participants obtained a college degree was used to define college graduate level. Lifetime MDD history was assessed by psychiatrists using an unstructured clinical interview. Number of recent stressful life events were self-reported using Life Experience Survey.³⁰⁰ Severity of depressive symptoms was assessed weekly using the 10-item clinician-rated MADRS (ranges from 0–60, higher scores indicating higher severity). A score of 7–19 and 20–34 is commonly used to indicate mild and moderate depression, respectively.³²⁷ Subjective daily VMS were assessed consistent with our previous approach, averaging number of VMS for the preceding week based on daily diary reporting.³²² Bi-weekly, participants completed the 7-item self-rated Insomnia Severity Index questionnaire to measure the severity of insomnia over the last two weeks.³²⁸ As scores reflected the severity of insomnia of the past two weeks, scores were used during two adjacent assessments to match the frequency of the weekly assessed predictors and outcome.

Statistical analyses

Sample characteristics were presented with means $\pm$ standard deviation (SD), median with interquartile range (IQR), or by frequencies with percentages, dependent on the type and distribution of the variable. As the data involved multiple within-person observations, we used generalized estimating equation (GEE) models with robust standard errors to determine the association of potential predictors with irritability, accounting for within-woman correlation.³⁰⁷ GEE models with different covariance structures were fitted using restricted maximum likelihood estimation, with “unstructured” chosen based on the Quasi-Akaike Information Criterion. All GEE models were adjusted for the time between visits, and both independent (if not dichotomous) and dependent variables were standardized by subtracting the sample mean of the variable and dividing the result by the variable’s sample SD (z-scoring) to ease comparability of the strengths of the associations.

First, we examined which variables predicted irritability. After univariate models were determined, multivariate models were examined by adjusting for predictors that achieved statistical significance in univariate analyses. For all

these analyses, missing datapoints on estradiol ($n = 2$, 0.5%), FSH ($n = 1$, 0.2%), progesterone ($n = 5$, 1.2%), severity of depressive symptoms ($n = 1$, 0.2%), VMS ($n = 61$, 14.5%), or severity of insomnia ($n = 56$, 13.3%) were not imputed. Hence, these missing's reduced the amount of available datapoints for some of the univariate models, and the multivariate model. Finally, the explained variance of irritability by variables that were significant in the multivariate analyses was determined by calculating the R^2 of each of the variables (standardized beta-coefficient $[\beta]$ variable²*100%).

Two post-hoc analyses were conducted. We first explored whether within-person mean estradiol might confound the association between variability of estradiol and irritability because the within-person coefficient of variation is affected substantially by the mean value for each participant. To account for this possibility, we added within personal-mean estradiol to the multivariate model. In a separate analysis, we explored to what extent the direction of the association between the weekly changes in estradiol and irritability varied between women, consistent with a recent study.³²⁹ We calculated for each woman the association of both crude and absolute values of weekly changes in estradiol with irritability and examined whether the association for each women was best described by absolute changes (increase and decrease in estradiol) or crude changes (increase or decrease, depending on the direction of the β) based on which β was largest in magnitude for a given woman.

Statistics were conducted using RStudio (R version 3.6.0; R Foundation for Statistical Computing, Vienna, Austria, 2016. URL: <https://www.R-project.org/>), with main packages 'gee' (version 4.13-19), 'gmodels' (version 2.18.1), and 'ggplot2' (version 3.2.1) packages. As multiple tests were performed, we calculated an adjusted False Discovery Rate p-value cut-off value to avoid the inflation of false-positive findings. A p-value of < 0.017 was considered statistically significant.

Results

Sample characteristics

The characteristics of 50 perimenopausal women in the study are shown in *Table 8.1*. Subjects had a mean (SD) age of 48.4 (3.9) years and were on average slightly overweight (mean BMI [SD], 27.6 [6.6]). Most women (82.0%) reported having moderate to severe irritability at least once during follow-up. On average, irritability severity was within the subclinical range (median irritability score 5.5,

interquartile range [IQR] 3 - 9), but it differed substantially week-to-week, both between and within individual women. On average, the values of between-person variability over the eight-week study period (mean coefficient of variation 72.9%) were greater than the within-person variability over the same timespan (mean coefficient of variation 63.7%). Although one quarter (28.0%) of women had a history of MDD, the entire sample was only mildly depressed during the eight-week study (mean MADRS score 10.8 [6.5]).

Predictors of irritability

Results of univariate and multivariate analyses are shown in *Table 8.2* and *Figure 8.1*. In univariate analyses, less variability in estradiol and higher depressive symptom severity were each associated with more irritability (β [95% confidence interval (CI)] within-person coefficient of variation estradiol -0.12 [-0.20, -0.03]; $p < 0.007$, and depressive symptom severity 0.40 [0.311, 0.49]; $p < 0.001$). More frequent VMS, younger age, and higher BMI each significantly predicted more irritability (β [95%CI] 0.12 [0.02, 0.21]; $p = 0.015$, -0.12 [-0.21, -0.04]; $p = 0.005$, and 0.24 [0.15, 0.34]; $p < 0.001$, respectively). Neither absolute values, nor weekly changes in levels of estradiol, progesterone and FSH were predictive of irritability. Variability in progesterone and FSH were also not associated with irritability. In multivariate analyses, lower estradiol variability, worse depressive symptom severity, younger age, and more frequent VMS remained independently associated with more irritability (β [95%CI] -0.23 [-0.32, -0.14]; $p < 0.001$, 0.45 [0.35, 0.56]; $p < 0.001$, -0.23 [-0.32, -0.14]; $p < 0.001$, 0.14 [0.05, 0.23]; $p = 0.002$, respectively). Depressive symptom severity explained the largest proportion of the variance in irritability (20.3%), followed by age and estradiol variability (both 5.3%), and VMS frequency (2.0%).

Sensitivity analyses

A different imputation approach for estradiol levels below the detection limit did not change the findings meaningfully (data not shown).

Table 8.1. Characteristics of 50 perimenopausal women.

	Perimenopausal women n = 50	
	mean / no.	SD / %
Age , mean (SD), y	48.4	3.9
Caucasian , no. (%)	27	54.0
College graduate , no. (%)	22	44.0
Body mass index , mean (SD)	27.2	6.6
Lifetime history of MDD , no. (%)	14	28.0
Depression severity , mean (SD)	10.8	6.5
Vasomotor symptom frequency , mean (SD)	2.6	3.1
Insomnia severity , mean (SD)	10.6	5.1
Serum reproductive hormone levels		
Estradiol, mean (SD), pg/ml	86.6	101.0
Progesterone, mean (SD), ng/ml	2.0	4.2
FSH, mean (SD), IU/L	34.5	36.6
One-week change estradiol, mean (SD), pg/ml ^a	0.1	103
One-week change progesterone, mean (SD), ng/ml ^a	0.1	5.5
One-week change FSH, mean (SD), IU/L ^a	-0.2	12.2
Variability estradiol, mean (SD) ^b	76.6	41.4
Variability progesterone, mean (SD) ^b	107.0	69.9
Variability FSH, mean (SD) ^b	49.1	25.0
Irritability severity , mean (SD)	6.4	4.8
Visits with at least moderate irritability (Hostility scale Symptom Questionnaire >7.7) , no (%)	143	44.6

Abbreviations: FSH, follicle stimulating hormone; MDD, major depressive disorder; SD, standard deviation. To convert estradiol pg/ml to pmol/l multiply by 3.67. To convert progesterone ng/ml to nmol/l multiply by 3.18. Age, Caucasian, college graduate, body mass index, history of MDD were determined at study entry. All other variables were assessed at every visit. ^a Differences in hormone levels between two successive visits (hormone level_{visit x} – hormone level_{visit x-1}). ^b Within-person coefficient of variation ([within-subject standard deviation [SD] of hormone level_{visit 1 to 9} / within-subject mean hormone level_{visit 1 to 9}]*100%)

Figure 8.1. Predictors of irritability that achieved the significance threshold in the multivariate analyses in 50 perimenopausal women with mild-to-moderate depressive symptoms.

Abbreviations: cv, coefficient of variation; MADRS, Montgomery Åsberg Depression Rating Scale; VMS, vasomotor symptoms. Significant predictors of irritability in perimenopausal women in multivariate analysis. Regression lines were fitted. Note that this is a presentation of the mean raw data per participant and that this does not fully correspond with the way the data was analyzed (multilevel).

Post Hoc Analyses

We explored whether the within-person mean estradiol might confound the association between variability of estradiol and irritability. As shown in *Figure 8.2*, there was a positive association between the mean and SD in estradiol levels for each participant. However, addition of mean estradiol to the multivariate model did not change the effect size of estradiol variability on irritability symptoms (change in b within-person coefficient of variation estradiol = 0.0%). Variability in estradiol still significantly predicted irritability (β [95%CI] -0.23 [-0.31, -0.14]; $p < 0.001$), but mean estradiol levels did not (-0.08 [-0.15, -0.001]; $p = 0.05$).

Table 8.2. Associations of potential predictors with severity of irritability symptoms in 50 perimenopausal women with mild-to-moderate depressive symptoms.

	Irritability (Hostility scale Symptom Questionnaire)					
	Univariate ^a			Multivariate ^b		
	β	95%CI	p	β	95%CI	p
Sex hormones						
Estradiol	-0.03	-0.11, 0.06	.58			
Progesterone	-0.04	-0.14, 0.05	.37			
FSH	0.02	-0.14, 0.09	.69			
Variability estradiol ^d	-0.12	-0.20, -0.03	.007	-0.23	-0.32, -0.14	<.001
Variability progesterone ^d	-0.003	-0.10, 0.09	.94			
Variability FSH ^d	-0.07	-0.16, 0.02	.13			
One-week change estradiol ^c	0.08	-0.01, 0.18	.09			
One-week change progesterone ^c	-0.01	-0.10, 0.08	.78			
One-week change FSH ^c	0.05	-0.07, 0.17	.42			
Age	-0.12	-0.21, -0.04	.005	-0.23	-0.32, -0.14	<.001
Caucasian race	-0.04	-0.24, 0.15	.64			
College	-0.13	-0.32, 0.07	.20			
Stressful life events	0.06	-0.16, 0.03	.19			
Body mass index	0.24	0.15, 0.34	<.001	0.12	0.02, 0.21	.017
Vasomotor symptom frequency	0.12	0.02, 0.21	.015	0.14	0.05, 0.23	.002
Insomnia severity	0.12	0.02, 0.22	.02			
Depression severity	0.40	0.31, 0.49	<.001	0.45	0.35, 0.56	<.001
Lifetime history of MDD	0.04	-0.18, 0.27	.070			

Abbreviations: β , standardized beta-coefficient; CI, confidence interval; FSH, follicle stimulating hormone; MDD, major depressive disorder. Estimates are bs determined by generalized estimating equation analyses. ^a Adjusted for days between visits. ^b Adjusted for ^a and for predictors that achieved statistical significance in univariate analyses. ^c Crude differences in hormone levels between two successive visits (hormone level_{visit x} - hormone level_{visit x-1}). ^d Within-person coefficient of variation ([within-subject standard deviation [SD] of hormone level_{visit 1 to 9} / within-subject mean hormone level_{visit 1 to 9}]*100%)

As a second post-hoc analysis, we explored whether inter-individual differences in the direction of the associations between the weekly changes in estradiol and irritability might explain why the overall variability variable of estradiol was a predictor of irritability, but the weekly change variable was not.³²⁹ The within-woman association of both crude and absolute values of weekly changes in estradiol with irritability varied greatly between women in strength and direction (range b for crude values of weekly changes: -26.50 to 5.09; for absolute values of weekly changes: -3.10 to 4.00). Analysis of the largest association in magnitude for each woman showed that some were sensitive to increases of estradiol (8.5%), some to decreases of estradiol (8.5%), and some to either increases or decreases (6.4%).

However, for the majority (76.6%) of women, these associations were not significant.

Figure 8.2. Variability of serum estradiol levels over eight weeks (a) and the association of mean serum estradiol levels with standard deviation (SD) estradiol (b), showing that women with more variability in estradiol also have on average higher mean estradiol levels over the visits compared to women with less estradiol variability.

Figure A shows lines that represent serum estradiol levels within individual women selected from the upper and lower quintiles of the within-person coefficient of variation of estradiol ($n = 20$). Figure B shows dots for the intersection of mean estradiol and the SD of estradiol for each of the participants ($n = 50$). The dots are labeled with the corresponding within-person coefficient of variation (in %) for estradiol.

Discussion

Irritability was highly prevalent in this cohort of mildly depressed perimenopausal women with weekly assessments over eight weeks. Women who had less variable serum estradiol levels, more frequent VMS, or who were younger, reported more irritability, associations that were independent of depressive symptom severity. Our results emphasize that irritability is an important symptom domain in mild depression during the perimenopause which can be disentangled from depressive symptoms and have hormonal risk factors that do not match those of depressive symptoms.

The finding that less variable estradiol levels, in contrast to crude levels, were predictive of more irritability, underlines the importance of measuring hormone serially in order to capture their dynamics, rather than obtain a single hormone level value. Importantly, the association was not driven by the propensity of low variability of estradiol to be accompanied by low mean levels of estradiol. Previous studies in perimenopausal women examining variability in estradiol in relation

to the specific mood symptoms of irritability are lacking. We and others have shown that greater weekly estradiol variability is associated with more depressive symptoms.^{117, 322} Hence, our findings might be extrapolated to suggest that, while suppression of estradiol variability might improve depressed mood, but not ameliorate irritability. This interpretation is supported by findings from two randomized controlled trials in 178 premenopausal and 725 early postmenopausal women, showing that estrogen supplementation improves depressive symptoms but worsens or does not have change irritability symptoms.³³⁰⁻³³² However, our findings are not consistent with those of a smaller trial ($n = 16$ for intervention and $n = 18$ for placebo) of perimenopausal depressed women showing that estradiol treatment improves both irritability and depressive symptoms more than placebo.³³³ Taken together, these findings suggest that different hormonal mechanisms may underlie irritability and depressive symptoms as mood symptoms that can be disentangled during the perimenopause, although further studies are needed to examine the specific mechanisms involved.

There are many potential ways through which estradiol may be involved in affect regulation and dysregulation. Evidence supporting this statement is reviewed in detail elsewhere,³³⁴ but include the following: (a) Estradiol can regulate gene transcription by intracytoplasmic receptor binding. These estrogen receptors are predominantly located in regions of the limbic system, such as the amygdala.⁴⁸ This system is important for emotion regulation and implicated in affective disorders such as depression, but also premenstrual dysphoric disorder in which irritability is a cardinal symptom.⁷⁹ (b) Estradiol can alter release of all classical neurotransmitters, such as serotonin.⁴⁸ The role of serotonin in the pathophysiology of affective disorders is supported by numerous findings using indirect measures of central serotonin and its transmission.⁷⁹ (c) Estradiol modulates systems implicated in the pathophysiology of affective disorders such as the hypothalamic-pituitary-adrenal-axis (HPA-axis) and the immune system.⁸² However, the exact mechanism underlying the results of the present study remain to be elucidated, and further speculation is beyond the scope of this publication.

The time span during which one might expect changes in estradiol to relate to irritability, is unknown. We observed that, in contrast with overall variability, week-to-week changes in estradiol were not associated with concurrent irritability levels. These findings suggest that weekly changes in estradiol do not explain changes in irritability during that same week. In studies of hormonal links

with perimenopausal depressive symptoms, others ³²⁹ have highlighted the importance of accounting for the between-women variability in the direction of the association between estradiol levels and mood, indicating that mood in some women may relate to changing levels regardless of the direction of the change while the direction of the change (increase versus decrease) might have a more important effect on mood for others. Hence, analyzing all women together, might “cancel out” a significant finding. We observed that a quarter of the women were sensitive to estradiol changes, but that the direction of the association between weekly estradiol change and irritability varied in direction and strength. While this variation in directions of the associations may have “cancelled out” an effect, the vast majority of women did not have any association of weekly change in estradiol level with irritability, which likely explains why we did not observe that irritability levels were associated with weekly estradiol changes.

In contrast to estradiol, progesterone levels were not associated with irritability. As the perimenopause is characterized by a reduction in progesterone peaks due to a reduction in the number of ovulatory cycles,³¹⁰ our results suggest that sensitivity to variability in progesterone is not a risk factor for irritability during the perimenopause. Alternatively, we might not have been able to detect variability progesterone as levels were collected weekly, which may have missed an ovulatory peak in progesterone during the intervening days.

Our observation that during the study women reported irritability of at least moderate severity at almost half of the visits highlights how prevalent irritability is among perimenopausal women who are mildly depressed. Though depression severity was the strongest predictor of irritability, it explained less than a quarter of the variance of irritability, indicating that irritability and depressive symptoms, when measured with validated questionnaires, are correlated but commonly represent distinct mood states. Irritability is a unique mood symptom in that it is not considered a core diagnostic symptom of MDD among adults.⁹ Moreover, MDD is typically characterized by apathy and flattened affect, the reverse of irritability. When women with MDD manifest irritability as one of the symptoms of a current depressive episode, they have more severe depressive and anxiety symptoms,³²⁰ and are more likely to have a chronic course,³¹⁷ and a lower quality of life³³⁵ than those with MDD but no irritability. In addition, early reductions in irritability predict favorable antidepressant treatment outcomes independently of depressive symptom severity.³³⁶ These observations suggest that depressive and

irritability symptoms are often comorbid and each warrant therapeutic attention to optimize well-being in perimenopausal women with affective illness.

Non-hormonal risk factors for perimenopausal irritability, such as VMS and younger age, are consistent with those observed for depression severity during the perimenopause.^{112, 117, 173, 292, 323, 337} However, some predictors of depressive symptomatology, such as higher BMI, more stressful life events, and a history of MDD,^{118, 322, 337} were not predictive of irritability in our sample. Previous research showed that a history of premenstrual syndrome (PMS), rather than of MDD, is associated with irritability in women transitioning from pre- to postmenopause,¹¹¹ Unfortunately, we were unable to examine the association with PMS as retrospective reports about PMS were missing in half of our study population. However, it is possible that perimenopausal irritability may have more in common with a history of PMS, which uniquely has irritability as a cardinal symptom, than with previous MDD. This supports our observation that perimenopausal irritability and depression are overlapping dysphoric mood states which can be disentangled in both their psychiatric history associations and in their concurrent reproductive hormone dynamics.

Strengths of this study include the weekly, concurrent assessment of irritability and depressive symptoms, reproductive hormones, and other potential predictors of irritability in a well-characterized cohort of perimenopausal women. This is important because concurrent examination of predictors with the outcome provides more robust insight into the real-time risk factors for irritability symptom during this dynamic phase of a woman's life. To the best of our knowledge, this is the first study to investigate hormone dynamics rather than absolute levels in relation to irritability, a critical advance in our understanding given the nature of hormonal fluctuations in the perimenopause. In addition, estradiol was assayed using the gold-standard LC/MS-MS, which has improved reliability over immunoassays in the low estradiol range.³¹⁵ This study also has some limitations. First, use of observational data precludes any causal inference. Second, the weekly assessments of predictors and irritability didn't capture day-to-day variability. Future studies should elucidate whether this difference in time-span is relevant. Finally, generalizability of results to pre- or postmenopausal women without mild-to-moderate depressive symptomatology might be limited owing to women's susceptibility for depressive symptomatology during the perimenopause and to the unique patterns of variability seen in reproductive hormones during this reproductive transition.

Findings from this study reveal the high prevalence of irritability in mildly depressed perimenopausal women. These women were more likely to report worse severity of irritability when they had less variable estradiol levels, were younger, and had more frequent VMS. Though some predictors overlap with those for depression severity, the decreased rather than increased variability in estradiol appears to be an irritability-specific risk factor. Although our findings need to be replicated, our results indicate that an irritable mood among depressed perimenopausal women may be related to neural processes responding to a differential estradiol profile.

Acknowledgments

This work was supported by the National Institutes of Mental Health [5R01MH082922]. H Joffe's research program is supported by grants from NIH, V Foundation, Merck, Pfizer, Que-Oncology, and NeRRe/KaNDy. H Joffe is also a consultant for NeRRe/KaNDy, Merck, Sojournix, Eisai, and Jazz Pharmaceutical. Other authors have nothing to disclose. The funding source had no role in the design and conduct of the study, collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Part II

Hormonal contraceptives
and depression

Chapter 9

Association of use of oral contraceptives with depressive symptoms among adolescents and young women

Anouk E de Wit
Sanne H Booij
Erik J Giltay
Hadine Joffe
Robert A Schoevers
Tineke AJ Oldehinkel

Adapted version from the article published in JAMA Psychiatry, October 2019.

Abstract

Aim: Use of oral contraceptive pills (OC)s have been associated with an increased risk of subsequent clinical depression in adolescents. However, the association with concurrent depressive symptoms remains unclear. Therefore, we aimed to investigate the association between OC use and depressive symptoms, and to examine the influence of age and which specific symptoms are affected.

Methods: Data from girls aged 16 to 25 from the prospective cohort study TRacking Adolescents' Individual Lives Survey were used. Depressive symptoms were assessed by the Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV oriented affective problems scale of the Youth- (aged 16 years) and Adult-Self-Report (aged 19, 22, and 25 years). Multilevel analyses were performed to examine the association between OC use and depressive symptoms. An interaction term was added to determine the influence of age.

Results: Data from a total of 1010 girls (743-903 girls, depending on the wave) were analyzed (mean [standard deviation (SD)] age at the first assessment of OC use, 16.3 [0.7]; mean [SD] age at the final assessment of OC use, 25.6 [0.6] years). OC users particularly differed from nonusers at age 16 years, with nonusers having a higher mean (SD) socioeconomic status (0.17 [0.78] vs -0.15 [0.71]) and more often being virgins (424 of 533 [79.5%] vs 74 of 303 [24.4%]). Although all users combined (mean [SD] ages, 16.3 [0.7] to 25.6 [0.6] years) did not show higher depressive symptom scores compared with nonusers, adolescent users (mean [SD] age, 16.5 [0.7] years) reported higher depressive symptom scores compared with their nonusing counterparts (mean [SD] age, 16.1 [0.6] years) (mean [SD] score, 0.40 [0.30] vs 0.33 [0.30]), which persisted after adjustment for age, socioeconomic status and ethnicity (beta-coefficient [95% confidence interval (CI)] interaction with age, -0.021 [-0.038 to -0.005]; $p = 0.0096$). Adolescent contraceptive users particularly reported more crying (odds ratio 1.89 [95%CI] [1.38-2.58]; $p < 0.001$), hypersomnia (1.68 [1.14-2.48]; $p = 0.006$), and more eating problems (1.54 [1.13-2.10]; $p = 0.009$) than nonusers.

Conclusion: Although OC use showed no association with depressive symptoms when all age groups were combined, 16-year-old girls reported higher depressive symptom scores when using OCs. Monitoring depressive symptoms in adolescents who are using OCs is important, as the use of OCs may affect their quality of life and put them at risk for nonadherence.

Introduction

Oral contraceptive pills (OC)s are often used by women to prevent pregnancies or diminish menstrual symptoms. Successful fertility control engenders many social and economic benefits for women, such as better education and personal autonomy.³³⁸ However, 32% to 60% of women discontinue OC use within six months for varying reasons, including mood changes.³³⁹⁻³⁴¹ Research findings regarding OC use and concurrent mood changes are inconsistent, and range from improved mood or having fewer mood swings³⁴²⁻³⁴⁸ to worsened mood^{124, 349-353} or having no effect at all.³⁵⁴⁻³⁶¹ These results suggest that some women may benefit from OC use whereas others do not, or are even negatively affected.

The above-described heterogeneity in findings may be related to differences in study populations, as previous research showed that the incidence of adverse OC effects may depend on age. In a large Swedish population-based observational study, adolescents were more likely than adult women to subsequently start using psychotropic drugs after initiation of OC use.³⁶² A comparable pattern was found in a Danish observational study that showed that adolescents were more prone than adult women to seek psychiatric help when using OCs.¹²⁴ This finding suggests that the first onset of clinical depression associated with OC use is particularly salient during early adolescence. Whether this phenomenon also holds true for the concurrent presence of depressive symptoms is unknown. Alternatively, the inconsistent results might also be because of the use of diverse (unvalidated) mood questionnaires.³⁶³ Assessment of separate symptoms with validated questionnaires is of interest, as a randomized clinical trial revealed that OC use increased mood swings and irritability symptoms, but improved depressive symptoms.³³⁰

Therefore, we aimed to examine the associations between OC use and concurrent depressive symptoms, test whether these associations are affected by adolescence, and determine which particular symptoms—if any—are associated with OC use. We used data from a large longitudinal survey of adolescents in the Netherlands, including 1010 adolescent girls whose depressive symptoms were assessed using validated and sensitive questionnaires for depressive symptoms in up to four waves for nine years.³⁶⁴

Methods

Study population

Data were derived from the Dutch population survey TRacking Adolescents' Individual Lives Survey (TRAILS), conducted from September 1, 2005, to December 31, 2016. TRAILS is a large prospective cohort designed to investigate the psychological, social, and physical development of adolescents. Children were recruited from primary schools in 2001 and 2002.¹³⁷ Exclusion criteria were serious health or language problems that would hamper full participation in the study. In total, 2230 children (mean [standard deviation (SD)] age, 11.1 [0.6] years; 1137 girls) were enrolled in the study. After baseline assessment, children were followed up at the median ages of 13 (T2), 16 (T3), 19 (T4), 22 (T5), and 25 (T6) years.³⁶⁵ All follow-up assessments had good retention rates (the proportion of the baseline sample) of 96.4% (2149 of 2230; T2), 81.4% (1816 of 2230; T3), 84.3% (1881 of 2230; T4), 79.7% (1778 of 2230; T5), and 72.6% (1618 of 2230; T6). The present study included girls and young women who had filled out at least one and up to four assessments of OC use and depressive symptoms from wave T3 to T6 (ages, 16-25 years). OC pill use was assessed at waves T3 to T6 with the question: "Do you use an oral contraceptive pill?" and as part of a medication list provided by the participants. Fourteen data points were excluded owing to discrepancies between the OC use question and the medication list. An additional 37 observations were excluded as participants were using sex steroids other than OCs or were pregnant at that point. These exclusions resulted in a final sample of 3317 data points from 1010 girls and young women (mean, 3.3 observations per participant). Excluded participants (T3, 107 of 952 [11.2%]; T4, 80 of 983 [8.1%]; T5, 111 of 937 [11.8%]; T6, 49 of 792 [6.2%]) more often had a lower socioeconomic status (difference in socioeconomic status score, 0.22 at age 16 and 19 years; 0.28 at age 22 years; 0.12 at age 25 years) than included girls at all waves (*Supplement 18*). The national ethical Central Committee on Research Involving Human Subjects approved the study design. At baseline (T1) parents provided written consent and children provided oral assent. At wave T2 and T3, both parents and children provided written consent and assent for participation in the new wave. From T4 to T6, children were at least 18 years of age and provided written consent.

Depressive symptoms

Depressive symptom scores at age 16 years were assessed with the Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV-oriented affective problems scale of the Youth Self-Report.³⁶⁶ This self-report version of the Child Behavior Checklist consists of items on crying, eating, sleeping, suicidal ideation, self-harm, feelings of worthlessness and guilt, energy, activity, sadness, and anhedonia. Items were rated as 0 = not true, 1 = a little or sometimes true, or 2 = very or often true for the preceding six months. At ages 19, 22, and 25 years, depressive symptoms were assessed with the affective problems scale of the Adult Self-Report,³⁶⁷ the adult version of the Youth Self-Report. This scale includes the above-mentioned items plus 2 items on decision-making and feelings of not being able to succeed. The scale scores represent the mean item scores (range, 0-2), with a score of more than 1 corresponding closely to the DSM-IV criteria for major depressive disorder (MDD).³⁶⁸ Cronbach α in our samples were 0.78 at age 16 years, 0.84 at age 19 years, 0.85 at age 22 years, and 0.87 at age 25 years.

Statistical analyses

Statistical analysis was performed from March 1, 2017, to May 31, 2019. Population characteristics of OC users and nonusers were compared by mean values of χ^2 tests or t tests, depending on the type of variable. First, the association between OC use and depressive symptoms was examined with a linear mixed model considering OC use as a fixed factor. A random intercept and random slope were added to the model as they lowered the Akaike information criterion. The random intercept accounted for between-participant variability of depressive symptoms and the random slope (modeled by age) allowed for individual slopes for depressive symptoms over time. Mixed models with different covariance structures were fitted using restricted maximum likelihood estimation and, in this case, variance components was chosen based on the Akaike information criterion. The final model was estimated with maximum likelihood estimation. Second, the interaction term age*OC use was added to the model to investigate whether age modified the association. Stratified analyses were performed to examine the association per age category using linear regression analyses. Third, associations between OC use and specific depressive symptoms were determined using binary logistic regression analyses in the age groups that showed significance in the

interaction analysis.

Exploratory Analyses

Women who experience psychological adverse effects are more likely to discontinue their OC use (the “healthy survivor effect”).¹²⁴ As this healthy survivor effect would underestimate the association between OC use and depressive symptoms, the same mixed model was repeated in girls who had never used OC before, in comparison with never users. As an alternative approach, duration of use (0, no OC use; 1, one month-three years; 2, three-six years; 3, six-nine years, and 4, nine-twelve years, based on the number of subsequent assessments with reported OC use) was added to the model as a covariate. To explore whether the potential association between OC use and depressive symptoms could be bidirectional, we adjusted the association between OC use and depressive symptoms at age 16 years for depressive symptom scores at age 13 years (T2). All analyses were conducted with and without adjustment for age (time varying), socioeconomic status, and ethnicity. Age was centered and divided by three to ease interpretation for the effect per each three-year increase in age. Socioeconomic status was assessed on a continuous scale based on the educational level, job, and income of the parents at study entry, with higher values indicating a better social and economic position. Ethnicity was categorized as having a Dutch or non-Dutch ancestry.

Data were analyzed using IBM SPSS Statistics, version 24 (IBM Corp) using two-sided tests. $p < 0.01$ for the main analyses and $p < 0.05$ for the exploratory analyses were considered statistically significant.

Results

Data from 743 to 903 girls, depending on the wave, were used for the analyses. Girls had a mean (SD) age of 16.3 (0.7) years at the first assessment of OC use (T3) and 25.6 (0.6) years at the final assessment (T6). Descriptive statistics of OC users and nonusers are shown in the *Table 9.1*. Oral contraceptive pill users differed particularly from nonusers at age 16 years. During this assessment, nonusers had a higher mean (SD) socioeconomic status than users (0.17 [0.78] vs -0.15 [0.71]), were more likely to be of non-Dutch ethnicity (76 of 536 [14.2%] vs 24 of 309 [7.8%]), and were more likely to be virgins (424 of 533 [79.5%] vs 74 of 303 [24.4%]). For the cohort as a whole, OC use was not associated with higher adjusted mean scores

on depressive symptoms (beta-coefficient [β] [95% confidence interval (CI)] 0.006 [-0.013 to 0.025]; $p = 0.52$). As depicted in *Figure 9.1*, age significantly affected the association (β [95%CI] age $\times$ OC use -0.021 [-0.038 to -0.005]; $p = 0.0096$); this association was driven by differences in 16-year-old girls. At this age, girls who used OC had higher concurrent depressive symptom scores than their nonusing counterparts (mean [SD] score, 0.40 [0.30] vs 0.33 [0.30]; B [95%CI], 0.075 [0.031-0.120]; $p < 0.001$).

Figure 9.1. Oral contraceptive (OC) use and depressive symptom scores by age groups.

Data are (adjusted) mean values with 1 standard error for the depressive symptom score through multilevel analysis (ie, mixed models). Analyses were adjusted for age, ethnicity, and socioeconomic status (beta-coefficient for interaction [B] = -0.021 [95%CI, -0.038 to -0.005; $p = 0.0096$]).

Table 9. 1. Socio-demographic and clinical characteristics of the participants per wave.

	16 years of age				19 years of age				22 years of age				25 years of age			
	No OC use n = 536	OC Use n = 309	p		No OC use n = 290	OC use n = 613	p		No OC use n = 283	OC use n = 543	p		No OC use n = 398	OC use n = 345	p	
Age, mean (SD)	16.1 (0.6)	16.5 (0.7)	<.001		18.3 (0.6)	19.0 (0.6)	.01		22.3 (0.6)	22.2 (0.7)	.004		25.6 (0.6)	25.6 (0.6)		
Dutch ethnicity, no. (%)	460 (85.8)	285 (92.2)	.006		235 (81.0)	564 (92.0)	<.001		235 (83.0)	501 (92.3)	<.001		352 (88.4)	317 (91.9)	.14	
Socioeconomic status, mean (SD) ^a	0.17 (0.78)	-0.15 (0.71)	<.001		0.11 (0.77)	0.03 (0.76)	.15		0.03 (0.78)	0.12 (0.74)	.13		0.07 (0.77)	0.16 (0.74)	.09	
Body mass index, mean (SD)	21.5 (3.1)	22.0 (3.1)	.01		23.3 (4.3)	23.0 (3.7)	.26		24.1 (4.7)	23.6 (4.2)	.12		24.1 (4.7)	24.0 (4.8)	.58	
Virgin, no. (%)	424 (79.5)	74 (24.4)	<.001		157 (54.7)	85 (13.9)	<.001		54 (19.1)	44 (8.1)	<.001		25 (6.3)	17 (4.9)	.43	
Age at menarche, mean (SD)	13.0 (1.2)	12.6 (1.1)	<.001		12.3 (1.2)	12.3 (1.1)	.97		12.3 (1.2)	12.9 (1.1)	.50		12.8 (1.1)	13.0 (1.1)	.01	
Antidepressant use, no. (%)	0 (0.0)	1 (0.3)	.37		4 (1.4)	6 (1.0)	.74		11 (2.0)	11 (2.0)	.17		16 (4.0)	15 (4.3)	.86	
Outcomes																
Depressive symptoms, mean (SD)	0.33 (0.30)	0.40 (0.30)	.001		0.36 (0.34)	0.36 (0.32)	.96		0.39 (0.34)	0.37 (0.32)	.29		0.45 (0.37)	0.42 (0.34)	.19	
Sadness, no. (%)	184 (34.4)	196 (35.5)	.76		106 (36.6)	231 (37.7)	.77		100 (35.3)	178 (32.8)	.49		182 (45.7)	146 (42.3)	.37	
Anhedonia, no. (%)	89 (16.7)	63 (20.6)	.16		39 (13.4)	99 (16.2)	.32		46 (16.3)	72 (13.3)	.25		91 (22.9)	53 (15.4)	.01	
Worthlessness, no. (%)	129 (24.2)	72 (23.5)	.87		80 (27.7)	139 (22.7)	.11		84 (29.7)	137 (25.2)	.19		140 (35.2)	101 (29.3)	.10	
Guilty, no. (%)	132 (24.8)	70 (22.9)	.56		60 (20.7)	110 (17.9)	.36		60 (21.2)	100 (18.5)	.35		109 (27.4)	72 (20.9)	.04	
Indecisiveness, no. (%)	-	-	-		150 (51.7)	286 (46.7)	.18		131 (46.3)	266 (49.1)	.46		225 (56.5)	181 (52.5)	.27	
Tiredness, no. (%)	277 (51.9)	183 (60.0)	.03		121 (41.7)	301 (49.1)	.04		142 (50.2)	321 (59.2)	.02		229 (57.2)	217 (62.9)	.15	
Loss of energy, no. (%)	242 (45.2)	160 (52.3)	.05		123 (42.4)	285 (46.5)	.25		150 (53.0)	261 (48.2)	.19		224 (56.3)	203 (58.8)	.50	
Insomnia, no. (%)	181 (33.8)	120 (39.3)	.12		96 (33.1)	214 (34.9)	.60		139 (49.3)	214 (39.5)	.008		158 (39.7)	142 (41.2)	.71	
Hypersomnia, no. (%)	77 (14.4)	73 (24.0)	<.001		77 (26.6)	177 (28.9)	.48		71 (25.2)	149 (27.5)	.51		118 (29.6)	115 (33.3)	.30	
Eating problems, no. (%)	225 (42.0)	144 (47.2)	.15		119 (41.0)	257 (41.9)	.83		120 (42.4)	216 (39.8)	.50		198 (49.7)	175 (50.7)	.83	
Non-succeeding, no. (%)	-	-	-		101 (34.8)	171 (27.9)	.04		95 (33.6)	146 (26.9)	.05		175 (44.0)	109 (31.6)	<.001	
Self-harm, no. (%)	32 (6.0)	21 (6.9)	.66		15 (5.2)	16 (2.6)	.05		9 (3.2)	11 (2.0)	.34		11 (2.8)	6 (1.7)	.46	
Suicidal ideation, no. (%)	37 (6.9)	22 (7.2)	.89		16 (5.5)	26 (4.2)	.40		13 (4.6)	21 (3.9)	.71		29 (7.3)	13 (3.8)	.04	
Crying, no. (%)	179 (33.5)	146 (48.0)	<.001		104 (35.9)	274 (44.7)	.01		126 (44.5)	268 (49.4)	.21		171 (43.0)	142 (41.2)	.66	

Abbreviations: OC, oral contraceptive pill; SD, standard deviation. Data are based on two-tailed χ^2 tests for dichotomous variables or t tests for continuous variables. All variables were assessed at every wave, except for socioeconomic status, ethnicity, and menarche of which the first two were determined at study entry (T1) and the latter at age 16 (T3). ^aContinuous scale based on the educational level, job, and income of the parents.

The depressive symptom item scores at age 16 years are depicted in *Figure 9.2*. Use of OCs was associated with more crying (odds ratio [OR] [95%CI] 1.89 [1.38-2.58]; $p < 0.001$), eating problems (1.54 [1.13-2.10]; $p = 0.009$), and hypersomnia (1.68 [1.14-2.48]; $p = 0.006$) compared with nonusers.

Figure 9.2. Depressive item scores according to oral contraceptive pill use in 16-year-olds.

Abbreviations: OC, oral contraceptive pill. Data are (adjusted) odds ratio with 95% confidence intervals for severe and moderate symptoms vs the reference group (no symptoms) through binary logistic regression analysis.

Exploratory analyses

Excluding girls who used OC previously (regardless of whether they continued or discontinued use) strengthened the association between OC use and depressive symptoms for the whole cohort (β [95%CI] for first-time OC use 0.021 [−0.005 to 0.046]; $p = 0.11$) (*Supplement 19*). Also, long-term OC use (up to three, six, nine, and twelve years) decreased the likelihood of higher depressive symptoms scores, although this finding was not significant (β [95%CI] −0.007 [−0.024 to 0.009]; $p = 0.39$). Adjustment for depressive symptom scores before OC use weakened the association in 16-year-old girls (β [95%CI] OC use 0.040 [0.001-0.079]; $p = 0.046$).

Post hoc analyses

To investigate whether the age-specific association may be explained by residual confounding or preexisting differences, we explored which differences between

OC users and nonusers were already present at the age of 13 years, when OC use was unlikely. At age 13 years, the 16-year-old user group was more often sexually active, experienced more stressful events, and more often had menstrual-related pain and acne than the 16-year-old nonuser group (*Supplement 20*). In subsequent analyses, these factors were added one by one to the adjusted model (that already included age, ethnicity, and socioeconomic status) to examine whether they individually affected the association. All four factors weakened the association (change in B for OC use: virginity, -9.3%; stressful events, -26.7%; menstrual-related pain, -21.3%; acne, -17.3%). However, none of the factors diminished the association between OC use and depressive symptoms at age 16 years.

Discussion

In this large cohort study of adolescents, we did not find support for an overall association between OC use and depressive symptoms among young women. However, 16-year-old girls using OCs did report higher concurrent depressive symptom scores compared with their nonusing counterparts. When individual depressive symptoms were examined, 16-year-old girls using OCs particularly reported more crying, hypersomnia, and eating problems, while the required symptoms for the diagnosis of depression—anhedonia and sadness—were unaffected. Adjusting the association between OC use and depressive symptoms for pre-OC use depressive symptoms in 16-year-old girls reduced the magnitude of the association, but did not eliminate the significance of the findings.

Earlier clues of a possible association between OC use in adolescence and vulnerability for depressive symptoms come from two large population-based studies.^{124, 362} These studies showed that adolescent girls using OC were the most likely to subsequently start using psychotropic drugs or be admitted to a psychiatric hospital compared with nonusers. Our study provides important evidence that this vulnerable group is the most likely to report more concurrent depressive symptoms.

This age-specific vulnerability is complex and likely bidirectional. Important emotion-related regions of the brain, such as the amygdala, prefrontal cortex, and hippocampus, are still maturing during adolescence and therefore may be particularly sensitive to sex hormone-related changes at age 16 years.³⁶⁹ In addition, ages 15 to 18 years may be a critical time for the development of depression in girls, as the sex gap in the incidence of depression dramatically

increases during this period.¹²⁸ Neurobiological studies support this idea, as sex hormones are capable of directly influencing gene expression in the cell nuclei in the brain.⁴⁷ Furthermore, the use of OCs affects sex hormone levels, including androgens and stress hormones.³⁷⁰ Although almost nothing is known about the effect of OC use on the maturation of the brain,³⁶⁹ these findings confirm its potential capability to do so.

In addition to the possibility that OC use leads to more depressive symptoms, the weakened strength of the association after adjustment for pre-OC use depressive symptoms in 16-year-old girls also points toward an association in the reverse direction. A previous study demonstrated that mood worsening among individuals using OCs was more likely in users with a history of depression.³⁶¹ This finding might suggest that a subset of women with depression may be at risk for OC-associated mood deterioration because they are particularly sensitive to interactions between cycling gonadal steroids and affect. Also, treatment with OCs is standard care for cycle-related mood problems³⁷¹ and for polycystic ovarian syndrome, which is associated with depressive symptoms.³⁷² Therefore, mood-related physical problems may also be involved in the association between OC use and depressive symptoms. Hence, some girls in our sample may have initiated OC use because they experienced mood problems.

Residual confounding could have led to overestimation of the association. Differences between OC users and nonusers that were likely present before the initiation of OC use weakened the strength of the association at age 16 years, although none of the differences completely diminished the association. However, future studies should replicate and consider the importance of these factors. In contrast, the *healthy survivor effect* may have led to underestimation of the association between OC use and depressive symptoms. In support of this possibility, comparing first-time users with never users yielded a stronger association. This finding is in line with findings from a Danish study in which the highest relative risks (RRs) for a diagnosis of depression were found after six months (RR, 1.5) of use, whereas RRs were smaller when individuals had used OC for longer periods (RR, 1.2 for use of 1-3 years; RR, 0.9 for use of four-six years; RR, 0.8 for use of seven-nine years).¹²⁴ Our finding, combined with these data, suggests that the absence of an association in older individuals may be owing to the healthy survivor effect. The possibility of an underestimation in our younger participants cannot be excluded either, because information on OC use before the age of 16 years was not available and thus could not be explored.

One may wonder why hypersomnia, eating problems, and increased crying, rather than core symptoms of depression such as sadness, were reported more commonly among adolescent OC users. These symptoms fit well in symptom profiles of adolescent depression, where, in contrast to adult depression, the emphasis is more on vegetative or physical disturbances (e.g., loss of energy, changes in weight, and appetite and sleep changes) than on anhedonia.³⁷³ The symptom specific findings are also in line with results from a randomized clinical trial, which showed that OC use worsened mood swings and irritability symptoms, but improved depressive symptoms such as sadness,³³⁰ suggesting that associations between OC use and mood may be symptom specific.

Strengths of this study are the availability of multiple time points when depressive symptoms were assessed with the use of well-validated instruments. Also, we were able to adjust for important covariates such as socioeconomic status and ethnicity. The assessment of depressive symptoms provides a complementary, more inclusive perspective on the association between OC use and depression across the severity spectrum to existing studies using proxies such as incident clinical depression. More important, subclinical depressive symptoms may also cause distress in adolescents, and many adolescents with psychiatric problems in the clinical range do not seek help.³⁷⁴ Moreover, the concurrent examination of OC use and depressive symptom severity over several years contributes to a better insight into the real-time risks associated with OC use.

This study has some limitations. Use of observational data precludes any causal inference. Also, this longitudinal analysis does not provide information about specific OCs. Because very few girls specified OC use as medication use, we were unable to conduct separate analyses for different kinds of OCs. However, we checked which OCs were used in a comparable cohort of girls born in the same year and postal code as the girls included in this study (Inter Active DataBase cohort).³⁷⁵ In that cohort, most of the girls used either monophasic (81.1%) or triphasic ethinyl estradiol and levonorgestrel (2.5%), followed by ethinyl estradiol and drospirenone (4.6%) and ethinyl estradiol and desogestrel (1.8%). Hence, testing for specific components was not useful, as probably almost all participants were using the same type of progestin. Finally, generalizability of results to other countries might be limited owing to differences in the acceptability of and access to contraception across societies.³⁷⁶ For example, unlike Dutch teenagers, not all teenagers in the United States have access to no-cost contraception.^{377, 378} Such

differences have implications for group characteristics of girls using OCs and, hence, associated outcomes.

Although we cannot be conclusive about the direction of the association, awareness of the possible existence of mood problems among OC-using adolescents is critical to the health and development of adolescents. Depressive symptoms may reduce quality of life³⁷⁹ and OC use adherence,³⁸⁰ because, for example, girls may attribute these depressive symptoms to the use of OCs, potentially resulting in unwanted teenage pregnancies. To lower the probability of an unintended conception, long-acting reversible contraceptives are recommended as a first-line option.^{381, 382} However, those are only used by a minority of teenagers in both the Netherlands and the United States.^{383, 384} We do not suggest limiting OC use to counterbalance this risk for depressive symptoms. In fact, previous research has shown benefits of OC use such as improvement of dysmenorrhea and premenstrual syndrome,³⁸¹ additionally, OC use is much safer than pregnancy and associated postpartum depression risks.³⁸²

In this study, 16-year-old OC users showed higher concurrent depressive symptom scores compared with their nonusing counterparts. Overall, adolescent girls are more likely to report increased crying, hypersomnia, and eating problems while taking OCs, but these symptoms seem to diminish once they enter adulthood. Monitoring these symptoms is important, as they may affect their quality of life and put them at risk for nonadherence.

Acknowledgements

We are grateful to everyone who participated in this research or worked on this project to make it possible. Written permission was obtained to express our very great appreciation to the staff of IADB.nl, especially to HJ Bos, B.Sc. of the Groningen Research Institute of Pharmacy of the University of Groningen, who made their database available for examining which kinds of OCs were likely used.

Participating centers of TRAILS include the University Medical Center and University of Groningen, the University of Utrecht, the Radboud Medical Center Nijmegen, and the Parnassia group, all in the Netherlands. TRAILS has been financially supported by various grants from the Netherlands Organization for Scientific Research NWO (Medical Research Council program grant GB-MW 940-38-011; ZonMW Brainpower grant 100-001-004; ZonMw Risk Behavior and Dependence grants 60-60600-97-118; ZonMw Culture and Health grant 261-

98-710; Social Sciences Council medium-sized investment grants GB-MaGW 480-01-006 and GB-MaGW 480-07-001; Social Sciences Council project grants GB-MaGW 452-04-314 and GB-MaGW 452-06-004; NWO large-sized investment grant 175.010.2003.005; NWO Longitudinal Survey and Panel Funding 481-08-013 and 481-11-001; NWO Vici 016.130.002 and 453-16-007/2735; NWO Gravitation 024.001.003), the Dutch Ministry of Justice (WODC), the European Science Foundation (EuroSTRESS project FP-006), the European Research Council (ERC-2017-STG-757364 and ERC-CoG-2015-681466), Biobanking and Biomolecular Resources Research Infrastructure BBMRI-NL (CP 32), the Gratama foundation, the Jan Dekker foundation, the participating universities, and Accare Centre for Child and Adolescent Psychiatry. We are grateful to everyone who participated in this research or worked on this project to make it possible. The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Chapter 10

Association between adolescent oral contraceptive use and future major depressive disorder: a prospective cohort study

Christine Anderl*
Anouk E de Wit*
Erik J Giltay
Tineke AJ Oldehinkel
Frances Chen
*Shared first author

Adapted version from the article published in Journal of Child Psychology and Psychiatry, July 2021.

Abstract

Aim: Because of the widespread use of oral contraceptive pills (OC)s and the devastating effects of depression both on an individual and a societal level, it is crucial to understand the nature of the previously reported relationship between OC use and depression risk. Insight into the impact of analytical choices on the association is important when interpreting available evidence. Hence, we examined the association between adolescent OC use and subsequent depression risk in early adulthood analyzing all theoretically justifiable models.

Methods: Data from the prospective cohort study Tracking Adolescents' Individual Lives Survey, among women aged 13 to 25 years were used. Adolescent OC use (ages 16 through 19 years) was used as a predictor and MDD in early adulthood (ages 20 to 25 years) as assessed by the Diagnostic and Statistical Manual of Mental Disorders-IV oriented Lifetime Depression Assessment Self-Report and the Composite International Diagnostic Interview, was used as an outcome. A total of 818 analytical models were analyzed using Specification Curve Analysis in 534 adolescent OC users and 191 nonusers.

Results: Overall, there was an association of adolescent OC use and an episode of MDD in early adulthood (median odds ratio $[OR]_{\text{median}} = 1.41$; $OR_{\text{min}} = 1.08$; $OR_{\text{max}} = 2.18$, $p < 0.001$), which was driven by the group of young women with no history of MDD ($OR_{\text{median}} = 1.72$; $OR_{\text{min}} = 1.21$; $OR_{\text{max}} = 2.18$, $p < 0.001$).

Conclusion: In summary, adolescent OC use was associated with a small but robust increased risk for experiencing an episode of MDD, especially among women with no history of MDD in adolescence. Understanding the potential side effects of OCs will help women and their doctors make informed choices when deciding among possible methods of birth control.

Introduction

The invention of oral contraceptive pill (OC) has been liberating for women and contributed substantially to the dramatic increase in women's education and participation in the labor market.³⁸⁵ Worldwide, more than 150 million women use OC³⁸⁶ and about 50 percent of women in modern societies do so during their teenage years.¹²⁵ While OCs are generally considered safe, several recent large-scale observational studies suggest that OC use, especially during adolescence, may increase a woman's risk for depressive symptoms and even Major Depressive Disorder (MDD).^{124, 125, 387} However, other studies reported no or even protective effects of OC use on depression risk.^{122, 123} Due to these seemingly conflicting findings, the putative relationship remains disputed. However, because of the widespread use of OCs and the devastating effects of MDD on the individual as well as the societal level,⁸ it is crucial to understand the nature of the relationship between OC use and MDD risk.

One understudied factor of the relationship between OC use and depression are potential enduring effects of OCs years after use, as most previous studies have focused on examining its short-term effects. The only published study that has specifically focused on long-term effects found that adolescent OC users more often had an MDD years later, even after cessation of use.¹²⁵ However, use of OCs was assessed retrospectively and the study—like most studies in the field—relied on cross-sectional data, limiting the possibility to control for important baseline differences between users and nonusers of OC and to establish the direction of the observed association between OC use and depression risk. Prospective studies can overcome these limitations and are critical for establishing the direction of the observed association between OC use and depression risk. In the present study, we therefore set out to overcome the limitations of previous studies by addressing the question of whether adolescent OC use predicts later episodes of MDD in a large prospective data set.

In addition, we opted for conducting a Specification Curve Analysis (SCA)³⁸⁸ to account for the multitude of theoretically justifiable operationalizations and analytic choices naturally available in large data sets³⁸⁹ which can lead to contradicting findings.³⁹⁰ To illustrate, a prior cross-sectional study investigating the relationship between adolescent OC use and depression reported a wide range of effect sizes for the relationship between ever use of OC and ever having a depressive disorder (odds ratios ranged between 1.1 and 1.9), depending on the specific covariates

that the analysis adjusted for.^{122, 123} SCA is a transparent analytical approach that allows to implement every theoretically justifiable analytical pathway in parallel and interpret their combined results as one entity. Using SCA also provides a way to explore which specific clusters of analytic decisions result in divergent patterns of findings. We expected that across specifications, adolescent OC use would predict higher risk for MDD in early adulthood.

Methods

Study population

Data were derived from the Dutch population survey TRacking Adolescents' Individual Lives Survey (TRAILS), a large prospective cohort study designed to investigate the psychological, social and physical development of adolescents. The national ethical 'Central Committee on Research Involving Human Subjects' approved the study design. Children were recruited from primary schools in 2001 and 2002.¹³⁷ Exclusion criteria were serious health or language problems that would hamper participation in the study. The participants (children) assented to participating in the study and one of their parents gave informed consent. In total, 2,230 children (mean age [standard deviation (SD)] 11.1 [0.6] years; 1,137 girls) were enrolled in the study. After study entry, children were followed up at mean [SD] ages of 13.6 [0.5] (T2, baseline), 16.3 [0.7] (T3), 19.1 [0.6] (T4), 22.3 [0.3] (T5), and 25.6 [0.6] (T6) years.³⁶⁵ For readability, we refer here to "baseline" instead of T2, and to ages 16, 19, and 25 instead of T3, T4, and T6, respectively. For the present study, we selected all biologically female participants with the following exceptions: We excluded participants (1) who had not completed the Lifetime Depression Assessment Self-Report (LIDAS)³⁹¹ at age 25 (T6; $n = 356$); (2) who had no available data on current OC use for both age 16 (T3) and age 19 (T4; $n = 18$); (3) who could not be conclusively classified as either adolescent users or nonusers because they had missing data at one assessment and reported that they were not currently using OCs at the other ($n = 32$); (4) who were using sex steroids other than OC at age 16 or 19 ($n = 0$); and (5) who had discrepancies between the OC use question and the medication list ($n = 6$). This resulted in a final sample of 725 women.

Specification curve analyses

The Specification Curve Analysis consisted of three steps.³⁸⁸ First, all theoretically justified specifications that could be used to relate adolescent OC use and MDD in young adulthood were identified. Second, the specifications were analyzed and the results were plotted. Third, (in)consistency of the resulting specification curve with the null-hypothesis that adolescent OC use did not predict higher risk for MDD was determined using an inferential specification curve.

Step 1. Identifying specifications

The number of specifications was the product of all justifiable options regarding the four core elements of the model: the sample, the predictor, the outcome, and potential sets of covariates. All specifications including details on the determination of covariates are listed in *Supplement 21*.

Sample. First, the analyses were performed on the entire sample of 725 women. To focus on first-onset MDD,³⁹² the analyses were subsequently repeated on the sample of 583 participants who had no history of MDD at or before age 19. For the details on the determination of MDD see the corresponding header.

Predictor: adolescent oral contraceptive pill use. All women who reported that they were currently using OCs at either age 16 or age 19, or both were classified as ‘adolescent OC users’. All women who reported that they were not, at both age 16 and age 19, were classified as ‘adolescent nonusers’. OC use was assessed with the question: “Do you use an oral contraceptive pill?” and as part of a medication list provided by the participants. At age 13, OC use was unlikely and hence this assessment was used as baseline.³⁸⁷

Outcome: major depressive disorder in young adulthood. The World Health Organization Composite International Diagnostic Interview (CIDI, version 3.0)³⁹³ at age 19 and the LIDAS³⁹¹ at age 25 were used to determine episodes of MDD after age 19. The CIDI is a structured diagnostic interview assessing Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV mental disorders, which has been shown to have good reliability and validity.^{393,394} When a lifetime MDD was endorsed, the age of onset was interrogated. The LIDAS³⁹¹ is a self-report instrument that includes all nine DSM-IV symptoms’ criteria for MDD, supplemented with items on impairment of functioning. When at least five out of nine symptoms, including one core symptom, were present and hampered functioning for at least two weeks somewhere in the past, a lifetime MDD was diagnosed. The LIDAS has been shown

to have adequate sensitivity and specificity in measuring lifetime MDD as assessed with the CIDI.³⁹¹

Combined data from the CIDI and LIDAS data were used to construct a set of justifiable outcome variables. Following previous work, we used the 12-month MDD prevalence at age 25 as our first outcome variable.¹²⁵ This variable should be least affected by recall bias, but has a small detection window (one year) and may hence miss some episodes. Therefore, we created a second outcome variable in which all years between the assessments at age 19 and 25 were used to define any new episodes of MDD.³⁹² For this outcome variable, women whose data at ages 19 and 25 were inconsistent ($n = 11$) were excluded. As the data on the number of episodes collected at age 25 was not specific enough to establish the timing of episodes in case of multiple recurrences, women who had already experienced two or more episodes at age 19 were excluded as well ($n = 23$). Third, to test the associations of OC use with first-onset MDD in early adulthood, we used lifetime MDD at age 25 based on LIDAS in the subsample of women with no history of MDD at age 19.¹²⁴

Selection of covariates. Potential covariates were selected from variables available in the dataset that had previously been suggested to entirely or partially account for the observed association between OC use and depression -either empirically or theoretically- in addition to basic demographic information. Age at baseline, ethnicity, and socioeconomic status were included in all but the crude model as fixed covariates; history of MDD before age 19 was included as additional fixed covariate in all models analyzing the entire sample. In addition to these fixed covariates, non-fixed covariates were entered into the models one by one and in combinations of two. This enabled us to examine whether differences in OC users and non-users could confound the association, while preventing overprediction of the model by limiting the number of covariates added. Some of non-fixed covariates were time-variant (body mass index [BMI], menstruation-related pain, acne, virginity, smoking, and stress), while other were time-invariant (age at menarche, sexual orientation, OC use at age 25, and depressive symptoms at baseline). The combinations of two were either entering both variables to the model or adding the multiplicative interaction of two time-variant (within the same age, i.e., age 13, 16, and 19) and/or time-invariant covariates.

Step 2: Implementing specifications

We tested a total of 818 models between adolescent OC use and development of MDD in early adulthood, using logistic regressions in R (version 1.2.1335). The specifications were ranked by their odds ratio (OR)s and plotted in a specification curve (see *Figure 10.1*).

Step 3: Joint statistical inferences

The hypothesis was tested using the resampling approach for non-experimental data described in.³⁸⁸ In particular, the null was forced on the data by creating 818 new outcome variables—one for each specification—by subtracting the respective empirically estimated effect of the predictor variable (i.e., adolescent OC use) on the empirical outcome variable (i.e., MDD in adulthood). We then sampled, with replacement, 725 rows of data (corresponding to the empirical sample size) and repeated this procedure 1,000 times, thereby creating 1,000 new data sets under resampling for which the null hypothesis was known to be true. Finally, we ran 1,000 SCA, one for each resampled data set under the null, using the same 818 specifications identified for the empirical data set and their corresponding new outcome variables. Significance was then computed in three ways: 1) by counting the number of times that the empirically observed median OR (without resampling under the null), estimated across all specifications, was smaller than or equal to the median OR of each single resampled specification curve under the null and dividing this number by 1,000; 2) by counting the number of times that the share of significant effects (all effects with $p < 0.05$) in the predicted direction in the empirical specifications was smaller or equal to the share of significant effects in the predicted direction of each single resampled specification curve under the null, again dividing this number by 1,000; and 3) by averaging the Z values associated with all p-values of each specification curve (called Stouffer's method) and counting for how many this value is equal or less extreme for the empirical specification curve compared to each of the resampled ones for which the null is known to be true, following the procedure described in,³⁸⁸ and again dividing this number by 1,000. Because subtracting the effect size from the binary empirical outcome variable results in non-binary outcome variables under the forced null, SCA for inferential tests were conducted using ordinary least squares regression for both the empirical and the resampled data.

Results

Data from 411 through 725 women, depending on the underlying sample and number of missing data points, were included in the 818 models. Most of the participating young women (73.7%) used OCs at least once during adolescence. Descriptive statistics of adolescent OC users and nonusers at baseline and during ages 16 and 19 are shown in *Table 10.1*. Nonusers had a higher mean SES than users, were less likely to be of Dutch ethnicity, were more likely to be virgins, and were less likely to have smoked at least once. However, there were no differences in history of MDD during adolescence (33 [17.3%] vs. 109 [20.4%]) and also no differences in depressive symptoms at baseline (0.29 [0.27] vs. 0.33 [0.29]). In accordance with common sense and prior literature (Burcusa & Iacono, 2007), the 12-month prevalence of MDD at age 25 was substantially higher in women with a history of MDD at age 19 (48 out of 142 women [33.8%]) compared to women with no such diagnosis (92 out of 583 women [15.9%]), $p < .001$. There were no differences in the lifetime prevalence of comorbid psychiatric disorders between adolescent OC users and nonusers at the age of 25. Only panic and anxiety disorders were somewhat more prevalent in adolescent OC users than in nonusers (4.5% and 7.7% vs. 1.0% and 3.7%, respectively) but these differences were not statistically significant. For details, see *Supplement 22*.

Figure 10.1 shows the results of the SCA on the association between adolescent OC use and MDD in early adulthood. Notably, all ORs were larger than 1, indicating a higher MDD risk for adolescent OC users, with a median effect size of $OR = 1.41$. However, effect sizes varied considerably across specifications ($OR_{min} = 1.08$; $OR_{max} = 2.18$). The majority of the higher effect sizes were based on the subsample of women without a history of MDD at age 19 ($OR_{median} = 1.72$; $OR_{min} = 1.21$; $OR_{max} = 2.18$). The variation in effect sizes between models was predominantly (79.5%) explained by the choice of sample; choice of outcome measure accounted for 15.4% and choices regarding the inclusion of specific covariates for 2.8% of the variation in effect sizes. See also *Figure 10.2*.

Table 10.1. Socio-demographic and clinical characteristics depending on history of oral contraceptive pill use during adolescence of 725 women.

		No Adolescent OC use n = 191	Adolescent OC use n = 534	p
	Wave			
Age, mean (SD)	Baseline	13.4 (0.5)	13.6 (0.5)	.003
Dutch Ethnicity, no. (%)	Baseline	159 (83.2)	498 (93.3)	<.001
Socioeconomic status, mean (SD)	Baseline	0.27 (0.74)	0.07 (0.76)	.002
Age at menarche, mean (SD)	Age 16	12.9 (1.2)	12.7 (1.1)	.009
Body mass index, mean (SD)	Baseline	19.2 (3.2)	19.2 (3.0)	.93
	Age 16	21.6 (3.3)	21.7 (2.9)	.70
	Age 19	22.8 (4.0)	23.2 (3.8)	.31
Menstrual pain, no. (%)	Baseline	55 (31.3)	207 (43.3)	.005
	Age 16	89 (52.0)	301 (69.4)	<.001
Acne, no. (%)	Baseline	68 (37.6)	178 (36.9)	.93
	Age 16	72 (42.1)	196 (45.2)	.53
Virginity, no. (%)	Baseline	180 (98.9)	494 (95.0)	<.001
	Age 16	166 (86.9)	249 (51.3)	<.001
	Age 19	126 (66.3)	69 (13.2)	<.001
Smoking, no. (%)	Baseline	1 (0.5)	21 (4.0)	.01
	Age 16	11 (5.8)	95 (19.5)	<.001
	Age 19	21 (11.0)	127 (24.2)	<.001
Stressful experiences, mean (SD)	Baseline	2.6 (2.5)	3.0 (2.8)	.05
	Age 16	3.7 (2.5)	4.2 (2.6)	.01
Heterosexuality, no. (%)	Age 19	180 (95.2)	487 (93.3)	.07
Oral contraceptive pill use, no. (%)	Age 16	-	222 (47.4)	-
	Age 19	-	501 (96.0)	-
	Age 25	71 (37.4)	255 (48.9)	.007
History of MDD, no. (%)	Age 19	33 (17.3)	109 (20.4)	.40
Depressive symptoms, mean (SD)	Baseline	0.29 (0.27)	0.33 (0.29)	.11

Abbreviations: MDD, major depressive disorder; SD, standard deviation. P-values are based on two-tailed chi-squared tests for dichotomous variables or t-tests for continuous variables.

Figure 10.1. Specification Curve Analysis on adolescent oral contraceptive pill use and future major depressive disorder.

Specification curve analysis for the association of adolescent oral contraceptive pill (OC) use with major depressive disorder (MDD) in early adulthood. The specification curve panel shows the odds ratio (OR) with 95% confidence interval (CI). The thick grey line depicts the ORs for adolescent OC use for developing MDD during early adulthood or not across all 818 implemented model specifications sorted by effect size). The thinner vertical lines depict the corresponding 95% CI. ORs with 95% CIs are colored blue in the case that they are significant, or grey in the case that they are not significant. The 'No. subjects' panel shows the number of participants included in the specific analyses, ranging from 411 to 725. The 'Sample' panel depicts which sample were used in each of the model specifications. Finally, the 'Outcome' panel shows which of the outcome measure was used in each of the specifications.

Figure 10.2. Specification Curve Analysis on adolescent oral contraceptive pill use and future major depressive disorder.

Specification of which covariates were used in each of the analyses from the specification curve analysis for the association of adolescent oral contraceptive pill (OC) use with major depressive disorder (MDD) in early adulthood. A grey tick mark indicates that the variable of the current row was included in a specification. Together, the ticks add up to the 818 specifications. A single specification may have any covariate or combination of two covariates (interaction or sum) at the specific time point (i.e., baseline, 16, or 19), where OC use 25, age at menarche, and sexual orientation were included at each of the time points. Each specification included the basic adjustments for socio-economic status, ethnicity, and age (basic model). History of MDD at age 19 was included in all specifications in the subgroup of women who were not diagnosed with a MDD before age 19.

The results of the resampling tests showed that not a single one of the 1,000 median ORs were equally large or larger in the resampled specification curves than the median OR of the empirical specification curve (median OR = 1.41, $p < 0.001$). Similarly, not a single one of the 1,000 resampled specification curves had an equal or larger share of significant effects than the empirical specification curve (16.6% of the empirical models showed significant effects, $p < 0.001$) or a more extreme average Z value across all specifications than the empirical specification curve, $p < 0.001$. Qualitatively similar but quantitatively stronger effects were found when we restricted the analyses to the subgroup of women who had a

first onset of MDD after the age of 19 years. Not a single one of the 1,000 median ORs were equally large or larger in the resampled specification curves than the median OR of the empirical specification curve (median OR = 1.72, $p < 0.001$) and neither the share of significant p -values across models (34.4% of the empirical models showed significant effects, $p < 0.001$) nor the average Z value ($p < 0.001$) was equally or more extreme in any of the 1,000 resampled specification curves than in the empirical specification curve.

Discussion

The present study is the first to examine long-term associations of OC use during adolescence with MDD in adulthood in a large, prospective cohort study. Overall, we found that women who had used OCs during adolescence showed an increased risk to experience an episode of MDD up to 6 years later. This finding was most pronounced in women who had not suffered from MDD before or during adolescence.

Generally, our findings converge with several recent studies suggesting that adolescent OC use is associated with increased risk for MDD.^{124, 125} However, while all ORs were consistently above 1 in all 818 models we ran, the effect sizes varied considerably (ORs between 1.1 and 2.2), depending on the specific sample, outcome variable, and control variables selected in the model. This supports the notion that if several plausible alternative operationalizations of key measures of interest are available in a given data set, partially arbitrary choices can severely impact the apparent outcomes and conclusions of a study.³⁹⁰

Notably, the relationship between adolescent OC use and MDD risk was less strong for women with a history of MDD at age 19. Apparently, adolescent OC use does not contribute strongly to a heightened risk for a future MDD episode if this risk is already heightened by a previous episode (also known as the scar hypothesis³⁹⁵). Alternatively or additionally, it may reflect a healthy survivor bias: adolescents who suffer from MDD—whether triggered by OC use or not—might have been more likely to stop using OCs after a short period of time. Hence, these women may have been incorrectly classified as nonusers, which would have falsely caused an overestimation of the MDD risk in the nonuser group. This misclassification bias may have directed odds ratios towards the null in the total group of 725 women.

Pre-existing differences between OC users and nonusers have been proposed to explain observed relationships between OC use and MDD risk. However, in this

sample only 2.8% of the variance in the specification curve (e.g., the variance in the effect sizes) was explained by the selection of covariates. This suggests that, although factors such as socio-economic status, prior sexual activity, depressive symptoms, and OC use at age 25 might be predictors of future MDD, they did not confound the relationship between OC use and risk for MDD to a large extent in our sample of 725 women.

The present study is observational, so experimental studies will be needed to test whether the observed association is causal. Previous randomized controlled trials have overall not found a consistent link between OC use and depression (or mood, more broadly).³⁹⁶ However, to our knowledge, they all focused on short term effects and none of them restricted their sample to first-time users. If adolescent OC use indeed increases depression risk in the long term, as the current study and two other studies suggest,^{124, 125} the (often large) subset of women who had used OCs during adolescence would already be at a heightened risk for MDD and contribute to the treatment-unrelated variance. In consequence, researchers would be less likely to find differences between the placebo and the active treatment groups and therefore unintentionally underestimate the effects of OC use on depression risk.

There are several plausible reasons why adolescent OC use may be more likely to increase risk for depression than adult OC use. First, important emotion-related regions of the brain, such as the amygdala, prefrontal cortex, and hippocampus, are still maturing during adolescence and may therefore be particularly sensitive to sex hormone-related changes during this period.³⁶⁹ Neurobiological studies support this idea, as sex hormones are capable of directly influencing gene expression in the cell nuclei in the brain. Via these sex steroid receptors, sex hormones during adolescence can not only have transient but also enduring effects.^{34, 35, 37, 77, 397} Second, adolescence may be a critical time for women to develop susceptibility for depression, as the sex gap in the incidence of depression finds its origin during this period.¹²⁸ Finally, adolescent OC use has been linked to alterations in physiological stress responses both concurrently and prospectively.^{398, 399} Because aberrant stress reactivity is a known risk factor for MDD,⁴⁰⁰ enduringly altered stress reactivity through adolescent OC use may be an additional potential route through which adolescent OC use may increase later vulnerability to MDD.

A major strength of our study is that SCA was used as a transparent way to investigate the relationship between adolescent OC use and later MDD risk.

Furthermore, the availability of data from a large prospective cohort study, allowing us to control for baseline differences between OC users and nonusers. Finally, this is the first prospective study in which MDD was assessed with standardized and validated assessments. Our study also had some limitations. First, we assessed the presence but not the duration of OC use. While this ruled out recall bias, it is possible that some of the participants started and stopped using OC between two assessments and were therefore incorrectly classified as nonusers. However, if anything, this should have weakened the observed association, rather than caused spurious associations between OC use and MDD risk. Second, the specific hormonal contraceptive formulations used by women in this study was mostly unknown. Third, generalizability of results to other countries might be limited owing to differences in the acceptability of and access to contraception across societies and differences in population characteristics such as ethnicity.³⁷⁶ For example, unlike Dutch adolescents, access to no-cost contraception is limited in the United States.^{377, 378, 401} This is reflected in the user rate differences between the countries. In this Dutch study, 73.7% of women used oral contraception during their teenage years, while 45.0% of women included in a study from the United States did so.¹²⁵ Such differences have implications for group characteristics of women using OCs and, hence, possibly also with associated outcomes.

Observational data precludes causal inference. Hence, until data from long-term experimental studies confirm the association, we do not suggest limiting OC use to counterbalance this risk for MDD. Access to safe and effective methods of birth control is a basic human right, and many women benefit from OC use as it improves dysmenorrhea and reproductive autonomy.³⁸⁷ At the same time, MDD has a high impact on the functioning of the individual and the people around them,⁸ and therefore, awareness of the possible existence of increased odds for MDD among OC-using adolescents is critical to the wellbeing of women.

In this large prospective cohort study, we found evidence that women who had used OCs during adolescence were more likely to have a first MDD episode as an adult. This finding was found across a multitude of analytic choices and not primarily accounted for by (pre-existing) differences between adolescent OC users and nonusers. This pattern of results may suggest that adolescence is a vulnerable developmental period during which OC use may increase women's likelihood to develop MDD later in life.

Acknowledgements

Participating centers of TRAILS include the University Medical Center and University of Groningen, the University of Utrecht, the Radboud Medical Center Nijmegen, and the Parnassia group, all in the Netherlands. TRAILS has been financially supported by various grants from the Netherlands Organization for Scientific Research NWO (Medical Research Council program grant GB-MW 940-38-011; ZonMW Brainpower grant 100-001-004; ZonMw Risk Behavior and Dependence grants 60-60600-97-118; ZonMw Culture and Health grant 261-98-710; Social Sciences Council medium-sized investment grants GB-MaGW 480-01-006 and GB-MaGW 480-07-001; Social Sciences Council project grants GB-MaGW 452-04-314 and GB-MaGW 452-06-004; NWO large-sized investment grant 175.010.2003.005; NWO Longitudinal Survey and Panel Funding 481-08-013 and 481-11-001; NWO Vici 016.130.002 and 453-16-007/2735; NWO Gravitation 024.001.003), the Dutch Ministry of Justice (WODC), the European Science Foundation (EuroSTRESS project FP-006), the European Research Council (ERC-2017-STG-757364 and ERC-CoG-2015-681466), , Biobanking and Biomolecular Resources Research Infrastructure BBMRI-NL (CP 32), the Gratama foundation, the Jan Dekker foundation, the participating universities, and Accare Centre for Child and Adolescent Psychiatry. We are grateful to everyone who participated in this research or worked on this project to make it possible. This work was also supported by a Feodor Lynen Research Fellowship from the Alexander von Humboldt-Foundation awarded to C.A. (DEU 1187856 FLF-P), and a Canadian Institutes of Health Research Project Grant (PJT-155935) awarded to F.S.C and C.A. The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Chapter 11

Hormonal contraceptive use and depressive symptoms: systematic review and network meta-analysis of randomized clinical trials

Anouk E de Wit
Ymkje Anna de Vries
Marrit K de Boer
Celeste Scheper
Ante A Fokkema
Robert A Schoevers
Erik J Giltay

Adapted version from the article published in BJPsych Open, June 2021.

Abstract

Aim: Observational studies suggest that hormonal contraceptive use may increase depressive symptoms in women, but it is unclear whether the effect is causal. Here we aimed to quantitatively examine the evidence from randomized clinical trials for the link between hormonal contraceptive use and depressive symptoms.

Methods: We performed a systematic review and network meta-analysis of randomized clinical trials comparing women randomized to any form of a hormonal contraceptive to women randomized to any other form of a (non-)hormonal contraceptive or placebo. We searched Cochrane Central Register of Controlled Trials (CENTRAL), PubMed, Web of Science, PsycINFO, Emcare, and EMBASE from inception to May 1st, 2020. Certainty of the evidence was assessed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach. A random effect Bayesian network meta-analysis was conducted with change in depressive symptoms between baseline and 3 cycles as outcome. Our study protocol was registered in the PROSPERO database (CRD42020193304).

Results: This review identified 3492 records, of which 14 trials were eligible and 12 could be included in the network meta-analysis. These trials included 5833 participants (mean age per study range: 16.8-32.4 years) and compared ten different interventions. Very low certainty evidence suggested that hormonal contraceptive use does not cause worsening of depressive symptoms (median standardized mean difference [SMD] [95% credible interval [CrI]] compared with placebo -0.04, range SMD [95%CrI] -0.17 [-0.46 to 0.13] to 0.13 [-0.28 to 0.56]).

Conclusion: This study suggests that hormonal contraceptive use does not lead to an increase of depressive symptoms in adult women. Future studies should include first-time users to confirm the results in young women.

Introduction

There has been considerable debate whether hormonal contraceptives influence the affective state. Although cohort studies have suggested that hormonal contraceptive use could increase the risk of depressive symptoms or the prescription of psychotropic drugs,^{124, 125, 362, 387} the use of observational data precludes statements about the causality of this relationship. A meta-analysis of experimental data could confirm or reject the causality of this observation, but no such meta-analysis has been performed to date.^{126, 402}

The execution of such an analysis is complicated by the availability of numerous hormonal contraceptives that each may have different effects.¹²⁶ The best-known subgroup of hormonal contraceptives is the oral contraceptive pill (OC) which contains synthetic forms of estrogen and progestin. Other examples of hormonal contraceptives are hormonal intrauterine devices and progestin-only implants (or pills of injections) that only contain a synthetic progestin. Some hormonal contraceptives have been suggested to possess a higher risk of depressive symptoms compared to other contraceptives, including those containing higher doses of synthetic estrogens,^{124, 403} and those containing progestins with androgenic features.^{404, 405} A cyclic versus a continuous regime of use may also increase the risk of developing depressive symptoms.⁴⁰⁶ Additionally, large observational studies have shown that teenage users in particular reported more depressive symptoms than non-users^{124, 362, 387} which might suggest that younger women are more vulnerable to the adverse effects of hormonal contraceptives independent of the specific formulation used. However, it may also point to a “healthy-user effect” that occurs when women who start hormonal contraceptive use as a teenager and experience side effects are more likely to discontinue use.³⁵⁷ If true, the stronger effect in younger women is explained by the higher proportion of first-time users, but not by younger age itself.

Given the substantial disease burden of depression, as well as the importance of hormonal contraceptives for the prevention of unwanted pregnancies, as well as the treatment of dysmenorrhea and acne, it is essential to examine whether hormonal contraceptives may cause depressive symptoms. A network meta-analysis of randomized clinical trials (RCT)s makes it possible to estimate the comparative harm of different hormonal contraceptives using both direct and indirect evidence, and hence to summarize and interpret the wider picture of the evidence for their putative effect on depressive symptoms.⁴⁰⁷ Here we conducted

a systematic review and network meta-analysis on data from RCTs that assessed the effects of hormonal contraceptives compared to any other (non-)hormonal contraceptive regime on depressive symptoms in women. Our primary aim was to examine whether any hormonal contraceptive compared to each other or placebo, or OCs with higher versus those with lower doses of estrogen, hormonal contraceptives with androgenic progestins versus those with anti-androgenic progestins, and OCs with a cyclic regime versus those with a continuous regime had a more negative impact on depressive symptoms. Secondly, we aimed to investigate whether younger women are especially at risk for such potential side effects.

Methods

Search strategy and selection criteria

RCTs were eligible for inclusion if they included premenopausal women; compared hormonal contraceptive regimes with each other or with non-hormonal contraceptives (like placebo or copper intrauterine device); and evaluated depressive symptoms using a validated (self-report or observer-rated) scale. hormonal contraceptives included OCs, combined injectable contraceptives, contraceptive patch, contraceptive ring, progestin-only pills, progestin-only injectables, progestin-only implants, and levonorgestrel-releasing intrauterine devices. Studies performed in transwomen, women with an oophorectomy, or hypogonadal women were excluded. Studies that investigated hormonal contraceptives as a treatment for premenstrual dysphoric disorder or syndrome were also excluded, since they target mood as a therapeutic goal and not as an unintended side effect.

We conducted a search for English-language papers in the Cochrane Central Register of Controlled Trials, PubMed, Web of Science, PsycINFO, Emcare, and EMBASE from their inception to May 1st, 2020. Reference lists of included trials were checked to identify other potentially eligible trials or ancillary publications. Roughly, we used multiple search terms to find “randomized clinical trials” that studied any form of “hormonal contraceptive” and included assessments that measured “depressive symptoms or depression”. See *Supplement 23* for the full electronic searches.

Two reviewers (CS and AF) independently scanned all retrieved citations by title, abstract, and full text according to the prespecified inclusion criteria.

Any discrepancies were resolved through discussion or recourse to a third reviewer (AdW). Two reviewers (CS and AF) extracted data on summary estimates independently for each eligible trial using a standardized, pilot tested data extraction form. At July 6th 2020, the first data was extracted. This study is registered with PROSPERO (CRD42020193304).

Data analysis

The reviewers (CS and AF) independently collected information on methodology (level of blinding, cross-over or parallel group design); interventions (formulation, dose, frequency, regime/route of administration); participants (subjects per group, dropouts, age, first time use of hormonal contraceptives, sexual activity, comorbidity); and outcomes (tools/scales, time points reported, phase of cycle reported). Outcomes included change in depressive symptoms or onset of a new depression between baseline and 3 cycles. The cycle represented the length of one completed treatment period as it was defined by the original research and therefore the duration of the cycle may slightly differ between trials (for example because some trials used a 24-days active and 4-days hormone free interval whereas others used a 21-days active and 7-days hormone free interval). When 3-cycles data were not available, other data as close to this point as possible was used (eligible range 1-48 cycles). Preference was given to validated depressive symptom questionnaires, but when such questionnaires were not used in a trial, data from negative affect or depressive symptom subscales from scales measuring related concepts were accepted. When depressive symptoms or depression had been measured with more than one standardized rating scale, the scale with the best psychometric properties (according to validity, reliability, responsiveness and interpretation) was chosen. However, in practice we did not have to choose between scales because none of the trials reported useful data for more than one scale. If results were reported separately for different cycle phases, effect sizes were averaged across phases. Intention-to-treat data was used whenever possible. Reviewers resolved discrepancies by discussion and, when necessary, with adjunction by a third party (AdW). We contacted study authors and drug manufactures to supplement incomplete reports of the original papers. The relative effect per comparison was summarized using the standardized mean difference (SMD) adjusted for small sample sizes (Hedges' g correction), with a 95% Credible Interval (CrI).⁴⁰⁸

We performed a random-effects Bayesian network meta-analysis to explore all direct and indirect comparisons, using the ‘*gemtc*’ package in R (version 4.0.3). Trials that were not connected to the network were excluded from the network meta-analysis, and we only reported the direct effects for such comparisons. We accounted for the correlations induced by multi-arm studies by using multivariate distributions. We used uninformative, default priors for model parameters, and ran Markov-Chain Monte-Carlo (MCMC) sampling. Simulations were run for four chains with an adaptive phase of 10,000, and a sampling phase of 500,000 iterations, thinned such that every 10th iteration was retained. The convergence of the models was checked using trace plots, density plots, and the Brooks-Gelman-Rubin diagnostic. Each type of hormonal contraceptive (including regime and dose) was treated as a separate node. A network plot was drawn, with thickness of the lines between nodes based on the number of direct comparisons investigated.

To facilitate interpretation of the findings, we calculated the median SMDs of all hormonal contraceptives and presented the range of SMDs of individual formulations. Then we calculated the median SMD of the following groups: OCs with higher versus those with lower doses of estrogen, hormonal contraceptives with androgenic progestins versus those with anti-androgenic progestins, and OCs with a cyclic regime versus those with a continuous regime. Finally, we examined whether studies that included on average younger women reported larger effect sizes compared to placebo than those that included older women, using a meta-regression analysis. The mean age of participants in the studies (in years) was standardized. The beta-coefficient with 95%CrI of age together with the change in deviance information criterion (DIC) was used to determine the impact of age on the relation between hormonal contraceptive use and depressive symptoms.

The overall certainty of evidence was determined using the Confidence in Network Meta-analysis (CINeMA) online tool, based on the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) for network meta-analysis.⁴⁰⁹⁻⁴¹¹ Judgements on the certainty of evidence were made for each of the following domains: within-study bias, reporting bias, indirectness, imprecision, heterogeneity, and incoherence. The risk of within study biases for the outcome measure depressive symptoms was assessed using the modified Cochrane Risk of Bias tool 2.0.⁴¹² For full details on the GRADE assessment, we refer to *Supplement 24*.

Results

Our systematic search identified 3492 citations published between 1961 and 2020, of which fifteen articles were considered eligible (*Figure 11.1*).^{330, 346, 403-405, 413-422} One additional trial was identified through hand-searching other articles.⁴²³ The information available in five of these articles was not sufficient to complete the data collection,^{413-415, 420, 424} but authors from three trials provided us with the missing data.^{414, 415, 420} One study was excluded after data extraction, as the baseline depressive symptom scores differed by more than five standard deviations between treatment groups, suggesting a failure of randomization and rendering the results uninterpretable.⁴²² Out of the remaining 14 studies, two had no connection with the network^{403, 423} and were hence only qualitatively described in this review.

Of the 12 trials included in the network meta-analysis, 11 had a parallel design, whereas one study had a cross-over design.^{330, 404, 405, 413-421} A total of ten different interventions were examined (nine active and one placebo). Six active interventions were OCs, two were progesterone-only pills, and one was a vaginal ring. Two multi-arm trials compared three interventions, of which one additionally included two separate samples.^{414, 415} The remaining were two-arm trials.^{330, 404, 405, 413, 416-421} A total of 5833 women were included, but across comparisons, the sample size per arm ranged from 14 to 2631 participants (*Table 11.1*). The weighted average duration of follow-up was 3.6 cycles (range 1-6) and the weighted mean drop-out rate was 16.6% (range 1.3%-24.2%). Women had a weighted mean age of 27.3 years (range 16.8-32.4), and were, if studies provided such information, often sexually active (weighted mean rate 63.2%).^{414-417, 419-421} Information on previous hormonal contraceptive use was provided in only five studies. Of these studies, two did not include first-time users at all,^{418, 420} whereas the others included between 1.3% and 33.0% of first-time users.^{330, 415, 417} Most studies excluded women with a current depression or antidepressant use,^{415, 417, 418, 420, 421} or did not provide information on baseline depression presence.^{346, 404, 405, 414, 416} Only a few studies specifically included women with current depression^{330, 419} or previously hormonal contraceptive-related negative affect,^{330, 418} in order to make the sample as representative of general users as possible.

Figure 11.1. PRISMA flow diagram.

*Ethinylestradiol/drospirenone (30ug/3mg), ethinylestradiol/etonogestrel (15ug/120ug), estradiol/dienogest (3;2;2;1mg/0;2;3;0mg), desogestrel (75ug), ethinylestradiol/levonorgestrel (30ug/150ug), levonorgestrel (30ug), estradiol/nomegestrel (1.5mg/2.5mg), ethinylestradiol/levonorgestrel (20ug/100ug), ethinylestradiol/desogestrel (20ug/150ug)

The network of treatment comparisons included ten individual nodes (*Figure 11.2*). Placebo and ethinylestradiol/levonorgestrel (30ug/150ug) were the most well-connected interventions, with eight other interventions directly linked to each. The network meta-analysis showed that none of the hormonal contraceptives worsened depressive symptoms compared to placebo (median SMD -0.04, range SMD [95%CrI] -0.17 [-0.46, 0.13] to 0.13 [-0.28, 0.56]) (*Figure 11.3*).

Figure 11.2. Network of treatment comparisons for depressive symptoms.

Each node represents different a different hormonal contraceptive or placebo. The thickness of lines between nodes is proportional to the number of studies investigating the direct comparison.

Table 11.1. Characteristics of the included randomized clinical trials.

	Age mean (SD)	Intervention Formulation	n*	Comparison Formulation	n*	1 st time users n* %**	Drop-outs n (%)	Duration n cycles	Type of trial	Analysis	Depression scale
Battaglia ⁴¹	26.5 (2.4)	Ethinylestradiol/drospirenone (30ug/3mg)	22	Ethinylestradiol/etonogestrel (15ug/120ug)	21	?	3 (7.0)	6	Parallel, single blinded	Complete cases	BDI
Davis ⁴⁰	30.5 (7.4)	Estradiol/dienogest (3;2;2;1mg/0;2;3;0mg)	97	Ethinylestradiol/levonorgestrel (30ug/150ug)	103	0.0	13 (6.1)	4	Parallel, double blinded	Complete cases	PGWBI dep
Elaut ¹⁴	23.1 (4.3)	Ethinylestradiol/etonogestrel (15ug/120mg) Desogestrel (75ug)	55 55	Ethinylestradiol/desogestrel (20ug/150ug)	55	?	40 (24.2)	3	Cross-over, unblinded	Intention-to- treat	SCL-90 dep
Engman ¹⁸	24.9 (4.2)	Ethinylestradiol/levonorgestrel (30ug/150ug)	14	Placebo	15	0.0	4 (11.4)	1	Parallel, double blinded	Complete cases	MADRS
Graham ⁴⁵	32.4 (3.7)	Ethinylestradiol/levonorgestrel (30ug/150ug) Levonorgestrel (30ug)	24 22	Placebo	25	1.3	4 (5.3)	3	Parallel, double blinded	Complete cases	BDI
Graham ⁴⁵	32.2 (4.2)	Ethinylestradiol/levonorgestrel (30ug/150ug) Levonorgestrel (30ug)	24 25	Placebo	25	33.0	1 (1.3)	3	Parallel, double blinded	Complete cases	BDI
Greco ⁴³	19.8 (1.8)	Ethinylestradiol/norgestimate (35ug/0.180.220.25mg)	31	Ethinylestradiol/norgestimate (25ug/0.180.220.25mg)	29	?	12 (20.0)	3	Parallel, single blinded	Complete cases	BDI
Kelly ⁴⁴	26.5 (6.1)	Ethinylestradiol/drospirenone (30ug/3mg)	193	Ethinylestradiol/levonorgestrel (30ug/150ug)	87	?	144 (34.0)	6	Parallel, single blinded	Complete cases	MDQ neg
Legro ⁴²	27.2 (5.1)	Ethinylestradiol/norethindrone acetate (20ug/1mg continuous)	31	Ethinylestradiol/norethindrone acetate (20ug/1mg)	31	?	11 (17.7)	6	Parallel, double blinded	Intention-to- treat	MDQ neg
Lundin ³⁶	24.3 (4.2)	Estradiol/nomegestrel (1.5mg/2.5mg)	80	Placebo	88	17.8	24 (11.9)	3	Parallel, double blinded	Complete cases	MADRS
O'Connell ⁴⁶	16.8 (2.1)	Ethinylestradiol/levonorgestrel (20ug/100ug)	34	Placebo	34	?	8 (10.5)	3	Parallel, double blinded	Complete cases	CES-D
Sangthawan ⁴⁵	26.8 (4.5)	Ethinylestradiol/drospirenone (30ug/3mg)	50	Ethinylestradiol/levonorgestrel (30ug/150ug)	49	?	5 (4.8)	6	Parallel, unblinded	Complete cases	WHAQ neg
Winkler ⁴⁷	28.3 (7)	Ethinylestradiol/desogestrel (20ug/150ug)	404	Ethinylestradiol/levonorgestrel (20ug/100ug)	384	7.6	239 (23.3)	6	Parallel, unblinded	Complete cases	PGWBI dep
Wijes ⁴⁶	27.6 (7.0)	Estradiol/nomegestrel (1.5mg/2.5mg)	2631	Ethinylestradiol/drospirenone (30ug/3mg)	891	?	?	3	Parallel, unblinded	Intention-to- treat	MDQ
Zethraeus ⁴⁹	23.7 (3.6)	Ethinylestradiol/levonorgestrel (30ug/150ug)	162	Placebo	167	?	11 (3.2)	3	Parallel, double blinded	Complete cases	BDI

Abbreviations: BDI, Beck Depression Inventory; CES-D, Center for Epidemiologic Studies Depression Scale; MADRS, Montgomery-Åsberg Depression Rating Scale; MDQ neg, Menstrual Distress Questionnaire negative affect subscale; PGWBI dep, Psychological General Well-Being Index depression subscale; SCL-90 dep, Symptom Checklist 90 depression subscale; SD, standard deviation; WHAQ neg, Women's Health Assessment Questionnaire negative affect subscale. Ethinylestradiol/dienogest (3;2;2;1mg/0;2;3;0mg) is a multiphasic oral contraceptive pill that has four different dosages of hormones throughout a 4-week cycle. The dosages before the / refer to the different dosages of ethinylestradiol, and the ones after the / to those of dienogest. The Graham study included two different samples of which a was from the Philippines, and b was from Scotland. *Number of women who were available for the analysis (so excluding the dropouts) in case of complete-cases analyses. **Percentage of the baseline sample that never had experience with hormonal contraceptive use before participating in the trial. As the majority of studies ran complete cases analyses, it's unknown what percentage of women included in the meta-analyses had previous user-experience.

Table 11.2. Comparisons for effects of hormonal contraceptives on depressive symptoms.

Desogestrel (75ug)	Ethinylestradiol /dienogest (3;2;2;1mg/0;2;3;0mg)	Ethinylestradiol /etonogestrel (15ug/120ug)	Ethinylestradiol /desogestrel (20ug/150ug)	Ethinylestradiol /levonorgestrel (20ug/100ug)	Ethinylestradiol /drospirenone (30ug/3mg)	Ethinylestradiol /levonorgestrel (30ug/150ug)	Placebo
0.09 ^a (-0.88, 1.07)	-0.04 ^a (-0.71, 0.72)	-0.02 ^a (-0.76, 0.66)	-0.08 ^a (-0.61, 0.46)	-0.01 ^a (-0.47, 0.46)	0.12 ^a (-0.52, 0.74)	-0.30 ^a (-0.71, 0.10)	0.13 ^a (-0.28, 0.56)
0.06 ^a (-0.75, 0.94)	-0.05 ^a (-0.91, 0.80)	-0.10 ^a (-0.85, 0.60)	-0.09 ^a (-0.71, 0.55)	0.11 ^a (-0.56, 0.76)	0.10 ^a (-0.22, 0.45)	-0.19 ^a (-0.70, 0.31)	-0.17 ^a (-0.46, 0.13)
0.04 ^a (-0.54, 0.62)	-0.13 ^a (-1.00, 0.74)	-0.11 ^a (-0.81, 0.55)	0.03 ^a (-0.59, 0.64)	0.21 ^a (-0.46, 0.88)	0.22 ^a (-0.40, 0.84)	-0.08 ^a (-0.79, 0.61)	0.05 ^a (-0.52, 0.63)
-0.03 ^a (-0.61, 0.55)	-0.14 ^a (-0.96, 0.69)	0.03 ^a (-0.44, 0.37)	0.13 ^a (-0.52, 0.80)	0.12 ^a (-0.37, 0.53)	0.12 ^a (-0.52, 0.74)	-0.17 ^a (-0.92, 0.58)	-0.04 ^a (-0.68, 0.63)
-0.04 ^a (-0.74, 0.66)	-0.02 ^a (-0.68, 0.62)	0.12 ^a (-0.37, 0.53)	-0.17 ^a (-0.92, 0.58)	0.21 ^a (-0.46, 0.88)	0.22 ^a (-0.40, 0.84)	-0.08 ^a (-0.79, 0.61)	0.05 ^a (-0.52, 0.63)
0.07 ^a (-0.71, 0.85)	0.08 ^a (-0.47, 0.63)	-0.22 ^a (-0.91, 0.46)	-0.18 ^a (-0.79, 0.35)	-0.09 ^a (-0.85, 0.65)	-0.19 ^a (-0.70, 0.31)	-0.30 ^a (-0.71, 0.10)	-0.17 ^a (-0.46, 0.13)
0.18 ^a (-0.62, 0.98)	-0.22 ^a (-0.91, 0.46)	-0.18 ^a (-0.79, 0.35)	-0.17 ^a (-0.92, 0.58)	-0.09 ^a (-0.85, 0.65)	-0.19 ^a (-0.70, 0.31)	-0.30 ^a (-0.71, 0.10)	-0.17 ^a (-0.46, 0.13)
-0.12 ^a (-1.00, 0.74)	-0.22 ^a (-0.91, 0.46)	-0.18 ^a (-0.79, 0.35)	-0.17 ^a (-0.92, 0.58)	-0.09 ^a (-0.85, 0.65)	-0.19 ^a (-0.70, 0.31)	-0.30 ^a (-0.71, 0.10)	-0.17 ^a (-0.46, 0.13)
0.01 ^a (-0.77, 0.81)	-0.09 ^a (-0.70, 0.54)	-0.05 ^a (-0.51, 0.35)	-0.04 ^a (-0.68, 0.63)	0.04 ^a (-0.60, 0.69)	-0.07 ^a (-0.44, 0.34)	-0.17 ^a (-0.46, 0.13)	0.13 ^a (-0.28, 0.56)

Data are standardized mean differences (SMD) with 95% Credible Intervals in the column-defining intervention compared with the row-defining intervention. Higher SMD values correspond with fewer depressive symptoms in the column-defining hormonal contraceptive. None of the results was significant and all evidence was of low certainty (marked with ^a). Ethinylestradiol/dienogest (3;2;2;1 mg/0;2;3;0mg) is a multiphasic oral contraceptive pill that has four different dosages of hormones throughout a 4-week cycle. The dosages before the / refer to the different dosages of ethinylestradiol, and the ones after the / to those of dienogest.

Figure 11.3. Network forest plot of results compared to placebo.

The size of the dots are proportional to the precision of the estimate ($1 / \text{standard error}$). Ethinylestradiol/dienogest (3;2;2;1mg/0;2;3;0mg) is a multiphasic oral contraceptive pill that has four different dosages of hormones throughout a 4-week cycle. The dosages before the “slash” refer to the different dosages of ethinylestradiol, and the ones after the “slash” to those of dienogest.

None of the specific formulations of hormonal contraceptives had a significantly more negative effect on depressive symptoms than any other specific formulation, suggesting that there was no group of hormonal contraceptives causing more depressive symptoms than another group. See *Table 11.2* for all the comparisons of effects of hormonal contraceptives on depressive symptoms. First, hormonal contraceptives with androgenic progestins (e.g. levonorgestrel and desogestrel) did not have a stronger association with depressive symptoms than those with anti-androgenic progestins (i.e., dienogest, nomegestrol, or drospirenone). For example, the median SMD for OCs containing levonorgestrel or desogestrel compared to those containing anti-androgenic progestins: -0.09 and 0.11, respectively (*Figure 11.4*). Second, OCs with higher doses of estrogen did not induce more depressive symptoms compared to those with lower doses

(Figure 11.5; median SMD for formulations with 30ug compared to those with 20ug ethinylestradiol: -0.16, respectively). This was also supported by a study that lacked a connection with the network.⁴⁰³ In that study, women who used a OC with a lower dose of estrogen (ethinylestradiol/norgestimate [25ug/180ug/215ug/250ug]) had on average fewer depressive symptoms than women using the same contraceptive with a higher dose of estrogen (35ug/180ug), but this was not statistically significant (mean difference in BDI score between low and high estrogen group = -2.26, $p = 0.09$). As the doses of estradiol and ethinylestradiol are hard to compare, and one the two studies that studied estradiol used an OC with a 4-phasic regime which makes it even harder to estimate the average level of estradiol to use, we limited our analyses to the direct and indirect effects of OCs including ethinylestradiol. Third, we aimed to compare continuous and intermittent regimens, but no such comparison was included in the network. However, another study that could not be included in the network showed that women who used ethinylestradiol/norgestimate (20ug/1mg) continuously did not have fewer depressive symptoms after six months of use than those who took it cyclically (mean difference in negative affect on Moos Menstrual Distress Questionnaire T-scores = -2.8, $p = 0.32$).⁴²³

The meta-regression analyses showed that on average, every year decrease in mean age of the women investigated in a trial was associated with 0.028 lower SMD of the hormonal contraceptive group compared to the placebo group (lower SMD meaning more depressive symptoms). However, this was not significant (standardized beta-coefficient for age [95%CrI] 0.238 [-0.232, 0.708], Δ DIC -3.6). This association weakened after exclusion of the study that solely included women with previous hormonal contraceptive-related negative affect (beta-coefficient for age [95%CrI] 0.193 [-0.254, 0.642], Δ DIC -4.0).⁴¹⁸

The certainty of evidence (GRADE) for all comparisons was judged to be “very low” (*Supplement 24 Table 1*). Most of these judgements were caused by within-study bias (*Supplement 24.1*), and by indirectness (*Supplement 24.3*) and imprecision (*Supplement 24.4*) of the effect sizes. The judgements made for each of the domains are described in detail in *Supplement 24*.

Two androgenic progestins included in the network (levonorgestrel and desogestrel) were compared to the anti-androgenic progestins included in the network (dienogest, drospirenone, norgestrel). Synthetic progestins, especially the older ones, do not only have affinity to the progesterone receptor but can also have androgenic or oestrogenic actions. From the synthetic progestins here, levonorgestrel has the strongest affinity to the androgen-receptor, followed by desogestrel.⁴²⁵ Newer progestins have higher progesterone potency and additional effects, such as anti-androgenic activity.⁴²⁵ The size of the dots are proportional to the precision of the estimate (1 / standard error). Ethinylestradiol/dienogest (3;2;1 mg/0;2;3;0mg) is a multiphasic oral contraceptive pill that has four different dosages of hormones throughout a 4-week cycle. The dosages before the / refer to the different dosages of ethinylestradiol, and the ones after the / to those of dienogest.

Figure 11.5. Network forest plot of results of oral contraceptive pills with a 30ug ethinylestradiol versus those with 20ug ethinylestradiol.

The size of the SMD dots are proportional to the precision of the estimate ($1 / \text{standard error}$).

Discussion

This systematic review examined 14 RCTs, of which 12 were included in a network meta-analysis to assess the effects of hormonal contraceptive use on depressive symptoms. Data from 5833 women were available and nine different hormonal contraceptives were compared to each other and placebo. The results suggest that, with a very low certainty of evidence, none of the investigated hormonal contraceptives caused more depressive symptoms compared to placebo.

Here we have provided quantitative evidence based on RCTs that hormonal contraceptives appear to have a limited effect on depressive symptoms. This finding is largely in line with the suggestions from previous qualitative reviews.^{126, 127, 363, 402} In contrast to those reviews, however, we could not confirm the idea that the dose of ethinylestradiol or the type of progestin may influence the effect of hormonal contraceptives on depressive symptoms. Although the data regarding the safety of 20ug versus 30ug ethinylestradiol OCs suggested less risk with 20ug formulas, this was not strong enough to endorse higher safety with 20ug pills. We cannot rule out that much higher doses of ethinylestradiol (e.g. 50ug) adversely affect the affective state, as none of the included trials studied such high doses, but these formulations are not recommended for contraception. Formulations containing androgenic progestins also did not cause more depressive symptoms than those with anti-androgenic formulations.

The suggestion that hormonal contraceptives have no effect on depressive symptoms stands in direct opposition to results from several observational studies showing that hormonal contraceptive users are at increased risk for depression

or report more depressive symptoms.^{124, 125, 362, 387} Residual confounding in observational studies could be one compelling explanation for the contradiction in findings. Indeed, one observational study showed that the association between OC use and depressive symptoms could at least be partly explained by pre-existing differences between oral contraceptive users and non-users which included differences in number of women being sexually active, number of stressful experiences, and number of women having acne or menstrual related pain.³⁸⁷ However, the absence of a significant effect in the current meta-analysis may also have arisen due to limitations of the currently available randomized trials. As some observational studies showed that an increased prevalence of depressive symptoms related to hormonal contraceptives was most pronounced in teenage women,^{124, 125, 362, 387} this may suggest that the absence of an effect in the, on average, adult women in our study could be due to the 'healthy-user effect'.³⁵⁷ We could not confirm that this bias has occurred in our analysis as effects of hormonal contraceptives on depressive symptoms were not stronger in studies that included on average younger women. However, age may be an imperfect proxy for first-time use, especially as there was only one trial that included teenage women, whereas the majority of trials included women in their mid-twenties. Hence, we cannot rule out that first-time use is a risk factor for experiencing depressive symptoms with hormonal contraceptive use.

Therefore, in order to provide a definite answer to the question whether hormonal contraceptives adversely affect depressive symptoms, more studies are needed that also include first-time users to ensure that findings are not biased by a "healthy-user effect". Women who have had experienced side effects of hormonal contraceptives during their first-time use are less likely to consent to participate in a randomized trial of hormonal contraceptives, making it unlikely that these side effects are going to be observed. As the potential randomization to a less reliable non-hormonal contraceptive such as a condom in young women would likely be ethically unacceptable given the significantly higher failure rate of condoms compared to hormonal contraceptives,⁴²⁶ such trials may be limited to head-to-head studies of hormonal contraceptives to yield those formulations with the smallest adverse effect on the affective state, if any. The availability of such data would not only improve the generalizability of the findings to this first-time user group, but also provide more precise estimates. This is essential as the reported effect size in this study was still of very low certainty of evidence according to the

GRADE criteria for network meta-analysis, mainly due to the limited number of trials per comparison and the limited sample sizes in each of the trials.

This study has some major strengths. The project is the first network meta-analysis on the effect of hormonal contraceptives on depressive symptoms and extends previous work that qualitatively reviewed observational and experimental studies.^{126, 127, 402} Moreover, it is substantially more comprehensive than previous reviews as it includes seven trials that were not incorporated previously,^{330, 414, 417-421} and includes data from three trials that have not been published before.^{414, 415, 420} Nevertheless, this study has also some limitations. Only one or a few trials were available for each comparison and many trials had a relatively small sample size, which is probably insufficient to detect small but clinically important effects on depressive symptoms for specific formulations. This limitation is reflected in our ratings of the certainty of evidence for each comparison. Also, the quality of our analysis is limited by inherent limitations of individual included trials. First, very few trials provided sufficient data on important moderators, making it difficult to assess whether the transitivity assumption for network meta-analysis held. Only a few studies included women that had previous negative experiences with hormonal contraceptive use which may have created a sample that has been less vulnerable to depressive symptoms than the average population of hormonal contraceptive users. Second, only one trial used an observer-rated scale to assess depressive symptoms whereas the other trials used self-reported questionnaires to assess this or a related concept (negative affect). While self-reported rating scales are efficient in terms of time and costs, and provide valuable insights in the subjectively experienced severity of symptoms, only observer-rated structured diagnostic assessment procedures can assess valid diagnoses of major depression disorders.

With our network meta-analysis, we provide quantitative evidence based on experimental data that hormonal contraceptive use is relatively safe regarding the effect on depressive symptoms, especially in adult women. However, few trials were available for each formulation of hormonal contraceptives, and certainty of evidence for each comparison was rated as very low. Only a limited number of first-time users or women with previously negative experiences with hormonal contraceptive use were included, which limits the generalizability of the findings to these groups. This implies that awareness of the possible existence of mood problems among hormonal contraceptive-users remains critical to the health and wellbeing of these women. Not only has depression has a high impact on

the functioning of the individual and the people around them,⁸ it also reduces hormonal contraceptive use adherence.³⁸⁰ However, by no means, we suggest to limit hormonal contraceptive use to counterbalance this potential, but not proven, risk for depressive symptoms. Access to safe and effective methods of birth control is a basic human right, and many women benefit from hormonal contraceptive use as it improves dysmenorrhea and reproductive autonomy.³⁸⁷ Moreover, hormonal contraceptive use is much safer than pregnancy and associated postpartum depression risks.³⁸² More RCTs investigating the safety of hormonal contraceptive use that include depressive symptoms as one of their outcomes are warranted to help women and their doctors make better informed choices when deciding among possible methods of contraception.

Acknowledgements

We thank Jan W. Schoones, MA, Information Specialist from Leiden University Medical Centre for assistance in developing the search strategy. We also are grateful for additional data that were provided to us by Cynthia A. Graham, professor at the University of Southampton, Els Elaut, professor at Gent University, and Bayer B.V..

There was no fundings for this project. The authors have nothing to disclose.

Data accessibility

The full dataset and the script for analyzing the data is freely available from: <https://osf.io/2ehyw/>.

Chapter 12

Efficacy of oral contraceptive pills for depressive symptoms and overall symptomatology in premenstrual syndrome: pairwise and network meta-analysis of randomized trials

Anouk E de Wit
Ymkje Anna de Vries
Marrit K de Boer
Celeste Scheper
Ante A Fokkema
Catharina AH Janssen
Erik J Giltay
Robert A Schoevers

Adapted version from the article published in
American Journal of Obstetrics & Gynecology, July 2021.

Abstract

Aim: Combined oral contraceptives are often considered a treatment option for women with premenstrual syndrome (PMS) or premenstrual dysphoric disorder (PMDD) also seeking contraception, but evidence for this treatment is scarce. We aimed to determine 1) the level of evidence for the efficacy of combined oral contraceptives in managing premenstrual depressive symptoms and overall premenstrual symptomatology, and 2) the comparative efficacy of combined oral contraceptives (PROSPERO registration number CRD42020205510).

Methods: We searched Cochrane Central Register of Controlled Trials (CENTRAL), PubMed, Web of Science, PsycINFO, Emcare, and EMBASE from inception to June 3rd, 2021. A protocol for this review was registered at PROSPERO (CRD42020205510). All randomized clinical trials that evaluated efficacy of OCs in women with PMS or PMDD were considered eligible for inclusion in the present meta-analysis. A random effect Bayesian pairwise and network meta-analysis was conducted with change in premenstrual depressive symptoms and overall premenstrual symptomatology between baseline and three cycles as outcome. Certainty of the evidence was assessed using the Grading of Recommendations Assessment, Development and Evaluation approach.

Results: Of 3493 records, nine eligible trials were included that studied 1205 women with PMS or PMDD (mean age per study range: 24.6-36.5 years). The pairwise meta-analysis revealed that OCs were more efficacious than placebo in treating overall premenstrual symptomatology (standardized mean difference SMD [95%CrI], 0.41 [0.17, 0.67]), but not premenstrual depressive symptoms specifically (SMD [95%CrI], 0.22 [-0.06, 0.47]). However, none of the OCs were more effective than each other in reducing premenstrual depressive symptoms and overall premenstrual symptomatology.

Conclusion: OCs may improve overall premenstrual symptomatology in women with PMS or PMDD, but not premenstrual depressive symptoms. There is no evidence for one OC being more efficacious than any other.

Introduction

Symptoms that occur in the luteal phase of the menstrual cycle and resolve after the onset of menstruation characterize the premenstrual syndrome (PMS) and its more severe variant premenstrual dysphoric disorder (PMDD).⁴²⁷⁻⁴³⁰ A broad range of physical and affective symptoms may be present, but in women with PMDD, severe affective symptoms such as depression are the primary complaint.⁴²⁷ Of menstruating women, up to 12.2% were affected by PMS and up to 5.3% by PMDD. In both disorders, cyclic changes in the production of progesterone are thought to trigger symptoms.⁴³¹⁻⁴³³ As combined oral contraceptives contain a synthetic estradiol and progestin that together suppress the hypothalamic-pituitary-ovarian axis, treatment with these drugs is often considered for women with PMS or PMDD who also seek contraception. However, the evidence for their efficacy remains scarce.

Currently, ethinylestradiol/drospirenone (20ug/3mg) in a 24-day regimen is the only combined oral contraceptive that has been approved by the US Food and Drug Association for the treatment of PMDD.⁴³⁴ This approval was based on the results of two randomized clinical trials in women with PMDD reporting that use of this formulation reduced overall premenstrual symptomatology compared to placebo.^{435, 436} However, other trials showed that the same ethinylestradiol/drospirenone combination in a different treatment regimen in women with PMS,⁴³⁷ and the same combination but with a slightly higher ethinylestradiol dosing (30ug) in women with PMDD, were not effective.⁴³⁸ Two trials investigating the continuous use of ethinylestradiol/levonorgestrel (20ug/90ug) in women with PMDD also did not show a consistent beneficial effect of on premenstrual depressive symptoms and overall premenstrual symptomatology.^{439, 440} Altogether, these inconsistent findings question the efficacy of combined oral contraceptives for treating premenstrual symptomatology in women with PMS and PMDD.

Here we sought to aggregate the results from clinical trials that randomized women with PMS or PMDD to combined oral contraceptives compared to any other hormonal or non-hormonal contraceptive treatment regimen or to placebo in order to examine the efficacy of combined oral contraceptives. We specifically aimed to determine 1) the efficacy of combined oral contraceptives for managing premenstrual depressive symptoms and overall premenstrual symptomatology in PMS and PMDD, and 2) whether some combined oral contraceptives are more efficacious than others.

Methods

Per recommended systematic review methodology this review was developed and registered a priori with PROSPERO (CRD42020205510).⁴⁴¹ In the original PROSPERO protocol, the primary aim was to conduct a network meta-analysis and only if this was not possible, to conduct a pairwise meta-analysis. Because all but one included randomized clinical trials had a placebo-controlled design, making it possible to conduct a (potentially better-powered) pairwise meta-analysis examining combined oral contraceptives as a group, the protocol was modified to include a pairwise meta-analysis as the primary analysis. We originally also aimed to investigate whether studies with a greater proportion of women with PMDD reported larger effect sizes compared to placebo than those that included a smaller proportion of women with PMDD using meta regression analyses. Since fewer than 10 studies could be included in this study, such analyses, however, were not possible.

Eligibility criteria

Inclusion criteria for studies were as follows (within a PICO-S framework):

1. **Population:** Premenopausal women diagnosed with PMS or PMDD. Diagnosis of PMDD was preferably made according to Diagnostic and Statistical Manual of Mental Disorders (DSM) IV or 5 criteria.^{427, 442} Other assessments of PMDD through the use of validated questionnaires were accepted if the criteria used required a significant rise of at least one emotional symptom (including but not limited to: depressed mood, anxiety or tension, affective lability, and anger or irritability) plus four additional symptoms during the luteal or premenstrual phase that subsided within the first three days of the menses. For PMS, the same criteria were used except that no additional symptoms were required for the diagnosis.
2. **Intervention:** A specific combined oral contraceptive regimen
3. **Comparison:** Another combined oral contraceptive regimen or (non-) hormonal contraceptives (e.g. hormonal or copper intrauterine device) or placebo
4. **Outcome:** Depressive symptoms using a validated (self-report or observer-rated) scale
5. **Study design:** Randomized controlled trials

Information sources, search strategy, and study selection

We conducted a search for studies published in English in the Cochrane Central Register of Controlled Trials, PubMed, Web of Science, PsycINFO, Emcare, and EMBASE from their inception to June 3rd, 2021. Reference lists of included trials and of related reviews (searched for by a separate search) were checked to identify other potentially eligible trials or ancillary publications. See *Supplement 23* for the full searches in Pubmed.

Two reviewers (CS and AF) independently scanned all retrieved citations by title, abstract, and full text according to the prespecified inclusion criteria. Any discrepancies were resolved through discussion or recourse to a third reviewer (AdW). Two reviewers (CS and AF) extracted data on summary estimates independently for each eligible trial using a standardized, pilot tested data extraction form.

Data extraction

The reviewers (CS and AF) independently collected information on methodology (level of blinding, cross-over or parallel group design), interventions (formulation, dose, frequency, regimen/route of administration), participants (number of subjects per group, number of dropouts, mean age, and proportion of participants with PMDD, with comorbidity, or who reported being sexually active), and outcomes (i.e., tools/scales, reported time points, phase of cycle reported). Reviewers resolved discrepancies by discussion and, when necessary, through adjudication by a third party (AdW). We contacted study authors and drug manufacturers to supplement incomplete data regarding outcomes. The relative effect per comparison was summarized using the standardized mean difference (SMD) adjusted for small sample sizes (Hedges' g correction), with a 95% Credible Interval (CrI).⁴⁰⁸ A CrI is an interval within which an unobserved parameter value falls with a particular probability given the evidence provided by the observed data. Using a standardized mean difference as an outcome, intervals containing zero suggest that the data are compatible with no effect or no difference between groups, as with the analogous 95% confidence intervals in frequentist statistics.

The primary outcome was change in premenstrual depressive symptoms or new-onset depression between baseline and three cycles. The secondary outcome was the change in overall premenstrual symptomatology measured over the same period. For the primary outcome, preference was given to validated depressive

symptom severity questionnaires. However, when such questionnaires were not used in a trial, data from negative affect or depressive symptom subscales from scales measuring related concepts were accepted. Similarly, overall premenstrual symptomatology measures were accepted when they had been assessed with validated symptom questionnaires for premenstrual complaints. This questionnaire had to assess both the affective and physical domain of PMS and provide a score that reflects the severity of premenstrual complaints. When symptoms of the primary or secondary outcome had been measured with more than one standardized rating scale, the scale with the best psychometric properties was chosen. If results were not separately reported for the premenstrual phase, effect sizes across phases were taken. Similarly, when 3-cycles data were not available, other data as close to this point as possible

Assessment of risk of bias

We assessed the studies' risk of bias using the modified Cochrane Risk of Bias tool 2.0.⁴¹² Additionally, we assessed the certainty of evidence contributing to network estimates of the primary outcome with the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework using the Confidence in Network Meta-analysis (CINeMA) tool.⁴⁰⁹⁻⁴¹¹ Judgements on the certainty of evidence were made for each of the following domains: within-study bias, reporting bias, indirectness, imprecision, heterogeneity, and incoherence. For full details on the GRADE assessment, we refer to *Supplement 25*.

Data synthesis

First, we performed a random-effects Bayesian pairwise meta-analysis with uninformative priors and Markov-Chain Monte-Carlo (MCMC) sampling, to determine the efficacy of OCs (as a group) compared to placebo in managing premenstrual depressive symptoms and overall premenstrual symptomatology. The '*brms*' package in R (version 4.0.3) was used to build the model. Simulations were run for four chains with an adaptive phase of 10,000 iterations and a sampling phase of 500,000 iterations, thinned such that every 10th iteration was retained. Model convergence was assured by checking whether the density of the replications were roughly similar to the ones of the observed data and checking the Potential Scale Reduction Factor.

Second, a random-effects Bayesian network meta-analysis with uninformative priors was used to determine the comparative efficacy of OCs in managing premenstrual depressive symptoms and overall premenstrual symptomatology. The 'gemtc' package in R (version 4.0.3) was used to build the model with the same method of sampling and running simulations as for the pairwise meta-analysis. The model accounted for correlations induced by multi-arm studies by using multivariate distributions. Model convergence was checked using trace plots, density plots, and the Brooks-Gelman-Rubin diagnostic. Each specific OC (including regimen and dose) was treated as a separate node. A network plot was drawn, with thickness of the lines between nodes based on the number of direct comparisons investigated.

Results

Our systematic search identified 3644 citations published between 1961 and June 3, 2021. One additional trial was identified through hand-searching other reports. After removal of duplicates, 3581 articles were screened based on title and abstract, of which 98 were screened based on the full-text. From these articles, ten studies were considered eligible,^{330, 435-440, 443-445} but one study was excluded due to insufficient data and no author response to a request for additional data.⁴⁴⁵ Nine studies were included in the network meta-analysis (eight placebo-controlled,^{330, 435-440, 443} one head-to head,⁴⁴⁴ see *Figure 12.1*) and eight placebo-controlled studies were included in the pairwise meta-analysis.^{330, 435-440, 443}

The nine included studies examined nine different interventions (eight different combined oral contraceptives and one placebo) among 1205 women (weighted mean age of 32.3 years [range 24.6-36.5]). Across comparisons, the sample size per arm ranged from 16 to 231 participants (*Table 12.1*). All studies evaluated the effect of the intervention premenstrually (in the final week before the onset of menses) after 3 months of treatment. During this follow-up period, drop-out was generally high with a weighted mean drop-out rate of 26.3% (range 0.0%-40.2%). Most studies included women with PMDD,⁴³⁵⁻⁴⁴⁰ but four studies predominantly included women with PMS.^{330, 437, 443, 444}

Figure 12.1. PRISMA flow diagram.

*Ethinylestradiol/drospirenone (20ug/3mg; continuous), ethinylestradiol/drospirenone (20ug/3mg; 21/7), ethinylestradiol/drospirenone (30ug/3mg; 21/7), ethinylestradiol/drospirenone (20ug/3mg; 21/7), ethinylestradiol/drospirenone (20ug/3mg; 24/4), ethinylestradiol/norethindrone (35ug/0.5;1;0.5;0mg; 21/7), estradiol/nomegestrol (1.5mg/2.5mg; 24/4), ethinylestradiol/levonorgestrel (20ug/90ug; continuous), ethinylestradiol/desogestrel (20ug/150ug; 24/4)

Only one trial used a validated depression-specific questionnaire to measure premenstrual depressive symptoms (Beck Depression Inventory).⁴³⁸ The other trials used the depressive symptom subscale of the Daily Record of Severity of Problems (DRSP),^{330, 435-437, 439, 440} or the negative affect subscale of the Daily Rating Form⁴⁴³ or Women's Health Assessment Questionnaire (WHAQ).⁴⁴⁴ Seven trials also provided data on the efficacy of interventions on overall premenstrual symptomatology. Five trials assessed overall premenstrual symptomatology with the DRSP total score.^{330, 435, 436, 439, 440} The two other trials used the WHAQ,⁴⁴⁴ and the Calendar of Premenstrual Experiences,⁴³⁸ respectively.

Risk of bias of included studies

Four (44.4%) of nine trials examined were rated as high risk of bias,^{330, 437, 438, 443} and the other trials were rated as moderate (53.6%) (*Supplement 25.1*).^{435, 436, 439, 440, 444} The certainty of evidence (GRADE) for all comparisons examined in the network meta-analysis was judged to be “very low” (*Supplement 25 Table 1*). Within-study bias (*Supplement 25.1*) and imprecision of the effect sizes (*Supplement 25.4*) caused most of these judgements. The judgements made for each of the domains are described in detail in *Supplement 25*.

Premenstrual depressive symptoms

Based on the results of eight placebo-controlled trials,^{330, 435-440, 443, 444} the pairwise meta-analysis revealed no evidence that OCs were effective for premenstrual depressive symptoms compared to placebo (SMD [95%CrI], 0.22 [-0.06, 0.47], upper panel *Figure 12.2*).

In the network meta-analysis one additional head-to-head trial could be included.⁴⁴⁴ This resulted in a network of nine individual nodes where most interventions were linked only with placebo and the majority of comparisons were only studied once (*Figure 12.3*). Therefore, comparisons between OCs and placebo were predominantly based on direct evidence, while comparisons among OCs were predominantly based on indirect evidence. In line with the pairwise meta-analysis findings, the network meta-analysis showed that none of the OCs (including the specific regimens) were more effective than placebo or each other (comparative efficacy shown in lower panel *Table 12.2* and visualized in the upper panel of *Figure 12.4*).

Figure 12.2. Forest plot of effects of oral contraceptive pills versus placebo on premenstrual depressive symptoms and overall premenstrual symptomatology.

Abbreviations: CrI, Credible Interval; OC, oral contraceptive pill; SMD, standardized mean difference. The effect sizes for each study are centered around the pooled effect size. Importantly, the effect sizes displayed are the estimates of the “true” effect size of study based on the Bayesian model (not the effect sizes reported in the original studies). The Freeman 2012 study describes four studies including two randomized clinical trials (the multinational and the North American trial). As the North American trial is also described in Halbreich 2012, we refer to the multinational trial when using Freeman 2012.

Figure 12.3. Network of treatment comparisons for premenstrual depressive symptoms.

Each node represents a different oral contraceptive pill or placebo. Thickness of lines between nodes is proportional to the number of randomly assigned patients contributing to direct comparisons. Ethinylestradiol/norethindrone (35ug/0.5;1;0.5;0mg) is a multiphasic oral contraceptive pill that has four different dosages of norethindrone throughout a 4-week cycle. The dosages before the 'slash' refer to the different dosages of the synthetic estrogen, and the ones after the 'slash' to those of the synthetic progestin.

Overall premenstrual symptomatology

A total of six placebo-controlled trials were included in the pairwise meta-analysis on the efficacy of OCs in reducing overall premenstrual symptomatology.^{330, 435-440,}

⁴⁴³ In contrast to the findings on premenstrual depressive symptoms, OCs were found to be moderately effective compared to placebo (SMD [95%CrI], 0.41 [0.17, 0.67], lower panel *Figure 12.2*).

However, when estimating the comparative efficacy of OCs in seven trials with the network meta-analysis,^{330, 435, 436, 438-440, 444} none of the OCs was more effective in improving overall premenstrual depressive symptomatology compared to placebo or compared to each other (comparative efficacy shown in upper panel *Table 12.2* and visualized in the lower panel of *Figure 12.4*).

Figure 12.4. Forest plot of treatment efficacy of oral contraceptive pill compared to placebo for premenstrual depressive and overall premenstrual symptomatology.

Abbreviations: CrI, Credible Interval; OC, oral contraceptive pill; SMD, Standardized mean difference. The size of the SMD dots are proportional to the precision of the estimate (1 / standard error). Ethinylestradiol/norethindrone (35ug/0.5;1;0.5;0mg) is a multiphasic oral contraceptive pill that has four different dosages of norethindrone throughout a 4-week cycle. The dosages before the / refer to the different dosages of the synthetic estrogen, and the ones after the / to those of the synthetic progestin.

Table 12.1. Characteristics of the included randomized clinical trials on the efficacy of oral contraceptive pills for women with premenstrual syndrome.

	Age mean (SD)	PMS/PMDD	Intervention Formulation	n*	Control Formulation	n*	Drop-outs %	RCT type	Analysis	Outcome scale Depression severity	PMS severity
Eisenlohr-Moul ⁴³⁷	32.7 (8.3)	64% PMS 36% PMDD	Ethinylestradiol/drospirenone (20ug/3mg, continuous) Ethinylestradiol/drospirenone (20ug/3mg, 21/7)	16 17	Placebo	22	24.0	Parallel, double blind	Complete cases	DRSP q1a/b/c	-
Freeman ⁴³⁸	31.0 (5.6)	PMDD	Ethinylestradiol/drospirenone (30ug/3mg, 21/7)	21	Placebo	28	40.2	Parallel, double blind	Complete cases	BDI	COPE
Freeman ^{**439}	36.5 (7.8)	PMDD	Ethinylestradiol/levonorgestrel (20ug/90ug, continuous)	34	Placebo	46	20.0	Parallel, double blind	Complete cases	DRSP q1a/b/ c+9+10	DRSP
Graham ⁴⁴³	29.5 (5.0)	PMS	Ethinylestradiol/norethindrone (35ug/0.5;1;0.5;0mg, 21/7)	20	Placebo	25	28.0	Parallel, double blind	Complete cases	DRF	-
Halbreich ⁴⁴⁰	36.1 (6.4)	PMDD	Ethinylestradiol/levonorgestrel (20ug/90ug, continuous)	133	Placebo	132	33.9	Parallel, double blind	Complete cases	DRSP q1a/b/ c+9+10	DRSP
Lundin ³³⁰	24.6 (4.1)	69% PMS 31% PMDD	Estradiol/norgestrel (1.5mg/2.5mg, 24/4)	24	Placebo	20	10.2	Parallel, double blind	Complete cases	DRSP q1a/b/c	DRSP
Pearlstein ⁴³⁶	31.5 (5.5)	PMDD	Ethinylestradiol/drospirenone (20ug/3mg, 24/4)	64	Placebo	64	60.9	Cross- over, double blind	Intention- to-treat	DRSP q1a/b/c	DRSP
Wichianpitaya ⁴⁴⁴	27.3 (5.7)	PMS	Ethinylestradiol/desogestrel (20ug/150ug, 24/4)	45	Ethinylestradiol/drospirenone (20ug/3mg, 24/4)	45	0.0	Head- to-head, unblind	Intention- to-treat	WHAQ	WHAQ
Yonkers ⁴³⁵	31.5 (5.7)	PMDD	Ethinylestradiol/drospirenone (20ug/3mg, 24/4)	231	Placebo	218	27.1	Parallel, double blind	Intention- to-treat	DRSP q1a/b/c	DRSP

BDI, Beck Depression Inventory; COPE, Calendar of Premenstrual Experiences; DRF, daily rating form; DRSP, daily record of severity of problems; PMDD, premenstrual dysphoric disorder; PMS, premenstrual syndrome; WHAQ, Women's Health Assessment Questionnaire negative affect subscale. Ethinylestradiol/norethindrone (35ug/0.5;1;0.5;0mg) is a multiphasic oral contraceptive pill that has four different dosages of norethindrone throughout a 4-week cycle. The dosages before the / refer to the different dosages of the synthetic estrogen, and the ones after the / to those of the synthetic progestin. *Number of women who were available for the analysis (so excluding the dropouts in case of complete-cases analyses). **This reference describes four studies including two randomized clinical trials (the multinational and the North American trial). As the North American trial is also described in another reference included in this meta-analysis,⁴⁴⁰ we refer to the multinational trial when using this reference.⁴³⁹

Table 12.2. Comparisons for efficacy of oral contraceptive pills on premenstrual depressive symptoms (lower triangle) and overall premenstrual symptomatology (upper triangle).

Oral contraceptive pill	Placebo	Premenstrual depressive symptoms	Overall premenstrual symptomatology
Ethinylestradiol /drospirenone (20ug/3mg; 24/4)	-0.04 [‡] (-1.20, 1.08)	-0.68 [‡] (-1.57, 0.21)	-0.19 [‡] (-1.35, 0.91)
0.67 [‡] (-0.63, 1.96)	Estradiol /norgestrel (1.5mg/2.5mg; 24/4)	-0.64 [‡] (-2.05, 0.84)	-0.15 [‡] (-1.52, 1.21)
0.08 [‡] (-0.97, 1.17)	Ethinylestradiol /drospirenone (20ug/3mg; 21/7)	-	-
-0.09 [‡] (-1.14, 1.01)	-0.84 [‡] (-2.13, 0.46)	-	-
0.74 [‡] (-0.08, 1.56)	-0.01 [‡] (-1.36, 1.29)	Ethinylestradiol /desogestrel (20ug/150ug; 24/4)	0.48 [‡] (-0.98, 1.89)
0.07 [‡] (-0.74, 0.82)	-0.68 [‡] (-1.78, 0.33)	0.83 [‡] (-0.55, 2.16)	0.20 [‡] (-0.91, 1.24)
-0.15 [‡] (-1.16, 0.91)	-0.89 [‡] (-2.16, 0.38)	0.15 [‡] (-0.96, 1.20)	-0.38 [‡] (-1.02, 0.20)
0.71 [‡] (-0.32, 1.77)	-0.04 [‡] (-1.32, 1.23)	-0.05 [‡] (-1.35, 1.23)	Ethinylestradiol /drospirenone (30ug/3mg; 21/7)
0.34 [‡] (-0.18, 0.91)	-0.41 [‡] (-1.32, 0.49)	-0.88 (-2.19, 0.45)	-0.28 [‡] (-1.24, 0.67)
	0.63 [‡] (-0.66, 1.92)	-0.03 [‡] (-1.34, 1.31)	Ethinylestradiol /norethindrone (35ug/0.5;1;0.5;0mg; 21/7)
	0.26 [‡] (-0.67, 1.19)	-0.40 [‡] (-1.36, 0.60)	-0.37 [‡] (-1.26, 0.53)
		0.27 [‡] (-0.24, 0.87)	0.49 [‡] (-0.40, 1.37)
			Placebo

Data are standardized mean differences (SMD) with 95% Credible intervals in the column-defining intervention compared with the row-defining intervention. Higher SMD values correspond with fewer symptoms in the column-defining hormonal contraceptive. None of the results were significant and all evidence was of low certainty (marked with [‡]). Ethinylestradiol/norethindrone (35ug/0.5;1;0.5;0mg) is a multiphasic oral contraceptive pill that has four different dosages of norethindrone throughout a 4-week cycle. The dosages before the / refer to the different dosages of the synthetic estrogen, and the ones after the / to those of the synthetic progestin.

Discussion

The results of this pairwise and network meta-analysis suggest that combined oral contraceptives, compared to placebo, may improve overall premenstrual symptomatology in both women with PMS or PMDD. However, we found no support for their efficacy in treating premenstrual depressive symptoms, and none of the combined oral contraceptives was more efficacious than any other.

The main strength of the present study is the use of a comprehensive literature search which enabled us to include seven randomized clinical trials that were not examined in the previous meta-analysis.^{330, 437-440, 443, 444} Also, the network meta-analysis enable us to estimate the comparative efficacy of combined oral contraceptives using both direct and indirect evidence.⁴⁰⁷ However, only one or two trials were available for each comparison and many trials had a relatively small sample size. The combination of these two factors limits the certainty of evidence for each comparison as was reflected by the GRADE assessment, which resulted in a "very low" certainty rating of the summary of evidence due to relatively large imprecision of the estimates from each study. Moreover, there were only two trials that used combined oral contraceptives as a comparison group. Therefore, the estimates of efficacy of one combined oral contraceptive versus another are primarily based on indirect evidence (mainly through placebo). Additionally, the limitations of individual included trials inevitably also limited the quality of our analysis. Only three out of the nine trials included in the network meta-analysis examined all participants randomized into a trial irrespective of what happened subsequently (also known as intention-to-treat analysis).^{436, 444, 446} This may have increased the risk of overestimation of the efficacy of combined oral contraceptives, as use of per-protocol or observed cases analysis violates the principles of randomization. Loss to follow-up was also substantial in almost all trials, which further question the prognostic balance afforded by randomization. Finally, unregistered trials that were not published (as a full journal article) were not included in the search. In case of publication bias against null findings, this may have resulted in an overestimate of effects. However, as we mainly report null findings (except for the effect of oral contraceptives compared to placebo for overall premenstrual symptomatology) we do not expect this to have a large effect on our results.

We are aware of only one previous meta-analysis that attempted to estimate the efficacy of combined oral contraceptives in treating PMDD. This meta-analysis

did not include measures on premenstrual depressive symptoms, but did report that compared to placebo, ethinylestradiol/drospirenone (20ug/3mg) use in a 24-day regimen was effective in reducing overall premenstrual symptomatology in women with PMDD.⁴⁴⁷ This finding therefore stands in direct opposition to results reported here, showing that the same formulation was not superior to other types of combined oral contraceptives or placebo in treating PMS and PMDD. However, there are important differences in the way the data were evaluated that could explain the difference in findings. First, the previous meta-analysis used a fixed-effect rather than a random-effects model for analyzing the data. Fixed-effects models assume that there is a single underlying “true” effect size, which requires that all factors which could influence the effect size are the same in all trials. As this assumption is rarely met, we argue that a random-effects model is more suitable. Second, the previous meta-analysis used frequentist inference testing rather than Bayesian inference testing. While both approaches are reasonable, the contradictory results suggest that the efficacy of ethinylestradiol/drospirenone (20ug/3mg) in a 24-day regimen is not very robust to differences in analytical strategy.

This meta-analysis shows that combined oral contraceptives represent a reasonable therapeutic option to treat overall premenstrual symptomatology in women with PMS or PMDD also seeking contraception. However, combined oral contraceptives were not effective in treating premenstrual depressive symptoms. Physicians should take this into account when prescribing combined oral contraceptives, as depressive symptoms are often the main symptoms in women with PMDD. In women who predominantly complain about depressive symptoms, other commonly used treatment options for women with PMDD, such as selective serotonin reuptake inhibitors (SSRIs) and cognitive-behavioral therapy, might be a more effective approach.^{448, 449} However, certainty of evidence for all available treatments is very low to moderate at best, which highlights that optimal treatment strategies are not yet well-established.

In contrast to what is often thought, there was no evidence for a difference in efficacy between formulations containing drospirenone vs. other formulations or a 24-day regimen vs. other regimens. This suggests that other arguments could justify the decision to choose one contraceptive over another in women with PMS or PMDD seeking treatment with a combined oral contraceptive. As drospirenone is associated with an increased risk for blood clots compared to other hormonal contraceptives,⁴⁵⁰ this may argue against preferentially prescribing a drospirenone-

containing contraceptive. However, the insufficient evidence in favor of a specific combined oral contraceptive over another in treating PMS or PMDD may be due to a lack of statistical power. This clearly highlights the need for additional, large randomized clinical trials in order to answer the question which treatment options are most effective in women with PMS and PMDD. Preferably, these trials should include both women with PMS and women with PMDD, to ensure that findings are generalizable to the full range of severity of premenstrual complaints which additionally would allow to examine whether combined oral contraceptives are equally effective in women with PMDD vs. PMS.

The results of this pairwise and network meta-analysis show that combined oral contraceptives may improve overall premenstrual symptomatology in women with PMS or PMDD, but there is currently no evidence to favor a specific combined oral contraceptive over another. Moreover, there was no evidence that combined oral contraceptives are effective in treating premenstrual depressive symptoms, which are often the primary complaints of women with PMDD. Few trials were available for each formulation of combined oral contraceptives, and certainty of evidence for each comparison was rated as very low. Therefore, more randomized clinical trials with head-to-head comparisons of combined oral contraceptives are needed to determine which formulation is most effective for premenstrual complaints in women with PMS and PMDD.

Acknowledgements

We thank Jan W. Schoones, MA, Information Specialist from Leiden University Medical Centre for assistance in developing the search strategy. We also are grateful for additional data that were provided to us by Cynthia A. Graham, professor at the University of Southampton, Cecilia Lundin, PhD student at Uppsala University, and Lynne Smith of Pfizer.

There was no fundings for this project. The authors have nothing to disclose.

Data accessibility

The full dataset and the script for analyzing the data is freely available from: <https://osf.io/8zbu3/>.

Part III

Researchers' role in science journalism

Chapter 13

From academia to society – improving the quality of science journalism

Anouk E de Wit

Adapted from a case study written for the Open Research Award of the University of Groningen Library (UB) and the Open Science Community Groningen (OSCG), October 2020. <https://www.rug.nl/library/open-access/blog/from-academia-to-society-improving-the-quality-of-science-journalism-02-10-2020>

Mass media are an important and effective channel for disseminating medical news. Unfortunately, messages often contain, intentionally or not, exaggerated claims that are not supported by the research itself. This is problematic as such stories may do harm to people who delay or avoid effective therapy based on such claims. Moreover, it may contribute to a general mistrust and misunderstanding of researchers and their scientific work. However, journalists are not always the ones to blame, and scientists themselves have an opportunity — and a responsibility — to improve the communication about their research. In this case study, I describe the challenges and difficulties that I encountered while trying to present the results of research on a sensitive topic: side effects of oral contraceptive pills.

Google oral contraceptive pill and depression, and you will find it's a hot topic. Online conversations stretch from intimate user posts talking about their personal experiences to political debates that are expressions of religious or social beliefs. So when I published the results of our observational study, reporting that young women on oral contraceptive pills had 21.2% higher depressive symptoms scores than their non-using counterparts,³⁸⁷ I felt obliged to convey our findings to society in a honest way. After all, I did not want our work to be taken as a recommendation for limiting oral contraceptive pill use.

I have learned a great deal from this experience, and have three tips for researchers to increase the likelihood that the media do not exaggerate research findings. First, write an honest press release and offer journalists some talking points. Often, exaggerated claims are already present in the press releases issued by universities. A study that analyzed press releases and their associated peer-review papers and news stories on exaggerations in advice to readers and causal statements drawn from observational research, showed that when caveats appeared in press releases, there was an increase in caveats in news articles (ORs 9.6 and 9.5 for caveats for advice and causal claims). Importantly, exaggerated claims were not associated with more journalistic uptake.¹³⁴ Hence, wanting to convey an honest message, starts with you. In our press release, our main message was that, because of the study design, a cause-effect relationship could not be established. Boiling down your findings to easy-to-understand talking points was great practice that contributed tremendously to my own understanding of our research. Fortunately, most journalists did a great job translating the message.⁴⁵¹
⁴⁵² However, an honest press release was not a closing solution, as some titles still contained exaggerated claims.^{453, 454}

However, conveying your message orally (for example, television) instead of in writing to a broad public is a different profession, and needs practice. Therefore, my second tip is to rehearse the talking points by discussing your research with others (preferably nonscientists). During my first interview with a journalist for a national news channel, he refused to include our research in the broadcast because the message was ‘too’ nuanced. A couple of interviews later however, I was able to simplify the findings and hence it was accepted for television.⁴⁵⁵

Finally, my last tip is to write or talk about your research yourself. It is odd that we often work on research for years, but eventually wait-and-see whether media will translate the findings well. A platform I used to convey my message is Dokter Media (www.doktermedia.nl). This non-profit platform aims to explain and indicate health-related science news in layman’s terms. I wrote a review⁴⁵⁶ and started a podcast⁴⁵⁷ about the meaning of our findings so those would be accessible to anyone interested. Motivated by my own experiences, I am now an editor at Dokter Media. With the focus on Open Science, I aim to raise awareness among scientists that they have an obligation to society, and an opportunity at our platform, to communicate about their research. In sum, scientists have an opportunity and a responsibility to communicate their findings to interested people from society. Though this can be difficult, providing talking points for journalists and rehearsing the essentials of your findings are good practice for conveying a clear message. If you want to talk about your research without the interference of others, Dokter Media is inviting researchers to write or talk about their findings in a nuanced way

Chapter 14

Discussion

The overall purpose of this thesis was to more precisely define how sex steroids relate to depression. Specifically, the aim was to answer questions about the strength and nature of these potential associations that have remained largely unanswered so far. In this last chapter, the findings will be put in perspective and their implications for clinical practice and future research will be reviewed. First, the main findings will be summarized, followed by an overview of strengths and limitations of the work. Then, based on these main findings, answers will be provided to the three questions that guided this thesis. Firstly, are endogenous sex steroids associated with depression, and if so, to what extent? Secondly, do hormonal contraceptives cause depression? And thirdly, what can researchers do to improve science journalism? Finally, starting-points for future research that is needed to further unravel the role of sex steroids in depression are proposed and discussed.

Summary of the main findings

Part I: Endogenous sex steroids and depression

The first part of this thesis, Chapters 2 through 8, focused on the associations of endogenous sex steroids with depression and on the methodological aspects of determining such associations.

Chapter 2: Plasma and salivary levels of testosterone were only moderately associated

Measuring testosterone in blood samples is the gold standard. If testosterone could be measured as reliably in saliva as in blood, this would provide a welcome, non-invasive and less stressful alternative. Therefore, the association between testosterone levels determined in blood and saliva was estimated in 2048 participants from Netherlands Study on Depression and Anxiety (NESDA). Although this association was statistically significant, it had a very moderate strength. Hence, the determination of testosterone in blood continues to be the method of choice.

Chapter 3: Androgen levels in adult women with and without major depressive disorder (MDD) were not different

MDD is more prevalent amongst women with medical conditions that are associated with elevated androgen levels. This suggests that higher circulating androgen levels in women could have detrimental effects on their mood. To test this hypothesis, the associations between androgen levels, and the presence and

course of MDD were examined among 1659 women from NESDA. Of all androgens studied, only free testosterone levels were slightly elevated in women with current MDD compared to those without MDD. However, none of the androgen levels predicted the onset or remission of MDD. Therefore, the findings do not support the idea that plasma androgens within the physiological range have an important deleterious nor protective effect on mood in women.

Chapter 4: Adult men with and without MDD had similar testosterone levels

Depressive symptoms are more prevalent in men with low testosterone levels. Yet, it was unknown whether this association could be ascribed to symptoms that overlap with hypogonadism, such as sexual dysfunction and loss of libido, or also to core emotional symptoms of MDD such as a depressed mood or anhedonia. Therefore, the associations between androgen levels and individual MDD symptoms and MDD were examined among 823 men from NESDA. It was shown that lower free testosterone levels in men were only associated with less interest in sex. Nevertheless, none of the androgens were associated with current, or the onset of first or recurrent MDD. These results suggest that androgen levels are of limited clinical relevance to MDD in adult men.

Chapter 5: Older men and women with MDD and aberrant testosterone levels were not at increased risk for metabolic syndrome compared to those with more average testosterone levels

Aberrant testosterone levels and MDD have both been associated with metabolic syndrome. Patients with both aberrant testosterone levels and MDD may therefore be more likely to develop metabolic syndrome. To test this hypothesis, the association between testosterone and metabolic syndrome, and the moderating effect of MDD on this association was examined in 478 older men and women from Netherlands Study on Depression in Older persons (NESDO). Indeed, lower testosterone levels in older men, and in contrast, higher testosterone levels in older women were associated with the metabolic syndrome. However, the associations did not differ between older men and women with and without MDD. This suggests that having MDD did not modify the risk for metabolic syndrome in women and men with testosterone levels at the higher or lower end of the spectrum, respectively.

Chapter 6: Adult men and women with higher or lower testosterone levels did not attempt more often suicide than those with more average testosterone levels

Small cross-sectional studies have shown that testosterone could be implicated in suicidality. However, prospective studies are needed to exclude reverse causation of this association as stress that coincides with a suicide attempt may alter androgen levels. For this thesis, the cross-sectional and longitudinal association between androgen levels and current and future suicidality was examined in 2861 adult men and women from NESDA. It was shown that none of the associations was statistically significant. Therefore, these findings did not support previous claims of such an association.

Chapter 7: Perimenopausal women with greater variability in estradiol had more depressive symptoms than perimenopausal women with less variability

Women are alleged to be at increased risk for depressive symptoms during the menopause transition, when levels of estradiol and progesterone fluctuate intensively. Therefore, the associations between variability in estradiol and progesterone, and depressive symptom were examined in 50 perimenopausal women. The results showed that greater estradiol variability and absence of progesterone levels consistent with ovulation, were associated with increased depressive symptoms in perimenopausal women from PERI. These results suggest that perimenopausal mood instability may be the result of fluctuations in estradiol and progesterone that occur during the menopause transition.

Chapter 8: Perimenopausal women with less variability in estradiol had more irritability symptoms than perimenopausal women with more variability

Hormonal factors that predispose women to irritability during the perimenopause have hardly been studied. In order to find hormonal risk factors, the associations estradiol, follicle stimulating hormone (FSH) and progesterone, and irritability were examined in 50 perimenopausal women from PERI. In contrast to the findings of Chapter 7, decreased rather than increased variability in estradiol levels was found to be associated with more irritability. This association was independent of the severity of depressive symptoms, which may suggest that irritability might be linked to different estradiol dynamics than depressive symptoms during the perimenopause.

Part II: Hormonal contraceptives and depression

The second part of this thesis focused on the effects of an exogenous source of sex steroids on depression: hormonal contraceptives. In the **Chapter 9 through 12**, the aim was to examine whether hormonal contraceptives are likely to cause depressive symptoms, and to determine who might be at risk for this effect.

Chapter 9: Adolescent oral contraceptive pill users reported more depressive symptoms than adolescent non-users

Oral contraceptive pill (OC) use have been associated with an increased risk of subsequent depression, especially in adolescent women. To study whether OCs also link to concurrent depressive symptoms, the depressive symptom severity of OC users was compared to those of non-users in a sample of 1010 young women from TRacking Adolescents' Individual Lives Survey (TRAILS). The results showed that although OC use was not associated with depressive symptoms when all observations from women aged between 16 and 25 were combined, 16-year-old girls reported higher depressive symptoms scores when using OCs. This study highlights that young women may be especially vulnerable to the potential depressogenic side effects of hormonal contraceptives.

Chapter 10: Adolescent oral contraceptive pill use was a risk factor for MDD later in life

Although OC use has been associated with depression in the short-term, long term associations remain largely unexplored. Therefore, the association between adolescent OC use and MDD later in life was examined in a sample of 725 women from TRAILS. To gain insight in the impact of analytical choices on the outcome of such study, all theoretically justifiable statistical models were studied. It was shown that adolescent OC use was risk a factor for experiencing an episode of MDD later in life. The results from this study strengthens the findings from Chapter 9 that young women may be especially at risk for the potential effects of OCs on MDD.

Chapter 11: Hormonal contraceptives did not cause depressive symptoms in adult women

As shown in Chapter 9 and Chapter 10, OC use was associated with a concurrent and longitudinal increased risk of (developing) depression. However, the observational nature of these studies precludes causal inference. With use of a network meta-analysis of randomized clinical trials with data from 5833 women,

the association between hormonal contraceptive use and depressive symptoms was examined. The results showed that hormonal contraceptives probably do not *cause* depressive symptoms. However, the generalizability of the findings is limited to a group of experienced hormonal contraceptive users, as only a very limited number of first-time users or women with previously negative experiences with hormonal contraceptive use were included in the examined trials.

Chapter 12: *Oral contraceptive pills did improve overall premenstrual symptomatology in women with premenstrual syndrome, but not premenstrual depressive symptoms specifically*

OCs are often considered a treatment option for women with premenstrual syndrome (PMS) or premenstrual dysphoric disorder (PMDD), but evidence remains scarce. Therefore, a pairwise and network meta-analysis of randomized clinical trials on the efficacy of OCs in women with PMS or PMDD was used to estimate the (comparative) efficacy of OCs on premenstrual depressive symptoms and overall premenstrual symptomatology. The results showed that OCs were more efficacious than placebo in treating overall premenstrual symptomatology, but not premenstrual depressive symptoms specifically. There is no evidence for one OC being more efficacious than another. Altogether, these results challenge the presumed efficacy of OCs on one of core complaints in women with PMDD; depressive symptoms.

Part III: Researchers' role in science journalism

Publication of the results of Chapter 9 made it clear that the potential deleterious effects of hormonal contraceptives are a sensitive topic that is likely to get media attention. The pitfalls and possibilities with regards to science journalism were therefore discussed in the final publication of this thesis, Chapter 13.

Chapter 13: *Researchers have multiple possibilities to improve science journalism*

Medical news messages often contain exaggerated claims that are not supported by the research itself. Journalists are not the only ones to blame for this as researchers themselves have an opportunity — and a responsibility — to improve the communication in the media about their research. In a case report, difficulties that researchers may experience while presenting the results of their research to a layman public were discussed. Moreover, tips were shared in order to provide researchers tools and possibilities to increase the likelihood that the media do

not misinterpret research findings. These tips included writing a clear and correct press release, and providing talking points to journalists.

Strengths and limitations of this research

In each Chapter the specific study strengths and limitations were reported. A summary of the overarching strengths and limitations of the research described in this thesis is given in the current section.

Strengths

This thesis provides methodological and empirical insights in the relationship between sex steroids and depression, and therefore adds to the scientific body of knowledge in the field. First, large cohort studies with well-phenotyped populations were used to determine the associations in **Chapter 2 through 10**. The participants of these studies were followed for many years with thorough diagnostic and other assessments, and low percentages of participants being lost-to-follow-up. Information on factors that impact sex steroids levels, such as body mass index (BMI) and specific medication use, were also measured and could therefore be included in the statistical models. Moreover, these cohorts each focused on subjects in different phases of life (TRAILS, adolescents; NESDA, adults; PERI, perimenopausal women; and NESDO, older persons). Together this offered the possibility to test hypotheses specific for certain ages. Second, liquid chromatography–tandem mass spectrometry (LC-MS/MS) measures were used to determine sex steroid levels in **Chapters 3, 4, 6, 7 and 8**. In addition to the increased reliability, LC-MS/MS measures also determine multiple sex steroids at the same time. This provided insights in associations on which very little has been published so far. Third, the use of rigorous and/or new analysis techniques made it possible to shed light on hypotheses that previously were difficult to test. For example, all theoretically justifiable models were tested using specification curve analysis to gain insight into the impact of analytical choices on the association of adolescent OC use and future MDD in **Chapter 10**. Also, state-of-the-art network meta-analysis techniques of randomized clinical trials were used to estimate the comparative harmful or beneficial effects of different OCs using both direct and indirect evidence in **Chapter 11 and 12**. This allowed for summarizing and interpreting the evidence for their putative effects on depressive symptoms from

a broader perspective. Altogether, the use of well-phenotyped cohort studies, state of the art sex steroid measurements and rigorous analysis techniques in this thesis contributed to the reliability of the results about associations that previously have not been thoroughly investigated.

Limitations

This thesis attempted to more precisely define how sex steroids relate to depression. Although many questions have been answered, some issues remain unresolved. These are the main limitations of this thesis. First, the majority of chapters (**Chapters 2 through 10**) yielded the results of observational cohort studies which precluded inferences on cause-effect relationships. Second, sex steroid level determinations in **Chapter 2 through 8** relied on the measurement of peripheral blood levels. To what extent these peripheral sex steroid levels reflected the processes in the central nervous system remains to be seen. Concentrations of neurosteroids generally exceed the circulating sex steroid levels due to accumulation and/or local synthesis.³⁴ Nevertheless, most of the steroid metabolites in the central nervous system are synthesized from peripheral steroid precursors or are directly transported through the blood-brain-barrier. Hence, peripheral sex steroids levels probably still reflect the exposure of the brain to sex steroids.⁴⁵⁸ Third, sex steroid levels were measured with radioimmunoassay's in **Chapter 2 and 5**. These measures are less reliable than LC-MS/MS analyses, in particular at lower concentrations (<5 nmol/L), such as found in women (Chapter 2 and 5) and hypogonadal men (Chapter 5).^{163, 164} Consequently, androgen levels in participants from NESDA were remeasured with LC-MS/MS in order to have most reliable measurements available for Chapter 3, 4, and 6.

Main findings; clinical and scientific perspectives

The aim of this thesis was to provide insights into the relation between sex steroids and depression, by focusing on three questions. Firstly, are endogenous sex steroids associated with depression, and if so, to what extent? Secondly, do hormonal contraceptives cause depressive symptoms? Thirdly, what can researchers do to improve science journalism? In this section answers to these three overarching questions will be provided by combining the main findings.

Are endogenous sex steroids associated with depression, and if so, to what extent?

Single androgen levels measured in adults are not associated with depression, but variability in estradiol levels may be implicated in depression in women

Results from **Chapter 3, 4, 6, 7 and 8**, showed that single measures of androgens were not predictive of depression. In contrast, variability of estradiol was associated with mood deterioration during the perimenopause as shown in **Chapter 7 and 8**. This underlines the importance of measuring sex steroid levels serially in order to capture their dynamics, rather than obtaining a single hormone level value.

If one would extrapolate the variability findings, these would imply that hormonal treatment could prevent or treat depressive symptoms perimenopausal women, but may at the same time ameliorate irritability symptoms. Indeed, there is some evidence that transdermal estradiol (combined with intermittent micronized progesterone) is more effective than a placebo in preventing the development of or treating depressive symptoms in perimenopausal women.^{333, 459, 460} However, the number of women included in these studies was small and none of the trials determined the variability of estradiol before and during the treatment. Hence, the hypothesis that reducing the magnitude of variability in estradiol represents a mechanism by which such treatment could benefit mood, could not be ascertained. Also, it is unknown who is, and importantly who is not, likely to benefit from hormonal treatment. If anything, the observation that the majority of women do not experience more depressive symptoms during the perimenopause compared to the premenopause, suggests that the majority of women do not experience changes in their mood as a consequence of the increased variability during the perimenopause.¹¹⁰ Hence, the majority of women will likely also not benefit from hormonal treatment.

There are no such things as male and female hormones

In **Chapter 3 and 4**, the hypothesis for the direction of the association between androgens and depression was in the opposite direction for men and women, respectively. This underlines that sex in itself should be viewed as an imperfect proxy for partially still unknown factors, such as sex steroids or sex-linked genes, as they explain the variation better than sex does. This will lead to a better understanding of the true sources of variation.⁴⁶¹ Focusing on these factors also implies that using terms as “female” and “male” hormones is rather stereotyping

than an accurate description. In fact, testosterone, and other sex steroids like estradiol and progesterone, are important for both sexes, although in a different manner. Although women have almost 14 times lower testosterone levels than men, women are more sensitive to the effects of testosterone.⁴⁶²

Do hormonal contraceptives cause depression?

Millions of women use hormonal contraceptives as means to prevent pregnancies or diminish menstrual symptoms. Successful fertility control brings many social and economic benefits for women, such as better education and personal autonomy. Still, side effects may occur. Over the last years the potential deleterious effects on mood have gained more attention. However, based on experimental data, it was shown in **Chapter 11** that most women can use these hormonal contraceptives safely with regards to the potential depressogenic side effects. Nevertheless, there may subgroups of women who experience an increase or decrease in depressive symptoms after they start taking hormonal contraceptives.

Adolescent women could be a subgroup of women that is sensitive to the depressogenic effects of hormonal contraceptives

The finding from Chapter 11 contrasts with the findings from **Chapters 9 and 10**. In these Chapters OC use in teenage women was associated with depressive symptoms, and future depression, respectively. However, there is one crucial difference in the populations that were examined in each of the studies. The women included in Chapter 9 and 10 were much younger than the women examined in Chapter 11. In other words, the absence of an effect in the latter Chapter, may have been a result of the on average much older age of women examined in the randomized clinical trials. Hence, adolescent women might be a subgroup of women that is sensitive to the depressogenic effects of hormonal contraceptives. Therefore it does remain important to monitor for depressive symptoms after these young women start taking hormonal contraceptives.

There is no evidence that women with PMS and PMDD are a subgroup of women that experiences positive impact on depressive symptoms with oral contraceptive pill use

One group of women who are thought to benefit from hormonal contraceptive use (OC use specifically) with regards to depressive symptoms, are women with PMS and PMDD. Indeed, the overall effect reported in **Chapter 12** suggested that

OCs may reduce depressive symptoms in these women. Yet, the imprecision of the estimate was too large to definitely conclude that women with premenstrual women indeed are a subgroup of women who experience a positive impact of OC use.

What can researchers do to improve science journalism?

Science journalists, public information officers and researchers depend on each other

The case described in **Chapter 13** showed that the creation of accurate and valid science news outlet involves the collaboration of journalists, public information officers, and researchers. This highlights that improving the communal understanding of each other's needs is important to improve science journalism. For researchers this would imply a couple of things. First, researchers need to understand the urgency of journalist's daily publication schedules and moreover that every science article by journalists is a media product that finds itself in a competitive situation with other media products and must therefore achieve a place in the attention economy of its target population.⁴⁶³ This competition is increasing due to the fragmentation of the public sphere. The channels upon which individuals historically relied for science news (news papers and magazines) are increasingly being superseded by newer sources like social media.⁴⁶⁴ Moreover, users pick and choose among channels and then give those selections only a quick scan before moving on which puts even more pressure on media to be noticed. Thus, being responsive as a researcher helps journalists to stick to their time-lines. Second, journalists often rely on information presented in press releases. Therefore, having an accurate and valid press release which is understandable for people without a scientific background (which is often true for journalists) is key.¹³³ Such press releases involves the collaboration of researchers with public information officers employed by the journal where the article will be published and/or those who are employed by the institution of the researcher.

Write an accurate and understandable press release

The personal experiences discussed in **Chapter 13** suggest that researchers should critically check press releases on exaggerating claims about their research. For example, researchers should check if the language of the press release unintentionally suggests a causal relationship when the research was of

observational nature and check whether the meaning of the included numbers is clear. For instance, citing odds ratios should be avoided as absolute risks and numbers needed to treat and harm are much easier to understand. If mentioning relative risks is unavoidable, sufficient context should be provided to make them comprehensible. Also, researchers should review whether the press release includes sentences explaining the sensitivity, the specificity, and the negative and positive predictive values when reporting about the results of diagnostic tests. In the case the press release reports about the result of a trial, researchers should make sure that the population is carefully described including what the use of this population means for the generalizability (external validity) of the results. Overall, hyperbolic language like “breakthrough” and “game-changing” should be avoided if this cannot be supported by the findings.

Future directions

This thesis provides methodological and empirical insights in the relationship between sex steroids and depression. The results of the studies suggested that both endogenous sex steroids and hormonal contraceptives may play a role in this relationship. However, risk factors that can determine in whom and under what circumstances these sex steroids are of importance, need to be defined more precisely. If anything, causality remains to be proven too. These issues are starting-points for future studies. In this section, future directions for studies on sex steroids and depression will be discussed. As to depression, such studies will hopefully provide a better insight into who might benefit from hormonal treatment and who would not. Finally, the section will end with the bringing up an important step to be taken in order to improve the science-society relationship.

Future studies on associations between endogenous sex steroids and depression

Androgens and depression: focus on adolescents and the determination of change scores

The generally negative findings on the association between single androgen measures and depression, could imply that these sex steroids are not involved in the etiology of depression. However, it is more likely that single measurements in adulthood are not sufficient to capture the impact of the sex steroids on brain

functioning. Although single sex steroid levels during adulthood could possibly reflect activational effects of sex steroids on brain functioning, they do not reflect their organizational impact. In contrast, during adolescence, sex steroids do have organizational effects with long-lasting neuroadaptations, causing crucial permanent changes in brain functioning and therefore behaviors later in life.^{34-37, 77, 397} Measuring androgens during this much earlier phase could therefore provide better insights in intraindividual differences in for example androgenization of the brain which may predispose individuals to psychopathology later in life.

Ideally, androgens should be measured at multiple timepoints in future studies. A single fasting morning sample provides a good approximation of the average androgen concentration but provides no insight in “normal” sex steroid levels for a particular person.⁴⁶⁵ “Normal” levels are usually defined as the set of values within which 95 percent of the normal population falls (that is, 95% prediction interval). However, within that range of levels, a certain level may still be “abnormally” high or low for a particular person because the variation of values within a person is smaller than the variation of values within 95% of the population. Therefore, multiple measurements of androgens within a time span of several hours would increase the precision of the androgen measure, and an additional measure of the androgen after for example a year would enable the possibility to calculate change scores. However, multiple measurements would require the smart use of stored blood samples and additional money which may be complicated in large cohort studies.

Non-invasive alternatives that could replace androgen determinations in blood are needed, especially for young adolescents. In **Chapter 2**, it was shown that salivary measures of testosterone with Salivettes® and radioimmunoassays did not correspond well with plasma levels. However, determinations with Salimetrics® (passive drooling) rather than Salivettes®, and LC-MS/MS rather than radioimmunoassays have proven to be reliable, also in children.⁴⁶⁶ Measurement of androgens in hair using LC-MS/MS would offer the possibility to estimate androgen production over multiple months.^{169, 223} However, in a recent large study in over 2000 women, almost half of the testosterone levels in hair were non-detectable due to low concentration in samples.⁴⁶⁷ As young adolescents have similar or even lower androgens levels than women, the reliability of testosterone measurements in hair is currently insufficient.

Estradiol, progesterone and depression in perimenopausal women: focus on selecting those who may benefit from hormonal treatment.

Earlier it was discussed that probably not all women are sensitive to changes in sex steroids, and therefore not all women are likely to benefit from hormonal treatment. Determining a woman's sensitivity profile to sex steroid changes with regards to depression would allow to select those women in whom hormonal treatment will possibly be useful.

A first step in determining a sensitivity profile is to increase the frequency of sex steroid measurements. Especially in the three years before menopause, estradiol fluctuates extensively and ovulations occur less predictively.³⁰⁹ Therefore, assessing the pattern of change in estradiol and progesterone based on daily estradiol measurements for an entire menstrual cycle or up to 50 days, whichever comes first, rather than inferring the patterns from samples obtained monthly or weekly, would improve the precision of variability profiles in perimenopausal women.³⁰⁹ Ideally, these levels should be measured in blood rather than in urine or saliva. Although the latter two options would be less painful and stressful for participants, these would also be less reliable.^{468, 469} Hopefully, less stressful ways to measure sex steroids even more often than daily will become available. Currently, such methods are already available for cortisol measures. Using a catheter placed under the skin on the abdomen, interstitial fluid can be collected to map how cortisol levels vary from hour to hour.^{470, 471} It is yet not possible to use a similar technique for sex steroid measures due to binding of testosterone to the collection tube. Nevertheless, if such problems were solved, this would provide better insights in the true variability of estradiol and progesterone.

More precise patterns of estradiol and progesterone could be linked to changes in mood if qualitative data about the participants' mood and activities are collected at the same time. These associations will provide insight in 1) how changes in sex steroids may link to mood, and 2) how the strength and direction of these associations may differ between women. The effect of these insights is twofold. First, knowing which women are potentially sensitive to sex steroid changes, and importantly, also who are not, allows for the search for risk factors that could predict a woman's sensitivity without extensively monitoring mood and sex steroid levels. Second, by conducting a trial in the sex steroid variability sensitive women, one could investigate the efficacy of hormonal treatment in a group of women who are more likely to benefit from such treatment. These trials should also include frequent sex steroid measurements to confirm the hypothesizes that a reduction of hormonal fluctuations improve depressive symptoms in these women.

Future studies on the effects of hormonal contraceptives on depression

Despite the myriad beneficial effects of hormonal contraceptives, observational evidence suggests that these drugs may deteriorate depressive symptoms in some women. It remains however controversial whether this association has a causal nature. More data of better quality is needed to provide a definite answer to this question.

Focus on first-time users and the specification of moderators when studying the depressogenic effects of hormonal contraceptives

One of the problems with the currently available data is that the study population of many of the available trials are highly selected and may not be particularly representative of the actual population of hormonal contraceptive users. Worldwide, more than 400 million women use hormonal contraceptives³⁸⁶ and about 50 percent of women in modern societies do so during adolescence.¹²⁵ However, the majority of trials have been conducted in adult women which may have induced a healthy user bias and consequently a too rosy picture of the effects of hormonal contraceptives on depressive symptoms. To make up for the underrepresentation of adolescent women and first-time users in randomized clinical trials, the priority of future trials must be to especially study these groups.

The conduction of clinical trials in adolescent women may involve some additional difficulties in comparison to the conduction of such trials in adult women. The most obvious difficulty is that the potential randomization to a placebo (condom) is likely to be ethically unacceptable given the significantly higher failure rate of condoms compared to hormonal contraceptives.⁴²⁶ Therefore, the possibilities for a placebo-controlled trial will be limited. Nevertheless, head-to-head trials will still be a good option with the additional advantage that such trials allow for the determination of the hormonal contraceptive with the most beneficial side effect profile. As randomized clinical trials require considerable costs and time, preferably such trials should include other assessments of for example sexual functioning as well so that one study could contribute towards multiple research questions simultaneously.

A possible design of a head-to-head trial would be to randomize adolescent women searching for contraception (for different reasons including acne, dysmenorrhea and contraception, most likely in the primary care setting) to several of the most used contraceptives. These contraceptive options would for example

include two oral contraceptive pills, and a hormonal and copper intrauterine device. With an assessment of depressive symptoms at baseline, and after three and six months, one can determine the comparative effects of (hormonal) contraceptives on mood. During these assessments there should also be data collected about potentially important factors that could moderate the relationship between hormonal contraceptives and depression. These factors would for example include first-time usership, indication of use, PMS or PMDD diagnosis, and previous user experiences. The results of these trials would allow to more precisely estimate the effect of hormonal contraceptives on mood and moreover to define subgroups of women at risk for the depressogenic effects of hormonal contraceptives.

Researchers without commercially interest should run these trials

The question remains who will conduct the clinical trials on the side effects of hormonal contraceptives. Expecting that multinational pharmaceutical companies will take the lead, would probably lead to disappointment. First, there is no incentive for pharmaceutical companies to study side effects in general, and certainly not when drugs are already on the market as it does not generate revenues. Second, data on the depressogenic effects of hormonal contraceptives will also not be collected as part of new drug research as pharmaceutical companies abandoned their hormonal and nonhormonal drug discovery research and development programs in the early 2000s. Reasons for this include the existence of many low-cost products in the market, the high bar for developing commercially successful new contraceptives with highly effective existing options already marketed, liability concerns since contraceptives are generally used for long periods of time by healthy women, and the costs of the such programs.⁴⁷²

The value of studying the depressogenic effects of hormonal contraceptives is not monetary, but the wellbeing of millions of women using them. Therefore, academic researchers who have other incentives than their industry counterparts are best suited to take this task. This would provide the additional advantage that research performed within academia is less likely to produce a positive, flattering result than industry funded trials.^{473, 474}

Future directions for science journalism

In this thesis, opportunities for researchers were discussed that could help them to improve their part in science journalism, such as by checking press releases more

critically. However, the objective of the Dutch ministry goes beyond this and aims at stimulating researchers to feature engaging dialogues with the public to improve the link between society and science.⁴⁷⁵ There are multiple initiatives that contribute to this objective (among them for example are Science Café meet-ups), but the problem remains that researchers are not given a serious incentive to communicate their science. Currently, academic publications are key to gaining tenure or to get promoted and therefore researchers are unlikely to take time away from this aspect of their job in order to participate in science communication. If institutions are serious about their science communication efforts, they should aim to assign real value to the communication efforts that researchers make, such as career advancement or a pay bump.

Conclusions

This thesis provides methodological and empirical insights in the relationship between sex steroids and depression. The use of well-phenotyped cohort studies, state of the art sex steroid measurements and rigorous analysis techniques in this thesis contributed to the reliability of the results about associations that previously have not been thoroughly investigated. In the previous Chapters, it became clear that sex steroids may have multiple and diverse roles in depression. The dynamics of sex steroids may play a role in the development and maintenance of depression, but their specific role also depends on factors such as the phase of life someone is in, and the biological sex of the person under investigation. Therefore, comprehensive knowledge on factors of influence on sex steroids, such as age, BMI, and medication use are vital to interpret the research findings correctly. Given the diurnal rhythm and variability of sex steroids within persons and the impact of the menstrual cycle, correct timing of sex steroid measurements with gold standard laboratory assessments is key to provide further insight in the impact of sex steroids on mood. Study designs with repeated sex steroids measurements within individuals may fit this line of research better than traditional nomothetic designs with one or few blood samplings. With regard to the mood effects of hormonal contraceptives, it became clear that contraceptive use in and of itself does not cause depressive symptoms in adult women. However, future randomized clinical trials on this relationship in younger women are needed to confirm or reject the hypothesis that hormonal contraceptive use causes depression in young women. These two suggestions may help provide even more definite answers to the question of whether, and more importantly, of how hormones can get you down.

References

1. Kessler RC. Gender differences in major depression: epidemiologic findings. In: Frank E, ed. *Gender and its Effect on Psychopathology*. American Psychiatric Press; 2000:61-62-84.
2. Bromet E, Andrade LH, Hwang I, et al. Cross-national epidemiology of DSM-IV major depressive episode. *BMC Medicine*. Jul 2011;9:90. doi:10.1186/1741-7015-9-90
3. Hasin DS, Sarvet AL, Meyers JL, et al. Epidemiology of Adult DSM-5 Major Depressive Disorder and Its Specifiers in the United States. *JAMA Psychiatry*. 04 2018;75(4):336-346. doi:10.1001/jamapsychiatry.2017.4602
4. De Graaf R, ten Have M, van Dorsselaer S. *De psychische gezondheid van de Nederlandse bevolking. NEMESIS-2, Opzet en eerste resultaten*. 2010.
5. Malhi GS, Mann JJ. Depression. *Lancet*. 11 2018;392(10161):2299-2312. doi:10.1016/S0140-6736(18)31948-2
6. Otte C, Gold SM, Penninx BW, et al. Major depressive disorder. *Nature Reviews Disease Primers*. 09 2016;2:16065. doi:10.1038/nrdp.2016.65
7. Ribeiro JD, Huang X, Fox KR, Franklin JC. Depression and hopelessness as risk factors for suicide ideation, attempts and death: meta-analysis of longitudinal studies. *British Journal of Psychiatry*. 05 2018;212(5):279-286. doi:10.1192/bjp.2018.27
8. Wittchen HU, Jacobi F, Rehm J, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol*. Sep 2011;21(9):655-79. doi:10.1016/j.euroneuro.2011.07.018
9. APA. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). Arlington: American Psychiatric Association; 2013.
10. Fergusson DM, Horwood LJ, Ridder EM, Beautrais AL. Subthreshold depression in adolescence and mental health outcomes in adulthood. *Arch Gen Psychiatry*. Jan 2005;62(1):66-72. doi:10.1001/archpsyc.62.1.66
11. Oxman TE, Sengupta A. Treatment of minor depression. *Am J Geriatr Psychiatry*. 2002 May-Jun 2002;10(3):256-64.
12. Beekman AT, Geerlings SW, Deeg DJ, et al. The natural history of late-life depression: a 6-year prospective study in the community. *Arch Gen Psychiatry*. Jul 2002;59(7):605-11. doi:10.1001/archpsyc.59.7.605
13. Bromet E, Andrade LH, Hwang I, et al. Cross-national epidemiology of DSM-IV major depressive episode. *BMC Med*. Jul 2011;9:90. doi:10.1186/1741-7015-9-90
14. Soares CN, Zitek B. Reproductive hormone sensitivity and risk for depression across the female life cycle: a continuum of vulnerability? *J Psychiatry Neurosci*. Jul 2008;33(4):331-43.
15. Angold A, Costello EJ, Worthman CM. Puberty and depression: the roles of age, pubertal status and pubertal timing. *Psychol Med*. Jan 1998;28(1):51-61.
16. Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. *Arch Gen Psychiatry*. Aug 2003;60(8):837-44. doi:10.1001/archpsyc.60.8.837
17. Tal R, Taylor H, Burney R, Mooney S, Giudice L. *Endocrinology of Pregnancy*. 2015.
18. Ko JY, Rockhill KM, Tong VT, Morrow B, Farr SL. Trends in Postpartum Depressive Symptoms - 27 States, 2004, 2008, and 2012. *MMWR Morb Mortal Wkly Rep*. Feb 2017;66(6):153-158. doi:10.15585/mmwr.mm6606a1

19. Cohen LS, Soares CN, Vitonis AF, Otto MW, Harlow BL. Risk for new onset of depression during the menopausal transition: the Harvard study of moods and cycles. *Archives of General Psychiatry*. 2006;63(4):385-390. doi:63/4/385 [pii]
20. Butler L, Santoro N. The reproductive endocrinology of the menopausal transition. *Steroids*. Jun 2011;76(7):627-35. doi:10.1016/j.steroids.2011.02.026
21. Soares CN. Mood disorders in midlife women: understanding the critical window and its clinical implications. *Menopause (New York, NY)*. 2014;21(2):198-206. doi:10.1097/GME.000000000000193 [doi]
22. Freeman EW, Sammel MD, Lin H, Nelson DB. Associations of hormones and menopausal status with depressed mood in women with no history of depression. *Archives of General Psychiatry*. 2006;63(4):375-382. doi:63/4/375 [pii]
23. Cole MG, Dendukuri N. Risk factors for depression among elderly community subjects: a systematic review and meta-analysis. *Am J Psychiatry*. Jun 2003;160(6):1147-56. doi:10.1176/appi.ajp.160.6.1147
24. Kessler RC, McGonagle KA, Swartz M, Blazer DG, Nelson CB. Sex and depression in the National Comorbidity Survey. I: Lifetime prevalence, chronicity and recurrence. *J Affect Disord*. 1993 Oct-Nov 1993;29(2-3):85-96.
25. Tamaria A, Bharti R, Sharma M, et al. Risk assessment for psychological disorders in postmenopausal women. *J Clin Diagn Res*. Dec 2013;7(12):2885-8. doi:10.7860/JCDR/2013/7580.3784
26. Dalsgaard S, Thorsteinsson E, Trabjerg BB, et al. Incidence Rates and Cumulative Incidences of the Full Spectrum of Diagnosed Mental Disorders in Childhood and Adolescence. *JAMA Psychiatry*. Nov 2019;doi:10.1001/jamapsychiatry.2019.3523
27. Rush AJ, Trivedi MH, Wisniewski SR, et al. Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression. *N Engl J Med*. Mar 2006;354(12):1231-42. doi:10.1056/NEJMoa052963
28. Belmaker RH, Agam G. Major depressive disorder. *N Engl J Med*. Jan 2008;358(1):55-68. doi:10.1056/NEJMra073096
29. Krishnan V, Nestler EJ. The molecular neurobiology of depression. *Nature*. Oct 2008;455(7215):894-902. doi:10.1038/nature07455
30. Jentsch MC, Van Buel EM, Bosker FJ, et al. Biomarker approaches in major depressive disorder evaluated in the context of current hypotheses. *Biomarkers in medicine*. 2015;9(3):277-297. doi:10.2217/bmm.14.114 [doi]
31. Becker J, Breedlove S, Crews D, McCarthy M. *Behavioral Endocrinology*. The MIT Press; 2002:776.
32. Rettberg JR, Yao J, Brinton RD. Estrogen: a master regulator of bioenergetic systems in the brain and body. *Front Neuroendocrinol*. Jan 2014;35(1):8-30. doi:10.1016/j.yfrne.2013.08.001
33. Juraska JM, Sisk CL, DonCarlos LL. Sexual differentiation of the adolescent rodent brain: hormonal influences and developmental mechanisms. *Horm Behav*. Jul 2013;64(2):203-10. doi:10.1016/j.yhbeh.2013.05.010
34. Woolley CS. Acute effects of estrogen on neuronal physiology. *Annu Rev Pharmacol Toxicol*. 2007;47:657-80. doi:10.1146/annurev.pharmtox.47.120505.105219

35. Gillies GE, McArthur S. Estrogen actions in the brain and the basis for differential action in men and women: a case for sex-specific medicines. *Pharmacol Rev.* Jun 2010;62(2):155-98. doi:10.1124/pr.109.002071
36. McHenry J, Carrier N, Hull E, Kabbaj M. Sex differences in anxiety and depression: role of testosterone. *Frontiers in neuroendocrinology.* 2014;35(1):42-57.
37. Schmidt BM, Gerdes D, Feuring M, Falkenstein E, Christ M, Wehling M. Rapid, nongenomic steroid actions: A new age? *Front Neuroendocrinol.* Jan 2000;21(1):57-94. doi:10.1006/frne.1999.0189
38. Oyola MG, Handa RJ. Hypothalamic-pituitary-adrenal and hypothalamic-pituitary-gonadal axes: sex differences in regulation of stress responsivity. *Stress.* 09 2017;20(5):476-494. doi:10.1080/10253890.2017.1369523
39. Ilie IR. *Introduction to endocrinology.* Springer; 2020. <https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=6112688><https://doi.org/10.1007/978-3-030-27382-8><https://link.springer.com/book/10.1007/978-3-030-27382-8><https://link.springer.com/10.1007/978-3-030-27382-8><http://www.vlebooks.com/vleweb/product/openreader?id=none&isbn=9783030273828>
40. Hammond GL. Plasma steroid-binding proteins: primary gatekeepers of steroid hormone action. *J Endocrinol.* 07 2016;230(1):R13-25. doi:10.1530/JOE-16-0070
41. Diotel N, Charlier TD, Lefebvre d'Helencourt C, et al. Steroid Transport, Local Synthesis, and Signaling within the Brain: Roles in Neurogenesis, Neuroprotection, and Sexual Behaviors. *Front Neurosci.* 2018;12:84. doi:10.3389/fnins.2018.00084
42. Shibuya K, Takata N, Hojo Y, et al. Hippocampal cytochrome P450s synthesize brain neurosteroids which are paracrine neuromodulators of synaptic signal transduction. *Biochim Biophys Acta.* Feb 2003;1619(3):301-16. doi:10.1016/s0304-4165(02)00489-0
43. Baulieu EE. Neurosteroids: a novel function of the brain. *Psychoneuroendocrinology.* Nov 1998;23(8):963-87. doi:10.1016/s0306-4530(98)00071-7
44. Paul SM, Purdy RH. Neuroactive steroids. *FASEB J.* Mar 1992;6(6):2311-22.
45. D'Arpe S, Di Felicianantonio M, Candelieri M, Franceschetti S, Piccioni MG, Bastianelli C. Ovarian function during hormonal contraception assessed by endocrine and sonographic markers: a systematic review. *Reprod Biomed Online.* Oct 2016;33(4):436-448. doi:10.1016/j.rbmo.2016.07.010
46. van der Heijden PAHH, Tibosch RMG, Geomini PMAJ, et al. What is the best drug treatment for premenopausal women with bleeding irregularities using the levonorgestrel-releasing intrauterine system? A systematic review. *Eur J Contracept Reprod Health Care.* Aug 2020;1-8. doi:10.1080/13625187.2020.1797663
47. Heberden C. Sex steroids and neurogenesis. *Biochem Pharmacol.* 10 2017;141:56-62. doi:10.1016/j.bcp.2017.05.019
48. Barth C, Villringer A, Sacher J. Sex hormones affect neurotransmitters and shape the adult female brain during hormonal transition periods. *Front Neurosci.* 2015;9:37. doi:10.3389/fnins.2015.00037
49. Morford JJ, Wu S, Mauvais-Jarvis F. The impact of androgen actions in neurons on metabolic health and disease. *Mol Cell Endocrinol.* 04 2018;465:92-102. doi:10.1016/j.mce.2017.09.001
50. Mhaouty-Kodja S. Role of the androgen receptor in the central nervous system. *Mol Cell Endocrinol.* 04 2018;465:103-112. doi:10.1016/j.mce.2017.08.001

51. Cunningham RL, Lumia AR, McGinnis MY. Androgen receptors, sex behavior, and aggression. *Neuroendocrinology*. 2012;96(2):131-40. doi:10.1159/000337663
52. Gundlah C, Lu NZ, Mirkes SJ, Bethea CL. Estrogen receptor beta (ERbeta) mRNA and protein in serotonin neurons of macaques. *Brain Res Mol Brain Res*. Jul 2001;91(1-2):14-22.
53. Osterlund MK, Hurd YL. Estrogen receptors in the human forebrain and the relation to neuropsychiatric disorders. *Prog Neurobiol*. Jun 2001;64(3):251-67.
54. Mitra SW, Hoskin E, Yudkovitz J, et al. Immunolocalization of estrogen receptor beta in the mouse brain: comparison with estrogen receptor alpha. *Endocrinology*. May 2003;144(5):2055-67. doi:10.1210/en.2002-221069
55. Perlman WR, Matsumoto M, Beltaifa S, et al. Expression of estrogen receptor alpha exon-deleted mRNA variants in the human and non-human primate frontal cortex. *Neuroscience*. 2005;134(1):81-95. doi:10.1016/j.neuroscience.2005.03.055
56. Bethea CL. Colocalization of progesterin receptors with serotonin in raphe neurons of macaque. *Neuroendocrinology*. Jan 1993;57(1):1-6. doi:10.1159/000126334
57. Rossetti MF, Cambiasso MJ, Holschbach MA, Cabrera R. Oestrogens and Progestagens: Synthesis and Action in the Brain. *J Neuroendocrinol*. 07 2016;28(7)doi:10.1111/jne.12402
58. Montague D, Weickert CS, Tomaskovic-Crook E, Rothmond DA, Kleinman JE, Rubinow DR. Oestrogen receptor alpha localisation in the prefrontal cortex of three mammalian species. *J Neuroendocrinol*. Jul 2008;20(7):893-903. doi:10.1111/j.1365-2826.2008.01743.x
59. Shughrue PJ, Merchenthaler I. Distribution of estrogen receptor beta immunoreactivity in the rat central nervous system. *J Comp Neurol*. Jul 2001;436(1):64-81.
60. Yokosuka M, Okamura H, Hayashi S. Postnatal development and sex difference in neurons containing estrogen receptor-alpha immunoreactivity in the preoptic brain, the diencephalon, and the amygdala in the rat. *J Comp Neurol*. Dec 1997;389(1):81-93.
61. Simonian SX, Spratt DP, Herbison AE. Identification and characterization of estrogen receptor alpha-containing neurons projecting to the vicinity of the gonadotropin-releasing hormone perikarya in the rostral preoptic area of the rat. *J Comp Neurol*. Aug 1999;411(2):346-58.
62. Laflamme N, Nappi RE, Drolet G, Labrie C, Rivest S. Expression and neuropeptidergic characterization of estrogen receptors (ERalpha and ERbeta) throughout the rat brain: anatomical evidence of distinct roles of each subtype. *J Neurobiol*. Sep 1998;36(3):357-78.
63. Chakraborty TR, Hof PR, Ng L, Gore AC. Stereologic analysis of estrogen receptor alpha (ER alpha) expression in rat hypothalamus and its regulation by aging and estrogen. *J Comp Neurol*. Nov 2003;466(3):409-21. doi:10.1002/cne.10906
64. Simerly RB, Chang C, Muramatsu M, Swanson LW. Distribution of androgen and estrogen receptor mRNA-containing cells in the rat brain: an in situ hybridization study. *J Comp Neurol*. Apr 1990;294(1):76-95. doi:10.1002/cne.902940107
65. Belcher SM. Regulated expression of estrogen receptor alpha and beta mRNA in granule cells during development of the rat cerebellum. *Brain Res Dev Brain Res*. Jun 1999;115(1):57-69.

66. Shughrue PJ, Scrimo PJ, Merchenthaler I. Evidence for the colocalization of estrogen receptor-beta mRNA and estrogen receptor-alpha immunoreactivity in neurons of the rat forebrain. *Endocrinology*. Dec 1998;139(12):5267-70. doi:10.1210/endo.139.12.6525
67. Creutz LM, Kritzer MF. Estrogen receptor-beta immunoreactivity in the midbrain of adult rats: regional, subregional, and cellular localization in the A10, A9, and A8 dopamine cell groups. *J Comp Neurol*. May 2002;446(3):288-300.
68. Guerra-Araiza C, Villamar-Cruz O, González-Arenas A, Chavira R, Camacho-Arroyo I. Changes in progesterone receptor isoforms content in the rat brain during the oestrous cycle and after oestradiol and progesterone treatments. *J Neuroendocrinol*. Oct 2003;15(10):984-90.
69. Auger CJ, De Vries GJ. Progesterone receptor immunoreactivity within steroid-responsive vasopressin-immunoreactive cells in the male and female rat brain. *J Neuroendocrinol*. Jul 2002;14(7):561-7.
70. Brock O, De Mees C, Bakker J. Hypothalamic expression of oestrogen receptor α and androgen receptor is sex-, age- and region-dependent in mice. *J Neuroendocrinol*. Apr 2015;27(4):264-76. doi:10.1111/jne.12258
71. Lu SF, McKenna SE, Cologer-Clifford A, Nau EA, Simon NG. Androgen receptor in mouse brain: sex differences and similarities in autoregulation. *Endocrinology*. Apr 1998;139(4):1594-601. doi:10.1210/endo.139.4.5863
72. Shah NM, Pisapia DJ, Maniatis S, Mendelsohn MM, Nemes A, Axel R. Visualizing sexual dimorphism in the brain. *Neuron*. Aug 2004;43(3):313-9. doi:10.1016/j.neuron.2004.07.008
73. Roselli CE, Handa RJ, Resko JA. Quantitative distribution of nuclear androgen receptors in microdissected areas of the rat brain. *Neuroendocrinology*. May 1989;49(5):449-53. doi:10.1159/000125151
74. McGinnis MY, Katz SE. Sex differences in cytosolic androgen receptors in gonadectomized male and female rats. *J Neuroendocrinol*. Mar 1996;8(3):193-7.
75. Wood RI, Newman SW. Androgen receptor immunoreactivity in the male and female Syrian hamster brain. *J Neurobiol*. Jun 1999;39(3):359-70.
76. Yu WH, McGinnis MY. Androgen receptors in cranial nerve motor nuclei of male and female rats. *J Neurobiol*. Jan 2001;46(1):1-10.
77. Del Río JP, Alliende MI, Molina N, Serrano FG, Molina S, Vigil P. Steroid Hormones and Their Action in Women's Brains: The Importance of Hormonal Balance. *Front Public Health*. 2018;6:141. doi:10.3389/fpubh.2018.00141
78. Rubinow DR, Schmidt PJ. Is there a role for reproductive steroids in the etiology and treatment of affective disorders? *Dialogues Clin Neurosci*. 09 2018;20(3):187-196.
79. Wharton W, Gleason CE, Olson SR, Carlsson CM, Asthana S. Neurobiological Underpinnings of the Estrogen - Mood Relationship. *Curr Psychiatry Rev*. Aug 2012;8(3):247-256. doi:10.2174/157340012800792957
80. Rainville JR, Tsyglakova M, Hodes GE. Deciphering sex differences in the immune system and depression. *Front Neuroendocrinol*. 07 2018;50:67-90. doi:10.1016/j.yfrne.2017.12.004
81. Panagiotakopoulos L, Neigh GN. Development of the HPA axis: where and when do sex differences manifest? *Front Neuroendocrinol*. Aug 2014;35(3):285-302. doi:10.1016/j.yfrne.2014.03.002

82. Pitsillou E, Bresnehan SM, Kagarakis EA, et al. The cellular and molecular basis of major depressive disorder: towards a unified model for understanding clinical depression. *Mol Biol Rep.* Oct 2019;doi:10.1007/s11033-019-05129-3
83. Biegon A, Kim SW, Alexoff DL, et al. Unique distribution of aromatase in the human brain: in vivo studies with PET and [N-methyl-11C]vorozole. *Synapse.* Nov 2010;64(11):801-7. doi:10.1002/syn.20791
84. Greenblatt RB, Nezhat C, Roesel RA, Natrajan PK. Update on the male and female climateric. *J Am Geriatr Soc.* Nov 1979;27(11):481-90.
85. Dokras A, Clifton S, Futterweit W, Wild R. Increased risk for abnormal depression scores in women with polycystic ovary syndrome: a systematic review and meta-analysis. *Obstet Gynecol.* Jan 2011;117(1):145-52. doi:10.1097/AOG.0b013e318202b0a4
86. Almeida OP, Yeap BB, Hankey GJ, Jamrozik K, Flicker L. Low free testosterone concentration as a potentially treatable cause of depressive symptoms in older men. *Archives of General Psychiatry.* 2008;65(3):283-289. doi:10.1001/archgenpsychiatry.2007.33 [doi]
87. Jankowska EA, Drohomirecka A, Ponikowska B, et al. Deficiencies in circulating testosterone and dehydroepiandrosterone sulphate, and depression in men with systolic chronic heart failure. *European journal of heart failure.* 2010;12(9):966-973. doi:10.1093/eurjhf/hfq108 [doi]
88. Joshi D, van Schoor NM, de Ronde W, et al. Low free testosterone levels are associated with prevalence and incidence of depressive symptoms in older men. *Clinical endocrinology.* 2010;72(2):232-240. doi:10.1111/j.1365-2265.2009.03641.x [doi]
89. Kushnir MM, Rockwood AL, Bergquist J. Liquid chromatography-tandem mass spectrometry applications in endocrinology. *Mass Spectrom Rev.* 2010 May-Jun 2010;29(3):480-502. doi:10.1002/mas.20264
90. Kelly WF, Kelly MJ, Faragher B. A prospective study of psychiatric and psychological aspects of Cushing's syndrome. *Clin Endocrinol (Oxf).* Dec 1996;45(6):715-20.
91. Kelly WF. Psychiatric aspects of Cushing's syndrome. *QJM.* Jul 1996;89(7):543-51. doi:10.1093/qjmed/89.7.543
92. Kushnir MM, Blamires T, Rockwood AL, et al. Liquid chromatography-tandem mass spectrometry assay for androstenedione, dehydroepiandrosterone, and testosterone with pediatric and adult reference intervals. *Clin Chem.* Jul 2010;56(7):1138-47. doi:10.1373/clinchem.2010.143222
93. Ford ES, Giles WH, Dietz WH. Prevalence of the metabolic syndrome among US adults: findings from the third National Health and Nutrition Examination Survey. *Jama.* 2002;287(3):356-359.
94. Schiffer L, Kempegowda P, Arlt W, O'Reilly MW. MECHANISMS IN ENDOCRINOLOGY: The sexually dimorphic role of androgens in human metabolic disease. *European journal of endocrinology.* 2017;177(3):R125-R143. doi:10.1530/EJE-17-0124 [doi]
95. Laaksonen DE, Niskanen L, Punnonen K, et al. Testosterone and sex hormone-binding globulin predict the metabolic syndrome and diabetes in middle-aged men. *Diabetes care.* 2004;27(5):1036-1041.
96. Brand JS, van der Tweel I, Grobbee DE, Emmelot-Vonk MH, van der Schouw YT. Testosterone, sex hormone-binding globulin and the metabolic syndrome: a systematic review and meta-analysis of observational studies. *International journal of epidemiology.* 2014;40(1):189-207. doi:10.1093/ije/dyq158 [doi]

97. Guducu N, Gormus U, Kutay SS, Kavak ZN, Telatar B. Endogenous sex hormones and their associations with cardiovascular risk factors in post-menopausal women. *J Endocrinol Invest*. 2013;36(8):588-592.
98. Penninx BWJH. Depression and cardiovascular disease: epidemiological evidence on their linking mechanisms. *Neuroscience & Biobehavioral Reviews*. 2016;
99. Delitala AP, Capobianco G, Delitala G, Cherchi PL, Dessole S. Polycystic ovary syndrome, adipose tissue and metabolic syndrome. *Archives of Gynecology and Obstetrics*. 2017;doi:10.1007/s00404-017-4429-2 [doi]
100. Stefansson J, Chatzittofis A, Nordström P, Arver S, Åsberg M, Jokinen J. CSF and plasma testosterone in attempted suicide. *Psychoneuroendocrinology*. 12 2016;74:1-6. doi:10.1016/j.psyneuen.2016.08.009
101. Markianos M, Tripodianakis J, Istikoglou C, et al. Suicide attempt by jumping: a study of gonadal axis hormones in male suicide attempters versus men who fell by accident. *Psychiatry Res*. Nov 2009;170(1):82-5. doi:10.1016/j.psychres.2008.08.001
102. Tripodianakis J, Markianos M, Rouvali O, Istikoglou C. Gonadal axis hormones in psychiatric male patients after a suicide attempt. *Eur Arch Psychiatry Clin Neurosci*. Apr 2007;257(3):135-9. doi:10.1007/s00406-006-0686-y
103. Zhang J, Jia CX, Wang LL. Testosterone differs between suicide attempters and community controls in men and women of China. *Physiol Behav*. Mar 2015;141:40-5. doi:10.1016/j.physbeh.2015.01.004
104. Perez-Rodriguez MM, Lopez-Castroman J, Martinez-Vigo M, et al. Lack of association between testosterone and suicide attempts. *Neuropsychobiology*. 2011;63(2):125-30. doi:10.1159/000318085
105. Butterfield MI, Stechuchak KM, Connor KM, et al. Neuroactive steroids and suicidality in posttraumatic stress disorder. *Am J Psychiatry*. Feb 2005;162(2):380-2. doi:10.1176/appi.ajp.162.2.380
106. Sher L, Grunebaum MF, Sullivan GM, et al. Testosterone levels in suicide attempters with bipolar disorder. *J Psychiatr Res*. Oct 2012;46(10):1267-71. doi:10.1016/j.jpsychires.2012.06.016
107. Sher L, Grunebaum MF, Sullivan GM, et al. Association of testosterone levels and future suicide attempts in females with bipolar disorder. *J Affect Disord*. Sep 2014;166:98-102. doi:10.1016/j.jad.2014.04.068
108. Hale GE, Robertson DM, Burger HG. The perimenopausal woman: endocrinology and management. *J Steroid Biochem Mol Biol*. Jul 2014;142:121-31. doi:10.1016/j.jsbmb.2013.08.015
109. Gordon JL, Girdler SS, Meltzer-Brody SE, et al. Ovarian hormone fluctuation, neurosteroids, and HPA axis dysregulation in perimenopausal depression: a novel heuristic model. *Am J Psychiatry*. Mar 2015;172(3):227-36. doi:10.1176/appi.ajp.2014.14070918
110. Maki PM, Kornstein SG, Joffe H, et al. Guidelines for the Evaluation and Treatment of Perimenopausal Depression: Summary and Recommendations. *J Womens Health (Larchmt)*. Feb 2019;28(2):117-134. doi:10.1089/jwh.2018.27099.mensocrec
111. Freeman EW, Sammel MD, Lin H, Gracia CR, Kapoor S. Symptoms in the menopausal transition: hormone and behavioral correlates. *Obstet Gynecol*. Jan 2008;111(1):127-36. doi:10.1097/01.AOG.0000295867.06184.b1

112. Bromberger JT, Assmann SF, Avis NE, Schocken M, Kravitz HM, Cordal A. Persistent mood symptoms in a multiethnic community cohort of pre- and perimenopausal women. *Am J Epidemiol*. Aug 2003;158(4):347-56. doi:10.1093/aje/kwg155
113. Fava GA, Grandi S, Canestrari R, Molnar G. Prodromal symptoms in primary major depressive disorder. *J Affect Disord*. Jun 1990;19(2):149-52.
114. Fava GA, Tossani E. Prodromal stage of major depression. *Early Interv Psychiatry*. Feb 2007;1(1):9-18. doi:10.1111/j.1751-7893.2007.00005.x
115. Hernandez M, Marangoni C, Grant MC, Estrada J, Faedda GL. Parental Reports of Prodromal Psychopathology in Pediatric Bipolar Disorder. *Curr Neuropsychopharmacol*. 2017;15(3):380-385. doi:10.2174/1570159X14666160801162046
116. Freeman EW, Sammel MD, Lin H, Nelson DB. Associations of hormones and menopausal status with depressed mood in women with no history of depression. *Arch Gen Psychiatry*. Apr 2006;63(4):375-82. doi:10.1001/archpsyc.63.4.375
117. Gordon JL, Eisenlohr-Moul TA, Rubinow DR, Schrubbe L, Girdler SS. Naturally Occurring Changes in Estradiol Concentrations in the Menopause Transition Predict Morning Cortisol and Negative Mood in Perimenopausal Depression. *Clin Psychol Sci*. Sep 2016;4(5):919-935. doi:10.1177/2167702616647924
118. Gordon JL, Rubinow DR, Eisenlohr-Moul TA, Leserman J, Girdler SS. Estradiol variability, stressful life events, and the emergence of depressive symptomatology during the menopausal transition. *Menopause*. Mar 2016;23(3):257-66. doi:10.1097/GME.0000000000000528
119. Bromberger JT, Kravitz HM, Chang YF, Cyranowski JM, Brown C, Matthews KA. Major depression during and after the menopausal transition: Study of Women's Health Across the Nation (SWAN). *Psychol Med*. Sep 2011;41(9):1879-88. doi:10.1017/S003329171100016X
120. Avis NE, Crawford S, Stellato R, Longcope C. Longitudinal study of hormone levels and depression among women transitioning through menopause. *Climacteric*. Sep 2001;4(3):243-9.
121. Woods NF, Smith-DiJulio K, Percival DB, Tao EY, Mariella A, Mitchell S. Depressed mood during the menopausal transition and early postmenopause: observations from the Seattle Midlife Women's Health Study. *Menopause (New York, NY)*. 2008;15(2):223-232. doi:10.1097/gme.0b013e3181450fc2 [doi]
122. Cheslack-Postava K, Keyes KM, Lowe SR, Koenen KC. Oral contraceptive use and psychiatric disorders in a nationally representative sample of women. *Arch Womens Ment Health*. Feb 2015;18(1):103-11. doi:10.1007/s00737-014-0453-4
123. McKetta S, Keyes KM. Oral contraceptive use and depression among adolescents. *Ann Epidemiol*. 01 2019;29:46-51. doi:10.1016/j.annepidem.2018.10.002
124. Skovlund CW, Mørch LS, Kessing LV, Lidegaard O. Association of Hormonal Contraception With Depression. *JAMA psychiatry*. 2016;doi:10.1001/jamapsychiatry.2016.2387 [doi]
125. Anderl C, Li G, Chen FS. Oral contraceptive use in adolescence predicts lasting vulnerability to depression in adulthood. *J Child Psychol Psychiatry*. Feb 2020;61(2):148-156. doi:10.1111/jcpp.13115

126. Schaffir J, Worly BL, Gur TL. Combined hormonal contraception and its effects on mood: a critical review. *Eur J Contracept Reprod Health Care*. Oct 2016;21(5):347-55. doi:10.1080/13625187.2016.1217327
127. Poromaa IS, Segebladh B. Adverse mood symptoms with oral contraceptives. *Acta Obstet Gynecol Scand*. Apr 2012;91(4):420-7. doi:10.1111/j.1600-0412.2011.01333.x
128. Hankin BL, Abramson LY, Moffitt TE, Silva PA, McGee R, Angell KE. Development of depression from preadolescence to young adulthood: emerging gender differences in a 10-year longitudinal study. *J Abnorm Psychol*. Feb 1998;107(1):128-40.
129. Schmidt PJ, Nieman LK, Danaceau MA, Adams LF, Rubinow DR. Differential behavioral effects of gonadal steroids in women with and in those without premenstrual syndrome. *N Engl J Med*. Jan 1998;338(4):209-16. doi:10.1056/NEJM199801223380401
130. Peters HP, Brossard D, de Cheveigné S, et al. Science communication. Interactions with the mass media. *Science*. Jul 2008;321(5886):204-5. doi:10.1126/science.1157780
131. Peters HP. Gap between science and media revisited: scientists as public communicators. *Proc Natl Acad Sci U S A*. Aug 2013;110 Suppl 3:14102-9. doi:10.1073/pnas.1212745110
132. Matthews S. Are YOU on the Pill? You're more likely to be depressed: Women who use contraception are up to 70 per cent more likely to be on antidepressants. *MailOnline*. <https://www.dailymail.co.uk/health/article-3812006/Are-contraception-likely-depressed-Women-use-skin-patches-70-cent-likely-depressed.html>
133. Viswanath K, Blake KD, Meissner HI, et al. Occupational practices and the making of health news: a national survey of US Health and medical science journalists. *J Health Commun*. Dec 2008;13(8):759-77. doi:10.1080/10810730802487430
134. Sumner P, Vivian-Griffiths S, Boivin J, et al. Exaggerations and Caveats in Press Releases and Health-Related Science News. *PLoS One*. 2016;11(12):e0168217. doi:10.1371/journal.pone.0168217
135. Penninx BW, Beekman AT, Smit JH, et al. The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *International journal of methods in psychiatric research*. 2008;17(3):121-140. doi:10.1002/mpr.256 [doi]
136. Comijs HC, van Marwijk HW, van der Mast RC, et al. The Netherlands study of depression in older persons (NESDO); a prospective cohort study. *BMC Research Notes*. 2011;4(1):524.
137. de Winter AF, Oldehinkel AJ, Veenstra R, Brunnekreef JA, Verhulst FC, Ormel J. Evaluation of non-response bias in mental health determinants and outcomes in a large sample of pre-adolescents. *European journal of epidemiology*. 2005;20(2):173-181.
138. Genazzani AR, Pluchino N, Freschi L, Ninni F, Luisi M. Androgens and the brain. *Maturitas*. 2007;57(1):27-30. doi:S0378-5122(07)00067-9 [pii]
139. Hofer P, Lanzenberger R, Kasper S. Testosterone in the brain: neuroimaging findings and the potential role for neuropsychopharmacology. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology*. 2013;23(2):79-88. doi:10.1016/j.euroneuro.2012.04.013 [doi]
140. de Ronde W, van der Schouw YT, Pols HA, et al. Calculation of bioavailable and free testosterone in men: a comparison of 5 published algorithms. *Clinical chemistry*. 2006;52(9):1777-1784. doi:clinchem.2005.063354 [pii]

141. Vermeulen A, Verdonck L, Kaufman JM. A critical evaluation of simple methods for the estimation of free testosterone in serum. *The Journal of clinical endocrinology and metabolism*. 1999;84(10):3666-3672. doi:10.1210/jcem.84.10.6079 [doi]
142. Arregger AL, Contreras LN, Tumilasci OR, Aquilano DR, Cardoso EM. Salivary testosterone: a reliable approach to the diagnosis of male hypogonadism. *Clinical endocrinology*. 2007;67(5):656-662. doi:CEN2937 [pii]
143. Cardoso EM, Contreras LN, Tumilasci EG, et al. Salivary testosterone for the diagnosis of androgen deficiency in end-stage renal disease. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association*. 2011;26(2):677-683. doi:10.1093/ndt/gfq439 [doi]
144. Lane AR, Hackney AC. Relationship between salivary and serum testosterone levels in response to different exercise intensities. *Hormones (Athens, Greece)*. 2014;doi:10.14310/horm.2002.1561 [doi]
145. Shirtcliff EA, Granger DA, Likos A. Gender differences in the validity of testosterone measured in saliva by immunoassay. *Hormones and behavior*. 2002;42(1):62-69. doi:S0018506X02917988 [pii]
146. Vittek J, L'Homedieu DG, Gordon GG, Rappaport SC, Southren AL. Direct radioimmunoassay (RIA) of salivary testosterone: correlation with free and total serum testosterone. *Life Sciences*. 1985;37(8):711-716.
147. Cadore E, Lhullier F, Brentano M, et al. Correlations between serum and salivary hormonal concentrations in response to resistance exercise. *Journal of sports sciences*. 2008;26(10):1067-1072. doi:10.1080/02640410801919526 [doi]
148. Gonzalez-Sanchez V, Moreno-Perez O, Garcia de Guadiana L, et al. Reference ranges for serum and salivary testosterone in young men of Mediterranean region. *Endocrinología y nutrición : organo de la Sociedad Espanola de Endocrinología y Nutrición*. 2015;62(1):4-10. doi:10.1016/j.endonu.2014.09.002 [doi]
149. Fiers T, Delanghe J, T'Sjoen G, Van Caenegem E, Wierckx K, Kaufman JM. A critical evaluation of salivary testosterone as a method for the assessment of serum testosterone. *Steroids*. 2014;86:5-9. doi:10.1016/j.steroids.2014.04.013 [doi]
150. Flyckt RL, Liu J, Frasure H, Wekselman K, Buch A, Kingsberg SA. Comparison of salivary versus serum testosterone levels in postmenopausal women receiving transdermal testosterone supplementation versus placebo. *Menopause (New York, NY)*. 2009;16(4):680-688. doi:10.1097/gme.0b013e318199d5c4 [doi]
151. Szydlarska D, Grzesiuk W, Kondracka A, Bartoszewicz Z, Bar-Andziak E. Measuring salivary androgens as a useful tool in the diagnosis of polycystic ovary syndrome. *Endokrynologia Polska*. 2012;63(3):183-190.
152. Baxendale PM, Jacobs HS, James VH. Salivary testosterone: relationship to unbound plasma testosterone in normal and hyperandrogenic women. *Clinical endocrinology*. 1982;16(6):595-603.
153. Granger DA, Schwartz EB, Booth A, Arentz M. Salivary testosterone determination in studies of child health and development. *Hormones and behavior*. 1999;35(1):18-27. doi:S0018-506X(98)91492-1 [pii]
154. Rilling JK, Worthman CM, Campbell BC, Stallings JF, Mbizva M. Ratios of plasma and salivary testosterone throughout puberty: production versus bioavailability. *Steroids*. 1996;61(6):374-378. doi:0039-128X(96)00043-8 [pii]

155. Giltay EJ, Enter D, Zitman FG, et al. Salivary testosterone: associations with depression, anxiety disorders, and antidepressant use in a large cohort study. *Journal of psychosomatic research*. 2012;72(3):205-213. doi:10.1016/j.jpsychores.2011.11.014 [doi]
156. Craig CL, Marshall AL, Sjoström M, et al. International physical activity questionnaire: 12-country reliability and validity. *Medicine and science in sports and exercise*. 2003;35(8):1381-1395. doi:10.1249/01.MSS.0000078924.61453.FB [doi]
157. Granger DA, Shirtcliff EA, Booth A, Kivlighan KT, Schwartz EB. The "trouble" with salivary testosterone. *Psychoneuroendocrinology*. 2004;29(10):1229-1240. doi:10.1016/j.psyneuen.2004.02.005 [doi]
158. Groschl M, Rauh M. Influence of commercial collection devices for saliva on the reliability of salivary steroids analysis. *Steroids*. 2006;71(13-14):1097-1100. doi:S0039-128X(06)00195-4 [pii]
159. Smeets-Janssen MM, Roelofs K, van Pelt J, et al. Salivary Testosterone Is Consistently and Positively Associated with Extraversion: Results from The Netherlands Study of Depression and Anxiety. *Neuropsychobiology*. 2015;71(2):76-84. doi:000369024 [pii]
160. Dabbs JM, Jr. Salivary testosterone measurements: reliability across hours, days, and weeks. *Physiology & Behavior*. 1990;48(1):83-86.
161. Diver MJ, Imtiaz KE, Ahmad AM, Vora JP, Fraser WD. Diurnal rhythms of serum total, free and bioavailable testosterone and of SHBG in middle-aged men compared with those in young men. *Clinical endocrinology*. 2003;58(6):710-717. doi:1772 [pii]
162. Gordon JL, Girdler SS, Meltzer-Brody SE, et al. Ovarian hormone fluctuation, neurosteroids, and HPA axis dysregulation in perimenopausal depression: a novel heuristic model. *The American Journal of Psychiatry*. 2015;172(3):227-236. doi:10.1176/appi.ajp.2014.14070918 [doi]
163. Rautenberg MW, Lentjes EGWM. Bepaling fertiliteitshormonen anno 2006. *Nederlands Tijdschrift voor Klinische Chemie en Laboratoriumgeneeskunde*. 2007;32(1):42-43-46.
164. Thijssen JHH. Bepalingen van geslachtshormonen in bloed anno 2000: juistheid en standaardisatie. *Nederlands Tijdschrift voor Klinische Chemie en Laboratoriumgeneeskunde*. 2000;25(6):346-347-351.
165. Hsing AW, Stanczyk FZ, Belanger A, et al. Reproducibility of serum sex steroid assays in men by RIA and mass spectrometry. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2007;16(5):1004-1008. doi:16/5/1004 [pii]
166. Ketha H, Kaur S, Grebe SK, Singh RJ. Clinical applications of LC-MS sex steroid assays: evolution of methodologies in the 21st century. *Current opinion in endocrinology, diabetes, and obesity*. 2014;21(3):217-226. doi:10.1097/MED.0000000000000068 [doi]
167. Huhtaniemi IT, Tajar A, Lee DM, et al. Comparison of serum testosterone and estradiol measurements in 3174 European men using platform immunoassay and mass spectrometry; relevance for the diagnostics in aging men. *European journal of endocrinology / European Federation of Endocrine Societies*. 2012;166(6):983-991. doi:10.1530/EJE-11-1051 [doi]

168. Khosla S, Amin S, Singh RJ, Atkinson EJ, Melton LJ, 3rd, Riggs BL. Comparison of sex steroid measurements in men by immunoassay versus mass spectroscopy and relationships with cortical and trabecular volumetric bone mineral density. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA*. 2008;19(10):1465-1471. doi:10.1007/s00198-008-0591-5 [doi]
169. Noppe G, de Rijke YB, Dorst K, van den Akker EL, van Rossum EF. LC-MS/MS-based method for long-term steroid profiling in human scalp hair. *Clinical endocrinology*. 2015;doi:10.1111/cen.12781 [doi]
170. Burger HG. Androgen production in women. *Fertil Steril*. Apr 2002;77 Suppl 4:S3-5. doi:10.1016/s0015-0282(02)02985-0
171. Marceau K, Ruttle PL, Shirtcliff EA, Essex MJ, Susman EJ. Developmental and contextual considerations for adrenal and gonadal hormone functioning during adolescence: Implications for adolescent mental health. *Dev Psychobiol*. Sep 2015;57(6):742-68. doi:10.1002/dev.21214
172. Noordzij M, Dekker FW, Zoccali C, Jager KJ. Measures of disease frequency: prevalence and incidence. *Nephron Clin Pract*. 2010;115(1):c17-20. doi:10.1159/000286345
173. Bromberger JT, Schott LL, Kravitz HM, et al. Longitudinal change in reproductive hormones and depressive symptoms across the menopausal transition: results from the Study of Women's Health Across the Nation (SWAN). *Archives of General Psychiatry*. 2010;67(6):598-607. doi:10.1001/archgenpsychiatry.2010.55 [doi]
174. Baischer W, Koinig G, Hartmann B, Huber J, Langer G. Hypothalamic-pituitary-gonadal axis in depressed premenopausal women: elevated blood testosterone concentrations compared to normal controls. *Psychoneuroendocrinology*. 1995;20(5):553-9.
175. Weber B, Lewicka S, Deuschle M, Colla M, Heuser I. Testosterone, androstenedione and dihydrotestosterone concentrations are elevated in female patients with major depression. *Psychoneuroendocrinology*. 2000;25(8):765-771. doi:S0306453000000238 [pii]
176. Oulis P, Masdrakis VG, Markianos M. Testosterone and dehydroepiandrosterone sulfate in female anxious and non-anxious major depression. *Int J Psychiatry Clin Pract*. Jan 2014;18(1):21-4. doi:10.3109/13651501.2013.845222
177. Matsuzaka H, Maeshima H, Kida S, et al. Gender differences in serum testosterone and cortisol in patients with major depressive disorder compared with controls. *International journal of psychiatry in medicine*. 2013;46(2):203-221.
178. Gallicchio L, Schilling C, Miller SR, Zacur H, Flaws JA. Correlates of depressive symptoms among women undergoing the menopausal transition. *J Psychosom Res*. Sep 2007;63(3):263-8. doi:10.1016/j.jpsychores.2007.02.003
179. Asselmann E, Kische H, Haring R, et al. Prospective associations of androgens and sex hormone-binding globulin with 12-month, lifetime and incident anxiety and depressive disorders in men and women from the general population. *J Affect Disord*. 02 2019;245:905-911. doi:10.1016/j.jad.2018.11.052
180. Kische H, Gross S, Wallaschofski H, et al. Associations of androgens with depressive symptoms and cognitive status in the general population. *PLoS One*. 2017;12(5):e0177272. doi:10.1371/journal.pone.0177272

181. Kische H, Pieper L, Venz J, et al. Longitudinal change instead of baseline testosterone predicts depressive symptoms. *Psychoneuroendocrinology*. Mar 2018;89:7-12. doi:10.1016/j.psyneuen.2017.12.013
182. Giltay EJ, van der Mast RC, Lauwen E, Heijboer AC, de Waal MWM, Comijs HC. Plasma Testosterone and the Course of Major Depressive Disorder in Older Men and Women. *The American Journal of Geriatric Psychiatry*. 2017;25(4):425-437.
183. World Health Organization Collaborating Centre for Drug Statistics M. *Anatomical Therapeutic Chemical (ATC) Classification*. 2007.
184. Vreeburg SA, Hoogendijk WJ, van Pelt J, et al. Major depressive disorder and hypothalamic-pituitary-adrenal axis activity: results from a large cohort study. *Arch Gen Psychiatry*. Jun 2009;66(6):617-26. doi:10.1001/archgenpsychiatry.2009.50
185. Wittchen HU. Reliability and validity studies of the WHO--Composite International Diagnostic Interview (CIDI): a critical review. *Journal of psychiatric research*. 1994;28(1):57-84. doi:0022-3956(94)90036-1 [pii]
186. APA. *Diagnostic and statistical manual of mental disorders, fourth edition*. 4th ed. American Psychiatric Publishing; 2001.
187. Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychological medicine*. 1996;26(3):477-486.
188. Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods*. Aug 2008;40(3):879-91.
189. Lakens D. Equivalence Tests: A Practical Primer for. *Soc Psychol Personal Sci*. May 2017;8(4):355-362. doi:10.1177/1948550617697177
190. Copay AG, Subach BR, Glassman SD, Polly DW, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J*. 2007 Sep-Oct 2007;7(5):541-6. doi:10.1016/j.spinee.2007.01.008
191. Bromberger JT, Kravitz HM. Mood and menopause: findings from the Study of Women's Health Across the Nation (SWAN) over 10 years. *Obstetrics and gynecology clinics of North America*. 2011;38(3):609-625. doi:10.1016/j.ogc.2011.05.011 [doi]
192. Somboonporn W, Davis SR, Council NHaMR. Testosterone effects on the breast: implications for testosterone therapy for women. *Endocr Rev*. Jun 2004;25(3):374-88. doi:10.1210/er.2003-0016
193. Bromberger JT, Schott LL, Kravitz HM, et al. Longitudinal change in reproductive hormones and depressive symptoms across the menopausal transition: results from the Study of Women's Health Across the Nation (SWAN). *Arch Gen Psychiatry*. Jun 2010;67(6):598-607. doi:10.1001/archgenpsychiatry.2010.55
194. Hardeveld F, Spijker J, De Graaf R, Nolen WA, Beekman AT. Prevalence and predictors of recurrence of major depressive disorder in the adult population. *Acta Psychiatr Scand*. Sep 2010;122(3):184-91. doi:10.1111/j.1600-0447.2009.01519.x
195. Amiri M, Kabir A, Nahidi F, Shekofteh M, Ramezani Tehrani F. Effects of combined oral contraceptives on the clinical and biochemical parameters of hyperandrogenism in patients with polycystic ovary syndrome: a systematic review and meta-analysis. *Eur J Contracept Reprod Health Care*. Feb 2018;23(1):64-77. doi:10.1080/13625187.2018.1435779

196. Zimmerman Y, Eijkemans MJ, Coelingh Bennink HJ, Blankenstein MA, Fauser BC. The effect of combined oral contraception on testosterone levels in healthy women: a systematic review and meta-analysis. *Hum Reprod Update*. 2014 Jan-Feb 2014;20(1):76-105. doi:10.1093/humupd/dmt038
197. Skiba MA, Bell RJ, Islam RM, Handelsman DJ, Desai R, Davis SR. Androgens During the Reproductive Years: What Is Normal for Women? *J Clin Endocrinol Metab*. Nov 2019;104(11):5382-5392. doi:10.1210/jc.2019-01357
198. van der Veen A, van Faassen M, de Jong WHA, van Beek AP, Dijck-Brouwer DAJ, Kema IP. Development and validation of a LC-MS/MS method for the establishment of reference intervals and biological variation for five plasma steroid hormones. *Clin Biochem*. Jun 2019;68:15-23. doi:10.1016/j.clinbiochem.2019.03.013
199. Hore BD. Hypogonadism presenting as a psychiatric disorder. *Br J Psychiatry*. Jul 1969;115(524):863-4. doi:10.1192/bjp.115.524.863-b
200. Höfer P, Lanzenberger R, Kasper S. Testosterone in the brain: neuroimaging findings and the potential role for neuropsychopharmacology. *Eur Neuropsychopharmacol*. Feb 2013;23(2):79-88. doi:10.1016/j.euroneuro.2012.04.013
201. Wu FC, Tajar A, Beynon JM, et al. Identification of late-onset hypogonadism in middle-aged and elderly men. *N Engl J Med*. Jul 2010;363(2):123-35. doi:10.1056/NEJMoa0911101
202. Corona G, Goulis DG, Huhtaniemi I, et al. European Academy of Andrology (EAA) guidelines on investigation, treatment and monitoring of functional hypogonadism in males: Endorsing organization: European Society of Endocrinology. *Andrology*. 09 2020;8(5):970-987. doi:10.1111/andr.12770
203. Bhasin S, Brito JP, Cunningham GR, et al. Testosterone Therapy in Men With Hypogonadism: An Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab*. 05 2018;103(5):1715-1744. doi:10.1210/jc.2018-00229
204. Markianos M, Tripodianakis J, Sarantidis D, Hatzimanolis J. Plasma testosterone and dehydroepiandrosterone sulfate in male and female patients with dysthymic disorder. *J Affect Disord*. Aug 2007;101(1-3):255-8. doi:10.1016/j.jad.2006.11.013
205. Seidman SN, Araujo AB, Roose SP, et al. Low testosterone levels in elderly men with dysthymic disorder. *Am J Psychiatry*. Mar 2002;159(3):456-9. doi:10.1176/appi.ajp.159.3.456
206. Barrett-Connor E, Von Muhlen DG, Kritz-Silverstein D. Bioavailable testosterone and depressed mood in older men: the Rancho Bernardo Study. *The Journal of clinical endocrinology and metabolism*. 1999;84(2):573-577. doi:10.1210/jcem.84.2.5495 [doi]
207. Monteagudo PT, Falcão AA, Verreschi IT, Zanella MT. The imbalance of sex-hormones related to depressive symptoms in obese men. *Aging Male*. 2016;19(1):20-6. doi:10.3109/13685538.2015.1084500
208. T'Sjoen GG, De Vos S, Goemaere S, et al. Sex steroid level, androgen receptor polymorphism, and depressive symptoms in healthy elderly men. *Journal of the American Geriatrics Society*. 2005;53(4):636-642. doi:10.1111/j.1532-5415.2005.03212.x [pii]
209. McIntyre RS, Mancini D, Eisfeld BS, et al. Calculated bioavailable testosterone levels and depression in middle-aged men. *Psychoneuroendocrinology*. 2006;31(9):1029-1035. doi:10.1016/j.psyneuen.2006.06.001 [pii]

210. Hintikka J, Niskanen L, Koivumaa-Honkanen H, et al. Hypogonadism, decreased sexual desire, and long-term depression in middle-aged men. *J Sex Med.* Jul 2009;6(7):2049-57. doi:10.1111/j.1743-6109.2009.01299.x
211. Shores MM, Sloan KL, Matsumoto AM, Mocer VM, Felker B, Kivlahan DR. Increased incidence of diagnosed depressive illness in hypogonadal older men. *Archives of General Psychiatry.* 2004;61(2):162-167. doi:10.1001/archpsyc.61.2.162 [doi]
212. Shores MM, Mocer VM, Sloan KL, Matsumoto AM, Kivlahan DR. Low testosterone levels predict incident depressive illness in older men: effects of age and medical morbidity. *The Journal of clinical psychiatry.* 2005;66(1):7-14.
213. Gao W, Bohl CE, Dalton JT. Chemistry and structural biology of androgen receptor. *Chem Rev.* Sep 2005;105(9):3352-70. doi:10.1021/cr020456u
214. Barrett-Connor E, von Mühlen D, Laughlin GA, Kripke A. Endogenous levels of dehydroepiandrosterone sulfate, but not other sex hormones, are associated with depressed mood in older women: the Rancho Bernardo Study. *J Am Geriatr Soc.* Jun 1999;47(6):685-91.
215. Wang C, Shiraishi S, Leung A, et al. Corrigendum to "Validation of a testosterone and dihydrotestosterone liquid chromatography tandem mass spectrometry assay: interference and comparison with established methods" [Steroids Volume 73, Issue 13, 12 December 2008, Pages 1345-1352]. *Steroids.* 07 2018;135:108. doi:10.1016/j.steroids.2018.05.006
216. Bhasin S, Seidman S. Testosterone Treatment of Depressive Disorders in Men: Too Much Smoke, Not Enough High-Quality Evidence. *JAMA Psychiatry.* Nov 2018;doi:10.1001/jamapsychiatry.2018.2661
217. Booth A, Johnson DR, Granger DA. Testosterone and men's depression: the role of social behavior. *J Health Soc Behav.* Jun 1999;40(2):130-40.
218. Cohen J. Statistical power analysis for the behavioral sciences. Revised edition ed. London: Academic Press; 1977.
219. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics.* 2001;29(4):1165-1188, 24.
220. Rastrelli G, Corona G, Tarocchi M, Mannucci E, Maggi M. How to define hypogonadism? Results from a population of men consulting for sexual dysfunction. *J Endocrinol Invest.* Apr 2016;39(4):473-84. doi:10.1007/s40618-015-0425-1
221. Anacker C, Zunszain PA, Carvalho LA, Pariante CM. The glucocorticoid receptor: pivot of depression and of antidepressant treatment? *Psychoneuroendocrinology.* Apr 2011;36(3):415-25. doi:10.1016/j.psyneuen.2010.03.007
222. Schweikert H-U. Estrogen in the Male: Nature, Sources, and Biological Effects. *Reference Module in Neuroscience and Biobehavioral Psychology.* Elsevier; 2019.
223. Gao W, Kirschbaum C, Grass J, Stalder T. LC-MS based analysis of endogenous steroid hormones in human hair. *J Steroid Biochem Mol Biol.* 09 2016;162:92-9. doi:10.1016/j.jsbmb.2015.12.022
224. Grundy SM, Cleeman JI, Daniels SR, et al. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. *Circulation.* 2005;112(17):2735-2752. doi:CIRCULATIONAHA.105.169404 [pii]

225. Durdiakova J, Ostatnikova D, Celec P. Testosterone and its metabolites--modulators of brain functions. *Acta Neurobiologiae Experimentalis*. 2011;71(4):434-454. doi:7147 [pii]
226. Chrysoshoou C, Panagiotakos D, Pitsavos C, et al. Low total testosterone levels are associated with the metabolic syndrome in elderly men: the role of body weight, lipids, insulin resistance, and inflammation; the Ikaria study. *The review of diabetic studies : RDS*. 2013;10(1):27-38. doi:10.1900/RDS.2013.10.27 [doi]
227. Kochanek KD, Murphy SL, Xu J, Tejada-Vera B. Deaths: Final Data for 2014. *National Vital Statistics Reports*. 2016;65(4)
228. Nicholson A, Kuper H, Hemingway H. Depression as an aetiologic and prognostic factor in coronary heart disease: a meta-analysis of 6362 events among 146 538 participants in 54 observational studies. *European heart journal*. 2006;27(23):2763-2774.
229. Mulvihill JS, Nicol GE, Dixon D, et al. Effect of Metabolic Syndrome on Late-Life Depression: Associations with Disease Severity and Treatment Resistance. (1532-5415 (Electronic))
230. Chung S-D, Kao L-T, Lin H-C, Xirasagar S, Huang C-C, Lee H-C. Patients receiving androgen deprivation therapy for prostate cancer have an increased risk of depressive disorder. *PloS one*. 2017;12(3):e0173266.
231. Rohr UD. The impact of testosterone imbalance on depression and women's health. *Maturitas*. 2002;41:25-46.
232. Yehuda R, Brand SR, Golier JA, Yang RK. Clinical correlates of DHEA associated with post-traumatic stress disorder. *Acta Psychiatrica Scandinavica*. 2006;114(3):187-193. doi:ACP801 [pii]
233. Comijs HC, Nieuwesteeg J, Kok R, et al. The two-year course of late-life depression; results from the Netherlands study of depression in older persons. *BMC psychiatry*. 2015;15(1):20.
234. Groenestege WM, Bui HN, ten Kate J, et al. Accuracy of first and second generation testosterone assays and improvement through sample extraction. *Clinical chemistry*. 2012;58(7):1154-1156. doi:10.1373/clinchem.2011.181735 [doi]
235. Heijboer AC, Savelkoul E, Kruit A, Endert E, Blankenstein MA. Inaccurate first-generation testosterone assays are influenced by sex hormone-binding globulin concentrations. *The Journal of Applied Laboratory Medicine: An AACC Publication*. 2016;1(2):194-201.
236. Bays HE, Dujovne CA, McGovern ME, et al. Comparison of once-daily, niacin extended-release/lovastatin with standard doses of atorvastatin and simvastatin (the ADVICOR Versus Other Cholesterol-Modulating Agents Trial Evaluation [ADVOCATE]). *The American Journal of Cardiology*. 2003;91(6):667-672.
237. Grundy SM, Vega GL, Yuan Z, Battisti WP, Brady WE, Palmisano J. Effectiveness and tolerability of simvastatin plus fenofibrate for combined hyperlipidemia (the SAFARI trial). *The American Journal of Cardiology*. 2005;95(4):462-468.
238. Vogelzangs N, Suthers K, Ferrucci L, et al. Hypercortisolemic depression is associated with the metabolic syndrome in late-life. *Psychoneuroendocrinology*. 2007;32(2):151-159.

239. Group SCR. Prevention of Stroke by Antihypertensive Drug Treatment in Older Persons With Isolated Systolic Hypertension. Final Results of the Systolic Hypertension in the Elderly Program (SHEP). *JAMA*. 1991;265(24):3255-3264. doi:10.1001/jama.1991.03460240051027
240. Bouchard C. BMI, fat mass, abdominal adiposity and visceral fat: where is the 'beef'? *International journal of obesity*. 2007;31(10):1552.
241. Chubb SA, Hyde Z, Almeida OP, et al. Lower sex hormone-binding globulin is more strongly associated with metabolic syndrome than lower total testosterone in older men: the Health in Men Study. *European journal of endocrinology*. 2008;158(6):785-792. doi:10.1530/EJE-07-0893 [doi]
242. Brand JS, van der Tweel I, Grobbee DE, Emmelot-Vonk MH, van der Schouw YT. Testosterone, sex hormone-binding globulin and the metabolic syndrome: a systematic review and meta-analysis of observational studies. *Int J Epidemiol*. Feb 2011;40(1):189-207. doi:10.1093/ije/dyq158
243. Golden SH, Ding J, Szklo M, Schmidt MI, Duncan BB, Dobs A. Glucose and insulin components of the metabolic syndrome are associated with hyperandrogenism in postmenopausal women: the atherosclerosis risk in communities study. *American Journal of Epidemiology*. 2004;160(6):540-548.
244. Sutton-Tyrrell K, Wildman RP, Matthews KA, et al. Sex-hormone-binding globulin and the free androgen index are related to cardiovascular risk factors in multiethnic premenopausal and perimenopausal women enrolled in the Study of Women Across the Nation (SWAN). *Circulation*. 2005;111(10):1242-1249. doi:111/10/1242 [pii]
245. Weinberg ME, Manson JE, Buring JE, et al. Low sex hormone-binding globulin is associated with the metabolic syndrome in postmenopausal women. *Metabolism*. 2006;55(11):1473-1480.
246. Handelsman DJ. Free Testosterone: Pumping up the Tires or Ending the Free Ride? *Endocrine reviews*. 2017;38(4):297-301.
247. Hammond GL. Diverse roles for sex hormone-binding globulin in reproduction. *Biology of reproduction*. 2011;85(3):431-441.
248. Fenske B, Kische H, Gross S, et al. Endogenous Androgens and Sex Hormone-Binding Globulin in Women and Risk of Metabolic Syndrome and Type 2 Diabetes. *The Journal of Clinical Endocrinology & Metabolism*. 2015;100(12):4595-4603.
249. Haring R, Völzke H, Spielhagen C, Nauck M, Wallaschofski H. The role of sex hormone-binding globulin and testosterone in the risk of incident metabolic syndrome. *European journal of preventive cardiology*. 2013;20(6):1061-1068.
250. Haring R, Teumer A, Volker U, et al. Mendelian randomization suggests non-causal associations of testosterone with cardiometabolic risk factors and mortality. *Andrology*. Jan 2013;1(1):17-23. doi:10.1111/j.2047-2927.2012.00002.x
251. Davis SR, Wahlin-Jacobsen S. Testosterone in women—the clinical significance. *The Lancet Diabetes & Endocrinology*. 2015;3(12):980-992.
252. Penninx BW, Milaneschi Y, Lamers F, Vogelzangs N. Understanding the somatic consequences of depression: biological mechanisms and the role of depression symptom profile. *BMC medicine*. 2013;11:129-7015-11-129. doi:10.1186/1741-7015-11-129 [doi]

253. Toufexis D, Rivarola MA, Lara H, Viau V. Stress and the reproductive axis. *Journal of neuroendocrinology*. 2014;26(9):573-586.
254. Yeap BB, Alfonso H, Chubb SAP, et al. Reference ranges and determinants of testosterone, dihydrotestosterone, and estradiol levels measured using liquid chromatography-tandem mass spectrometry in a population-based cohort of older men. *The Journal of Clinical Endocrinology & Metabolism*. 2012;97(11):4030-4039.
255. Nock MK, Borges G, Bromet EJ, et al. Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. *Br J Psychiatry*. Feb 2008;192(2):98-105. doi:10.1192/bjp.bp.107.040113
256. WHO. *Preventing suicide: a global imperative*. 2014. <https://apps.who.int/iris/bitstream/handle/10665/131056/97892?sequence=1>
257. Hendin H, Lipschitz A, Maltzberger JT, Haas AP, Wyncoop S. Therapists' reactions to patients' suicides. *Am J Psychiatry*. Dec 2000;157(12):2022-7. doi:10.1176/appi.ajp.157.12.2022
258. Lenz B, Röther M, Bouna-Pyrrou P, Mühle C, Tektas OY, Kornhuber J. The androgen model of suicide completion. *Prog Neurobiol*. Jan 2019;172:84-103. doi:10.1016/j.pneurobio.2018.06.003
259. van Heeringen K, Mann JJ. The neurobiology of suicide. *Lancet Psychiatry*. Jun 2014;1(1):63-72. doi:10.1016/S2215-0366(14)70220-2
260. Ludwig B, Roy B, Wang Q, Birur B, Dwivedi Y. The Life Span Model of Suicide and Its Neurobiological Foundation. *Front Neurosci*. 2017;11:74. doi:10.3389/fnins.2017.00074
261. McCloskey MS, Ammerman BA. Suicidal behavior and aggression-related disorders. *Curr Opin Psychol*. 08 2018;22:54-58. doi:10.1016/j.copsyc.2017.08.010
262. Turecki G, Brent DA. Suicide and suicidal behaviour. *Lancet*. Mar 2016;387(10024):1227-39. doi:10.1016/S0140-6736(15)00234-2
263. Coccaro EF. Testosterone and Aggression: More Than Just Biology? *Biol Psychiatry*. Aug 15 2017;82(4):234. doi:10.1016/j.biopsych.2017.06.025
264. Carré JM, McCormick CM, Hariri AR. The social neuroendocrinology of human aggression. *Psychoneuroendocrinology*. Aug 2011;36(7):935-44. doi:10.1016/j.psyneuen.2011.02.001
265. Archer J, Graham-Kevan N, Davies M. Testosterone and aggression: A reanalysis of Book, Starzyk, and Quinsey's (2001) study. *Aggression and Violent Behavior*; 2004(10):241-261. doi:10.1016/j.javb.2004.01.001
266. Solis E, Antypa N, Conijn JM, Kelderman H, Van der Does W. Psychometric properties of the Leiden Index of Depression Sensitivity (LEIDS). *Psychol Assess*. Feb 2017;29(2):158-171. doi:10.1037/pas0000326
267. Bostwick JM, Pankratz VS. Affective disorders and suicide risk: a reexamination. *Am J Psychiatry*. Dec 2000;157(12):1925-32. doi:10.1176/appi.ajp.157.12.1925
268. Beck A, Steer R. *Beck scale for Suicide Ideation; Manual*. 1991.
269. Beck A, Kovacs M, Weissman A. Assessment of suicidal intention: the Scale for Suicide Ideation. *J Consult Clin Psychol*. Apr 1979;47(2):343-52.
270. Roland BC, Morris JL, Zelhart PF. Proposed relation of testosterone levels to male suicides and sudden deaths. *Psychol Rep*. Aug 1986;59(1):100-2. doi:10.2466/pr0.1986.59.1.100

271. Mobasseri N, Babaei F, Karimian M, Nikzad H. Androgen receptor (*EXCLI J*. 2018;17:1167-1179. doi:10.17179/excli2018-1744
272. Gruber AJ, Pope HG. Psychiatric and medical effects of anabolic-androgenic steroid use in women. *Psychother Psychosom*. 2000;69(1):19-26. doi:10.1159/000012362
273. Strauss RH, Liggett MT, Lanese RR. Anabolic steroid use and perceived effects in ten weight-trained women athletes. *JAMA*. May 1985;253(19):2871-3.
274. Onakomaiya MM, Henderson LP. Mad men, women and steroid cocktails: a review of the impact of sex and other factors on anabolic androgenic steroids effects on affective behaviors. *Psychopharmacology (Berl)*. Feb 2016;233(4):549-69. doi:10.1007/s00213-015-4193-6
275. Carre JM, Geniole SN, Ortiz TL, Bird BM, Videto A, Bonin PL. Exogenous Testosterone Rapidly Increases Aggressive Behavior in Dominant and Impulsive Men. *Biol Psychiatry*. Aug 15 2017;82(4):249-256. doi:10.1016/j.biopsych.2016.06.009
276. Dreher JC, Dunne S, Pazderska A, Frodl T, Nolan JJ, O'Doherty JP. Testosterone causes both prosocial and antisocial status-enhancing behaviors in human males. *Proc Natl Acad Sci U S A*. 10 2016;113(41):11633-11638. doi:10.1073/pnas.1608085113
277. Davison SL, Bell R, Donath S, Montalto JG, Davis SR. Androgen levels in adult females: changes with age, menopause, and oophorectomy. *J Clin Endocrinol Metab*. Jul 2005;90(7):3847-53. doi:10.1210/jc.2005-0212
278. Burger HG, Dudley EC, Cui J, Dennerstein L, Hopper JL. A prospective longitudinal study of serum testosterone, dehydroepiandrosterone sulfate, and sex hormone-binding globulin levels through the menopause transition. *J Clin Endocrinol Metab*. Aug 2000;85(8):2832-8. doi:10.1210/jcem.85.8.6740
279. Rothman MS, Carlson NE, Xu M, et al. Reexamination of testosterone, dihydrotestosterone, estradiol and estrone levels across the menstrual cycle and in postmenopausal women measured by liquid chromatography-tandem mass spectrometry. *Steroids*. Jan 2011;76(1-2):177-82. doi:10.1016/j.steroids.2010.10.010
280. Elmlinger MW, Kühnel W, Wormstall H, Döller PC. Reference intervals for testosterone, androstenedione and SHBG levels in healthy females and males from birth until old age. *Clin Lab*. 2005;51(11-12):625-32.
281. Kaufman JM, Lapauw B, Mahmoud A, T'Sjoen G, Huhtaniemi IT. Aging and the Male Reproductive System. *Endocr Rev*. 08 2019;40(4):906-972. doi:10.1210/er.2018-00178
282. Handelsman DJ, Yeap B, Flicker L, Martin S, Wittert GA, Ly LP. Age-specific population centiles for androgen status in men. *Eur J Endocrinol*. Dec 2015;173(6):809-17. doi:10.1530/EJE-15-0380
283. Jones ME, Schoemaker MJ, Rae M, et al. Reproducibility of estradiol and testosterone levels in postmenopausal women over 5 years: results from the breakthrough generations study. *Am J Epidemiol*. May 2014;179(9):1128-33. doi:10.1093/aje/kwu027
284. Hsing AW, Stanczyk FZ, Bélanger A, et al. Reproducibility of serum sex steroid assays in men by RIA and mass spectrometry. *Cancer Epidemiol Biomarkers Prev*. May 2007;16(5):1004-8. doi:10.1158/1055-9965.EPI-06-0792
285. Cauley JA, Gutai JP, Kuller LH, Powell JG. Reliability and interrelations among serum sex hormones in postmenopausal women. *Am J Epidemiol*. Jan 1991;133(1):50-7. doi:10.1093/oxfordjournals.aje.a115801

286. Rodríguez MR, Nuevo R, Chatterji S, Ayuso-Mateos JL. Definitions and factors associated with subthreshold depressive conditions: a systematic review. *BMC Psychiatry*. Oct 2012;12:181. doi:10.1186/1471-244X-12-181
287. Cuijpers P, van Straten A, Smit F, Mihalopoulos C, Beekman A. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. *Am J Psychiatry*. Oct 2008;165(10):1272-80. doi:10.1176/appi.ajp.2008.07091422
288. Worsley R, Bell R, Kulkarni J, Davis SR. The association between vasomotor symptoms and depression during perimenopause: a systematic review. *Maturitas*. Feb 2014;77(2):111-7. doi:10.1016/j.maturitas.2013.11.007
289. Joffe H, Hall JE, Soares CN, et al. Vasomotor symptoms are associated with depression in perimenopausal women seeking primary care. *Menopause*. 2002 Nov-Dec 2002;9(6):392-8.
290. Joffe H, Crawford SL, Freeman MP, et al. Independent Contributions of Nocturnal Hot Flashes and Sleep Disturbance to Depression in Estrogen-Deprived Women. *J Clin Endocrinol Metab*. Oct 2016;101(10):3847-3855. doi:10.1210/jc.2016-2348
291. Bromberger JT, Kravitz HM, Youk A, Schott LL, Joffe H. Patterns of depressive disorders across 13 years and their determinants among midlife women: SWAN mental health study. *J Affect Disord*. Dec 2016;206:31-40. doi:10.1016/j.jad.2016.07.005
292. Avis NE, Brambilla D, McKinlay SM, Vass K. A longitudinal analysis of the association between menopause and depression. Results from the Massachusetts Women's Health Study. *Ann Epidemiol*. May 1994;4(3):214-20. doi:10.1016/1047-2797(94)90099-x
293. Rupprecht R. Neuroactive steroids: mechanisms of action and neuropsychopharmacological properties. *Psychoneuroendocrinology*. Feb 2003;28(2):139-68. doi:10.1016/s0306-4530(02)00064-1
294. Joffe H, Cohen LS. Estrogen, serotonin, and mood disturbance: where is the therapeutic bridge? *Biol Psychiatry*. Nov 1998;44(9):798-811.
295. Gordon JL, Girdler SS, Meltzer-Brody SE, et al. Response to Eskola et al. *Am J Psychiatry*. Aug 2015;172(8):797. doi:10.1176/appi.ajp.2015.15030377r
296. Harlow SD, Gass M, Hall JE, et al. Executive summary of the Stages of Reproductive Aging Workshop + 10: addressing the unfinished agenda of staging reproductive aging. *J Clin Endocrinol Metab*. Apr 2012;97(4):1159-68. doi:10.1210/jc.2011-3362
297. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. Apr 1979;134:382-9.
298. Kellner R. A symptom questionnaire. *J Clin Psychiatry*. Jul 1987;48(7):268-74.
299. Gilbody S, Richards D, Brealey S, Hewitt C. Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. *J Gen Intern Med*. Nov 2007;22(11):1596-602. doi:10.1007/s11606-007-0333-y
300. Sarason IG, Johnson JH, Siegel JM. Assessing the impact of life changes: development of the Life Experiences Survey. *J Consult Clin Psychol*. Oct 1978;46(5):932-46.
301. Mundt JC, Katzelnick DJ, Kennedy SH, Eisfeld BS, Bouffard BB, Greist JH. Validation of an IVRS version of the MADRS. *J Psychiatr Res*. Apr 2006;40(3):243-6. doi:10.1016/j.jpsychires.2005.05.002

302. Davidson J, Turnbull CD, Strickland R, Miller R, Graves K. The Montgomery-Asberg Depression Scale: reliability and validity. *Acta Psychiatr Scand*. May 1986;73(5):544-8. doi:10.1111/j.1600-0447.1986.tb02723.x
303. Herrmann N, Black SE, Lawrence J, Szekely C, Szalai JP. The Sunnybrook Stroke Study: a prospective study of depressive symptoms and functional outcome. *Stroke*. Mar 1998;29(3):618-24. doi:10.1161/01.str.29.3.618
304. Nelson RE, Grebe SK, OKane DJ, Singh RJ. Liquid chromatography-tandem mass spectrometry assay for simultaneous measurement of estradiol and estrone in human plasma. *Clin Chem*. Feb 2004;50(2):373-84. doi:10.1373/clinchem.2003.025478
305. Siekmann L. Determination of oestradiol-17 beta in human serum by isotope dilution-mass spectrometry. Definitive methods in clinical chemistry, II. *J Clin Chem Clin Biochem*. Aug 1984;22(8):551-7.
306. Newton KM, Carpenter JS, Guthrie KA, et al. Methods for the design of vasomotor symptom trials: the menopausal strategies: finding lasting answers to symptoms and health network. *Menopause*. Jan 2014;21(1):45-58. doi:10.1097/GME.0b013e31829337a4
307. Herring AH. Applied Longitudinal Analysis, by Garrett M. Fitzmaurice, Nan M. Laird, and James H. Ware, John Wiley & Sons, 2011: ISBN 978-0-470-38027-7, 740 pp., \$125. Taylor & Francis; 2013.
308. Leiva R, Bouchard T, Boehringer H, Abulla S, Ecochard R. Random serum progesterone threshold to confirm ovulation. *Steroids*. Sep 2015;101:125-9. doi:10.1016/j.steroids.2015.06.013
309. Santoro N, Crawford SL, El Khoudary SR, et al. Menstrual Cycle Hormone Changes in Women Traversing Menopause: Study of Women's Health Across the Nation. *J Clin Endocrinol Metab*. 07 2017;102(7):2218-2229. doi:10.1210/jc.2016-4017
310. O'Connor KA, Ferrell R, Brindle E, et al. Progesterone and ovulation across stages of the transition to menopause. *Menopause*. 2009 Nov-Dec 2009;16(6):1178-87. doi:10.1097/gme.0b013e3181aa192d
311. Maki PM, Kornstein SG, Joffe H, et al. Guidelines for the evaluation and treatment of perimenopausal depression: summary and recommendations. *Menopause*. Oct 2018;25(10):1069-1085. doi:10.1097/GME.0000000000001174
312. Meltzer-Brody S, Colquhoun H, Riesenberger R, et al. Brexanolone injection in post-partum depression: two multicentre, double-blind, randomised, placebo-controlled, phase 3 trials. *Lancet*. 09 2018;392(10152):1058-1070. doi:10.1016/S0140-6736(18)31551-4
313. Payne JL, Palmer JT, Joffe H. A reproductive subtype of depression: conceptualizing models and moving toward etiology. *Harv Rev Psychiatry*. 2009;17(2):72-86. doi:10.1080/10673220902899706
314. Joffe H, Petrillo LF, Koukopoulos A, et al. Increased estradiol and improved sleep, but not hot flashes, predict enhanced mood during the menopausal transition. *J Clin Endocrinol Metab*. Jul 2011;96(7):E1044-54. doi:10.1210/jc.2010-2503
315. Demers LM. Testosterone and estradiol assays: current and future trends. *Steroids*. Dec 2008;73(13):1333-8. doi:10.1016/j.steroids.2008.05.002

316. Mauas V, Kopala-Sibley DC, Zuroff DC. Depressive symptoms in the transition to menopause: the roles of irritability, personality vulnerability, and self-regulation. *Arch Womens Ment Health*. Aug 2014;17(4):279-89. doi:10.1007/s00737-014-0434-7
317. Fava M, Hwang I, Rush AJ, Sampson N, Walters EE, Kessler RC. The importance of irritability as a symptom of major depressive disorder: results from the National Comorbidity Survey Replication. *Mol Psychiatry*. Aug 2010;15(8):856-67. doi:10.1038/mp.2009.20
318. Toohey MJ, DiGiuseppe R. Defining and measuring irritability: Construct clarification and differentiation. *Clin Psychol Rev*. Apr 2017;53:93-108. doi:10.1016/j.cpr.2017.01.009
319. Berk L, Hallam KT, Venugopal K, et al. Impact of irritability: a 2-year observational study of outpatients with bipolar I or schizoaffective disorder. *Bipolar Disord*. 2017;19(3):184-197. doi:10.1111/bdi.12486
320. Perlis RH, Fava M, Trivedi MH, et al. Irritability is associated with anxiety and greater severity, but not bipolar spectrum features, in major depressive disorder. *Acta Psychiatr Scand*. Apr 2009;119(4):282-9. doi:10.1111/j.1600-0447.2008.01298.x
321. Reinares M, Vieta E, Colom F, et al. What really matters to bipolar patients' caregivers: sources of family burden. *J Affect Disord*. Aug 2006;94(1-3):157-63. doi:10.1016/j.jad.2006.04.022
322. Joffe H, de Wit A, Coborn J, et al. Impact of Estradiol Variability and Progesterone on Mood in Perimenopausal Women With Depressive Symptoms. *J Clin Endocrinol Metab*. Mar 2020;105(3):doi:10.1210/clinem/dgz181
323. Bromberger JT, Meyer PM, Kravitz HM, et al. Psychologic distress and natural menopause: a multiethnic community study. *Am J Public Health*. Sep 2001;91(9):1435-42. doi:10.2105/ajph.91.9.1435
324. Benasi G, Fava GA, Rafanelli C. Kellner's Symptom Questionnaire, a Highly Sensitive Patient-Reported Outcome Measure: Systematic Review of Clinimetric Properties. *Psychother Psychosom*. 2020;89(2):74-89. doi:10.1159/000506110
325. Grusso P, Quatraro RM. Maternity blues in Italian primipara women: symptoms and mood states in the first fifteen days after childbirth. *Health Care Women Int*. Jul 2013;34(7):556-76. doi:10.1080/07399332.2012.708373
326. Helsel D. *Statistics for censored environmental data using minitab and r*. 2nd ed. Wiley; 2011.
327. McDowell I. *Measuring Health: A guide to rating scales and questionnaires*. 3rd ed. Oxford University Press; 2006.
328. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med*. Jul 2001;2(4):297-307.
329. Gordon JL, Sander B, Eisenlohr-Moul TA, Sykes Tottenham L. Mood sensitivity to estradiol predicts depressive symptoms in the menopause transition. *Psychol Med*. Mar 2020;1-9. doi:10.1017/S0033291720000483
330. Lundin C, Danielsson KG, Bixo M, et al. Combined oral contraceptive use is associated with both improvement and worsening of mood in the different phases of the treatment cycle-A double-blind, placebo-controlled randomized trial. *Psychoneuroendocrinology*. Feb 2017;76:135-143. doi:10.1016/j.psyneuen.2016.11.033

331. Gleason CE, Dowling NM, Wharton W, et al. Effects of Hormone Therapy on Cognition and Mood in Recently Postmenopausal Women: Findings from the Randomized, Controlled KEEPS-Cognitive and Affective Study. *PLoS Med.* Jun 2015;12(6):e1001833; discussion e1001833. doi:10.1371/journal.pmed.1001833
332. Santoro N, Allshouse A, Neal-Perry G, et al. Longitudinal changes in menopausal symptoms comparing women randomized to low-dose oral conjugated estrogens or transdermal estradiol plus micronized progesterone versus placebo: the Kronos Early Estrogen Prevention Study. *Menopause.* 03 2017;24(3):238-246. doi:10.1097/GME.0000000000000756
333. Schmidt PJ, Nieman L, Danaceau MA, et al. Estrogen replacement in perimenopause-related depression: a preliminary report. *Am J Obstet Gynecol.* Aug 2000;183(2):414-20. doi:10.1067/mob.2000.106004
334. Rubinow DR, Girdler SS. Hormones, heart disease, and health: individualized medicine versus throwing the baby out with the bathwater. *Depress Anxiety.* Jun 2011;28(6):E1-E15. doi:10.1002/da.20833
335. Perlis RH, Fraguas R, Fava M, et al. Prevalence and clinical correlates of irritability in major depressive disorder: a preliminary report from the Sequenced Treatment Alternatives to Relieve Depression study. *J Clin Psychiatry.* Feb 2005;66(2):159-66; quiz 147, 273-4.
336. Jha MK, Minhajuddin A, South C, Rush AJ, Trivedi MH. Irritability and Its Clinical Utility in Major Depressive Disorder: Prediction of Individual-Level Acute-Phase Outcomes Using Early Changes in Irritability and Depression Severity. *Am J Psychiatry.* 05 2019;176(5):358-366. doi:10.1176/appi.ajp.2018.18030355
337. Bromberger JT, Matthews KA, Schott LL, et al. Depressive symptoms during the menopausal transition: the Study of Women's Health Across the Nation (SWAN). *J Affect Disord.* Nov 2007;103(1-3):267-72. doi:10.1016/j.jad.2007.01.034
338. Sonfield A, Hasstedt K, Kavanaugh M, Anderson R. *The social and economic benefits of women's ability to determine whether and when to have children.* 2013. https://www.guttmacher.org/sites/default/files/report_pdf/social-economic-benefits.pdf
339. Westhoff CL, Heartwell S, Edwards S, et al. Oral contraceptive discontinuation: do side effects matter? *American Journal of Obstetrics and Gynecology.* 2007;196(4):412. e1-412. e7.
340. Sanders SA, Graham CA, Bass JL, Bancroft J. A prospective study of the effects of oral contraceptives on sexuality and well-being and their relationship to discontinuation. *Contraception.* 2001;64(1):51-58.
341. Rosenberg MJ, Waugh MS. Oral contraceptive discontinuation: a prospective evaluation of frequency and reasons. *American Journal of Obstetrics and Gynecology.* 1998;179(3):577-582.
342. Kristjansdottir J, Olsson GI, Sundelin C, Naessen T. Self-reported health in adolescent girls varies according to the season and its relation to medication and hormonal contraception--a descriptive study. *The European journal of contraception & reproductive health care : the official journal of the European Society of Contraception.* 2013;18(5):343-354. doi:10.3109/13625187.2013.821107 [doi]
343. Toffol E, Heikinheimo O, Koponen P, Luoto R, Partonen T. Further evidence for lack of negative associations between hormonal contraception and mental health. *Contraception.* 2012;86(5):470-480. doi:10.1016/j.contraception.2012.02.014 [doi]

344. Sucato GS, Land SR, Murray PJ, Cecchini R, Gold MA. Adolescents' experiences using the contraceptive patch versus pills. *Journal of pediatric and adolescent gynecology*. 2011;24(4):197-203. doi:10.1016/j.jpog.2011.02.001 [doi]
345. Ott MA, Shew ML, Ofner S, Tu W, Fortenberry JD. The influence of hormonal contraception on mood and sexual interest among adolescents. *Archives of Sexual Behavior*. 2008;37(4):605-613. doi:10.1007/s10508-007-9302-0 [doi]
346. O'Connell K, Davis AR, Kerns J. Oral contraceptives: side effects and depression in adolescent girls. *Contraception*. 2007;75(4):299-304. doi:S0010-7824(06)00363-5 [pii]
347. Coffee AL, Sulak PJ, Kuehl TJ. Long-term assessment of symptomatology and satisfaction of an extended oral contraceptive regimen. *Contraception*. 2007;75(6):444-449. doi:S0010-7824(07)00074-1 [pii]
348. Kutner SJ, Brown WL. Types of oral contraceptives, depression, and premenstrual symptoms. *The Journal of nervous and mental disease*. 1972;155(3):153-162.
349. Gingnell M, Engman J, Frick A, et al. Oral contraceptive use changes brain activity and mood in women with previous negative affect on the pill--a double-blinded, placebo-controlled randomized trial of a levonorgestrel-containing combined oral contraceptive. *Psychoneuroendocrinology*. 2013;38(7):1133-1144. doi:10.1016/j.psyneuen.2012.11.006 [doi]
350. Hall KS, White KO, Rickert VI, Reame N, Westhoff C. Influence of depressed mood and psychological stress symptoms on perceived oral contraceptive side effects and discontinuation in young minority women. *Contraception*. 2012;86(5):518-525. doi:10.1016/j.contraception.2012.04.010 [doi]
351. Wiebe E, Brotto L, MacKay J. Characteristics of women who experience mood and sexual side effects with use of hormonal contraception. *Journal of obstetrics and gynaecology Canada*. 2011;33(12):1234-1240. doi:S1701-2163(16)35108-8 [pii]
352. Kulkarni J. Depression as a side effect of the contraceptive pill. *Expert opinion on drug safety*. 2007;6(4):371-374. doi:10.1517/14740338.6.4.371 [doi]
353. Herzberg BN, Johnson AL, Brown S. Depressive symptoms and oral contraceptives. *British medical journal*. 1970;4(5728):142-145.
354. Cheslack-Postava K, Keyes KM, Lowe SR, Koenen KC. Oral contraceptive use and psychiatric disorders in a nationally representative sample of women. *Archives of women's mental health*. 2015;18(1):103-111. doi:10.1007/s00737-014-0453-4 [doi]
355. Berenson AB, Odom SD, Breitkopf CR, Rahman M. Physiologic and psychologic symptoms associated with use of injectable contraception and 20 microg oral contraceptive pills. *American Journal of Obstetrics and Gynecology*. 2008;199(4):351.e1-351.12. doi:10.1016/j.ajog.2008.04.048 [doi]
356. Rosenthal SL, Cotton S, Ready JN, Potter LS, Succop PA. Adolescents' attitudes and experiences regarding levonorgestrel 100 mcg/ethinyl estradiol 20 mcg. *Journal of pediatric and adolescent gynecology*. 2002;15(5):301-305. doi:S1083318802001961 [pii]
357. Oinonen KA, Mazmanian D. Effects of oral contraceptives on daily self-ratings of positive and negative affect. *Journal of psychosomatic research*. 2001;51(5):647-658. doi:S0022-3999(01)00240-9 [pii]
358. Oddens BJ. Women's satisfaction with birth control: a population survey of physical and psychological effects of oral contraceptives, intrauterine devices, condoms, natural family planning, and sterilization among 1466 women. *Contraception*. 1999;59(5):277-286. doi:S0010-7824(99)00034-7 [pii]

359. Rubino-Watkins MF, Doster JA, Franks S, et al. Oral contraceptive use: implications for cognitive and emotional functioning. *The Journal of nervous and mental disease*. 1999;187(5):275-280.
360. Walker A, Bancroft J. Relationship between premenstrual symptoms and oral contraceptive use: a controlled study. *Psychosomatic medicine*. 1990;52(1):86-96.
361. Joffe H, Cohen LS, Harlow BL. Impact of oral contraceptive pill use on premenstrual mood: predictors of improvement and deterioration. *Am J Obstet Gynecol*. Dec 2003;189(6):1523-30.
362. Zettermark S, Perez Vicente R, Merlo J. Hormonal contraception increases the risk of psychotropic drug use in adolescent girls but not in adults: A pharmacoepidemiological study on 800 000 Swedish women. *PLoS One*. 2018;13(3):e0194773. doi:10.1371/journal.pone.0194773
363. Böttcher B, Radenbach K, Wildt L, Hinney B. Hormonal contraception and depression: a survey of the present state of knowledge. *Arch Gynecol Obstet*. Jul 2012;286(1):231-6. doi:10.1007/s00404-012-2298-2
364. Ormel J, Oldehinkel AJ, Sijtsma J, et al. The TRacking Adolescents' Individual Lives Survey (TRAILS): design, current status, and selected findings. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2012;51(10):1020-1036.
365. Oldehinkel AJ, Rosmalen JGM, Buitelaar JK, et al. Cohort profile update: the tracking adolescents' individual lives survey (TRAILS). *International journal of epidemiology*. 2014;44(1):76-76n.
366. Achenbach T, Rescorla L. Manual for the ASEBA School-Age Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families; 2001.
367. Achenbach T, Rescorla L. Manual for the ASEBA Adult Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families; 2003.
368. Achenbach TM, Dumenci L, Rescorla LA. Ratings of relations between DSM-IV diagnostic categories and items of the Adult Self-Report (ASR) and Adult Behavior Checklist (ABCL). *Research Center for Children, Youth&FamiliesAdres: wwwASEBAorg*. 2003;
369. Cahill L. How does hormonal contraception affect the developing human adolescent brain? *Current Opinion in Behavioral Sciences*. 2019;23doi:10.1016/j.cobeha.2018.06.015
370. Eisenhofer G, Peitzsch M, Kaden D, et al. Reference intervals for plasma concentrations of adrenal steroids measured by LC-MS/MS: Impact of gender, age, oral contraceptives, body mass index and blood pressure status. *Clin Chim Acta*. Jul 2017;470:115-124. doi:10.1016/j.cca.2017.05.002
371. Casper R, Yonkers K. Treatment of premenstrual syndrome and premenstrual dysphoric disorder. UpToDate; 2018.
372. Martin KA, Anderson RR, Chang RJ, et al. Evaluation and Treatment of Hirsutism in Premenopausal Women: An Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab*. 04 2018;103(4):1233-1257. doi:10.1210/jc.2018-00241
373. Rice F, Riglin L, Lomax T, et al. Adolescent and adult differences in major depression symptom profiles. *J Affect Disord*. 01 2019;243:175-181. doi:10.1016/j.jad.2018.09.015
374. Jorg F, Visser E, Ormel J, Reijneveld SA, Hartman CA, Oldehinkel AJ. Mental health care use in adolescents with and without mental disorders. *Eur Child Adolesc Psychiatry*. May 2016;25(5):501-8. doi:10.1007/s00787-015-0754-9

375. Visser ST, Schuiling-Veninga CC, Bos JH, de Jong-van den Berg LT, Postma MJ. The population-based prescription database IADB.nl: its development, usefulness in outcomes research and challenges. *Expert Rev Pharmacoecon Outcomes Res.* Jun 2013;13(3):285-92. doi:10.1586/erp.13.20
376. Higgins JA, Smith NK. The Sexual Acceptability of Contraception: Reviewing the Literature and Building a New Concept. *J Sex Res.* 2016 May-Jun 2016;53(4-5):417-56. doi:10.1080/00224499.2015.1134425
377. Anticonceptiemiddelen (Zvw). National Health Care Institute; 2019.
378. Insurance Coverage of Contraceptives. Guttmacher Institute; 2019.
379. Rogers J, Hengartner MP, Angst J, Ajdacic-Gross V, Rössler W. Associations with quality of life and the effect of psychopathology in a community study. *Soc Psychiatry Psychiatr Epidemiol.* Sep 2014;49(9):1467-73. doi:10.1007/s00127-014-0841-0
380. Hall KS, Moreau C, Trussell J, Barber J. Role of young women's depression and stress symptoms in their weekly use and nonuse of contraceptive methods. *J Adolesc Health.* Aug 2013;53(2):241-8. doi:10.1016/j.jadohealth.2013.02.009
381. ACOG Practice Bulletin No. 110: noncontraceptive uses of hormonal contraceptives. *Obstet Gynecol.* Jan 2010;115(1):206-18. doi:10.1097/AOG.0b013e3181cb50b5
382. Adolescence Co. Contraception for adolescents. *Pediatrics.* Oct 2014;134(4):e1244-56. doi:10.1542/peds.2014-2299
383. Kavanaugh ML, Jerman J. Contraceptive method use in the United States: trends and characteristics between 2008, 2012 and 2014. *Contraception.* Jan 2018;97(1):14-21. doi:10.1016/j.contraception.2017.10.003
384. de Graaf H, van den Borne M, Nikkelen S, Twisk D, Meijer S. *Seksuele gezondheid van jongeren in Nederland anno 2017.* 2017. <https://seksonderje25e.nl/files/uploads/Onderzoeksboek%20Seks%20onder%20je%2025e%202017.pdf>
385. Sonfield A, Hasstedt K, Kavanaugh M, Anderson R. *The Social and Economic Benefits of Women's Ability To Determine Whether and When to Have Children.* 2013. www.guttmacher.org/pubs/social-economic-benefits.pdf<[www.guttmacher.org/pubs=""](http://www.guttmacher.org/pubs=) social-economic-benefits.pdf=""></[www.guttmacher.org/pubs=""](http://www.guttmacher.org/pubs=)>
386. United Nations DoEaSA, Population Division. Contraceptive Use by Method 2019. *Data Booklet (ST/ESA/SERA/435).* 2019. https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2020/Jan/un_2019_contraceptiveusebymethod_databooklet.pdf
387. de Wit AE, Booij SH, Giltay EJ, Joffe H, Schoevers RA, Oldehinkel AJ. Association of Use of Oral Contraceptives With Depressive Symptoms Among Adolescents and Young Women. *JAMA Psychiatry.* Oct 2019;doi:10.1001/jamapsychiatry.2019.2838
388. Simonsohn U, Simmons JP, Nelson LD. Specification curve analysis. *Nat Hum Behav.* Jul 2020;doi:10.1038/s41562-020-0912-z
389. Orben A, Przybylski AK. The association between adolescent well-being and digital technology use. *Nat Hum Behav.* 02 2019;3(2):173-182. doi:10.1038/s41562-018-0506-1
390. Botvinik-Nezer R, Holzmeister F, Camerer CF, et al. Variability in the analysis of a single neuroimaging dataset by many teams. *Nature.* 06 2020;582(7810):84-88. doi:10.1038/s41586-020-2314-9

391. Bot M, Middeldorp CM, de Geus EJ, et al. Validity of LIDAS (Lifetime Depression Assessment Self-report): a self-report online assessment of lifetime major depressive disorder. *Psychol Med*. 01 2017;47(2):279-289. doi:10.1017/S0033291716002312
392. Vrijen C, Hartman CA, Oldehinkel AJ. Reward-Related Attentional Bias at Age 16 Predicts Onset of Depression During 9 Years of Follow-up. *J Am Acad Child Adolesc Psychiatry*. 03 2019;58(3):329-338. doi:10.1016/j.jaac.2018.06.009
393. Kessler RC, Abelson J, Demler O, et al. Clinical calibration of DSM-IV diagnoses in the World Mental Health (WMH) version of the World Health Organization (WHO) Composite International Diagnostic Interview (WMHCIDI). *Int J Methods Psychiatr Res*. 2004;13(2):122-39. doi:10.1002/mpr.169
394. Haro JM, Arbabzadeh-Bouchez S, Brugha TS, et al. Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health surveys. *Int J Methods Psychiatr Res*. 2006;15(4):167-80. doi:10.1002/mpr.196
395. Burcusa SL, Iacono WG. Risk for recurrence in depression. *Clin Psychol Rev*. Dec 2007;27(8):959-85. doi:10.1016/j.cpr.2007.02.005
396. Lundin C, Danielsson KG, Bixo M, et al. Combined oral contraceptive use is associated with both improvement and worsening of mood in the different phases of the treatment cycle-A double-blind, placebo-controlled randomized trial. *Psychoneuroendocrinology*. 02 2017;76:135-143. doi:10.1016/j.psyneuen.2016.11.033
397. Vigil P, Del Río JP, Carrera B, ArÁnguiz FC, Rioseco H, Cortés ME. Influence of sex steroid hormones on the adolescent brain and behavior: An update. *Linacre Q*. Aug 2016;83(3):308-329. doi:10.1080/00243639.2016.1211863
398. Bouma EM, Riese H, Ormel J, Verhulst FC, Oldehinkel AJ. Adolescents' cortisol responses to awakening and social stress; effects of gender, menstrual phase and oral contraceptives. The TRAILS study. *Psychoneuroendocrinology*. Jul 2009;34(6):884-93. doi:10.1016/j.psyneuen.2009.01.003
399. Sharma R, Smith SA, Boukina N, et al. Use of the birth control pill affects stress reactivity and brain structure and function. *Horm Behav*. 08 2020;124:104783. doi:10.1016/j.yhbeh.2020.104783
400. LeMoult J. From Stress to Depression: Bringing Together Cognitive and Biological Science. *Curr Dir Psychol Sci*. Dec 2020;29(6):592-598. doi:10.1177/0963721420964039
401. Committee opinion no. 615: Access to contraception. *Obstet Gynecol*. Jan 2015;125(1):250-255. doi:10.1097/01.AOG.0000459866.14114.33
402. Worly BL, Gur TL, Schaffir J. The relationship between progestin hormonal contraception and depression: a systematic review. *Contraception*. 06 2018;97(6):478-489. doi:10.1016/j.contraception.2018.01.010
403. Greco T, Graham CA, Bancroft J, Tanner A, Doll HA. The effects of oral contraceptives on androgen levels and their relevance to premenstrual mood and sexual interest: a comparison of two triphasic formulations containing norgestimate and either 35 or 25 microg of ethinyl estradiol. *Contraception*. Jul 2007;76(1):8-17. doi:10.1016/j.contraception.2007.04.002

404. Kelly S, Davies E, Fearn S, et al. Effects of oral contraceptives containing ethinylestradiol with either drospirenone or levonorgestrel on various parameters associated with well-being in healthy women: a randomized, single-blind, parallel-group, multicentre study. *Clin Drug Investig*. 2010;30(5):325-36. doi:10.2165/11535450-000000000-00000
405. Sangthawan M, Taneepanichskul S. A comparative study of monophasic oral contraceptives containing either drospirenone 3 mg or levonorgestrel 150 microg on premenstrual symptoms. *Contraception*. Jan 2005;71(1):1-7. doi:10.1016/j.contraception.2004.07.010
406. Coffee AL, Kuehl TJ, Willis S, Sulak PJ. Oral contraceptives and premenstrual symptoms: comparison of a 21/7 and extended regimen. *Am J Obstet Gynecol*. Nov 2006;195(5):1311-9. doi:10.1016/j.ajog.2006.05.012
407. Higgins JP, Welton NJ. Network meta-analysis: a norm for comparative effectiveness? *Lancet*. Aug 2015;386(9994):628-30. doi:10.1016/S0140-6736(15)61478-7
408. Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychol Methods*. Mar 2002;7(1):105-25. doi:10.1037/1082-989x.7.1.105
409. Brignardello-Petersen R, Bonner A, Alexander PE, et al. Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis. *J Clin Epidemiol*. 01 2018;93:36-44. doi:10.1016/j.jclinepi.2017.10.005
410. Nikolakopoulou A, Higgins JPT, Papakonstantinou T, et al. CINeMA: An approach for assessing confidence in the results of a network meta-analysis. *PLoS Med*. 04 2020;17(4):e1003082. doi:10.1371/journal.pmed.1003082
411. Papakonstantinou T, Nikolakopoulou A, Higgins JPT, Egger M, Salanti G. CINeMA: Software for semiautomated assessment of the confidence in the results of network meta-analysis. *Campbell systematic reviews*. 2020;
412. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 08 2019;366:l4898. doi:10.1136/bmj.l4898
413. Guang-Sheng F, Mei-Lu B, Li-Nan C, et al. Efficacy and safety of the combined oral contraceptive ethinylestradiol/drospirenone (Yasmin) in healthy Chinese women: a randomized, open-label, controlled, multicentre trial. *Clin Drug Investig*. 2010;30(6):387-96. doi:10.1007/BF03256908
414. Elaut E, Buysse A, De Sutter P, et al. Relation of androgen receptor sensitivity and mood to sexual desire in hormonal contraception users. *Contraception*. May 2012;85(5):470-9. doi:10.1016/j.contraception.2011.10.007
415. Graham CA, Ramos R, Bancroft J, Maglaya C, Farley TM. The effects of steroidal contraceptives on the well-being and sexuality of women: a double-blind, placebo-controlled, two-centre study of combined and progestogen-only methods. *Contraception*. Dec 1995;52(6):363-9. doi:10.1016/0010-7824(95)00226-x
416. Witjes H, Creinin MD, Sundström-Poromaa I, Martin Nguyen A, Korver T. Comparative analysis of the effects of norgestrel acetate/17 β -estradiol and drospirenone/ethinylestradiol on premenstrual and menstrual symptoms and dysmenorrhea. *Eur J Contracept Reprod Health Care*. 2015;20(4):296-307. doi:10.3109/13625187.2015.1016154

417. Winkler UH, Ferguson H, Mulders JA. Cycle control, quality of life and acne with two low-dose oral contraceptives containing 20 microg ethinylestradiol. *Contraception*. Jun 2004;69(6):469-76. doi:10.1016/j.contraception.2003.12.017
418. Engman J, Sundström Poromaa I, Moby L, Wikström J, Fredrikson M, Gingnell M. Hormonal Cycle and Contraceptive Effects on Amygdala and Salience Resting-State Networks in Women with Previous Affective Side Effects on the Pill. *Neuropsychopharmacology*. 02 2018;43(3):555-563. doi:10.1038/npp.2017.157
419. Zethraeus N, Dreber A, Ranehill E, et al. A first-choice combined oral contraceptive influences general well-being in healthy women: a double-blind, randomized, placebo-controlled trial. *Fertil Steril*. 05 2017;107(5):1238-1245. doi:10.1016/j.fertnstert.2017.02.120
420. Davis SR, Bitzer J, Giraldi A, et al. Change to either a nonandrogenic or androgenic progestin-containing oral contraceptive preparation is associated with improved sexual function in women with oral contraceptive-associated sexual dysfunction. *J Sex Med*. Dec 2013;10(12):3069-79. doi:10.1111/jsm.12310
421. Battaglia C, Morotti E, Persico N, et al. Clitoral vascularization and sexual behavior in young patients treated with drospirenone-ethinyl estradiol or contraceptive vaginal ring: a prospective, randomized, pilot study. *J Sex Med*. Feb 2014;11(2):471-80. doi:10.1111/jsm.12392
422. Shahnazi M, Farshbaf Khalili A, Ranjbar Kochaksaraei F, et al. A comparison of second and third generations combined oral contraceptive pills' effect on mood. *Iran Red Crescent Med J*. Aug 2014;16(8):e13628. doi:10.5812/ircmj.13628
423. Legro RS, Pauli JG, Kunselman AR, et al. Effects of continuous versus cyclical oral contraception: a randomized controlled trial. *J Clin Endocrinol Metab*. Feb 2008;93(2):420-9. doi:10.1210/jc.2007-2287
424. Elovainio M, Teperi J, Aalto AM, et al. Depressive symptoms as predictors of discontinuation of treatment of menorrhagia by levonorgestrel-releasing intrauterine system. *Int J Behav Med*. 2007;14(2):70-5. doi:10.1007/BF03004171
425. Rowlands S. Newer progestogens. *J Fam Plann Reprod Health Care*. Jan 2003;29(1):13-6. doi:10.1783/147118903101197188
426. Trussell J, Aiken A, Micks E, Guthrie K. Efficacy, safety, and personal considerations. *Contraceptive technology*. 21 ed. Ayer Company Publishers, Inc.; 2018.
427. Association AP. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). Arlington: American Psychiatric Association; 2013.
428. Yonkers KA, O'Brien PM, Eriksson E. Premenstrual syndrome. *Lancet*. Apr 2008;371(9619):1200-10. doi:10.1016/S0140-6736(08)60527-9
429. Wittchen H-U, Becker E, Lieb R, Krause P. Prevalence, incidence and stability of premenstrual dysphoric disorder in the community. *Psychol Med*. Jan 2002;32(1):119-32. doi:10.1017/s0033291701004925
430. Potter J, Bouyer J, Trussell J, Moreau C. Premenstrual syndrome prevalence and fluctuation over time: results from a French population-based survey. *J Womens Health (Larchmt)*. 2009 Jan-Feb 2009;18(1):31-9. doi:10.1089/jwh.2008.0932
431. Rapkin AJ, Korotkaya Y, Taylor KC. Contraception counseling for women with premenstrual dysphoric disorder (PMDD): current perspectives. *Open Access J Contracept*. 2019;10:27-39. doi:10.2147/OAJC.S183193

432. BJOG. Management of premenstrual syndrome. 2016.
433. Hofmeister S, Bodden S. Premenstrual Syndrome and Premenstrual Dysphoric Disorder. *Am Fam Physician*. Aug 2016;94(3):236-40.
434. Administration UFaD. Yaz (drospirenone/ethinyl estradiol) for Oral Contraception and Premenstrual Dysphoric Disorder. FDA. Accessed January 7, 2021. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2006/021873_yaz_toc.cfm
435. Yonkers KA, Brown C, Pearlstein TB, Foegh M, Sampson-Landers C, Rapkin A. Efficacy of a new low-dose oral contraceptive with drospirenone in premenstrual dysphoric disorder. *Obstet Gynecol*. Sep 2005;106(3):492-501. doi:10.1097/01.AOG.0000175834.77215.2e
436. Pearlstein TB, Bachmann GA, Zacur HA, Yonkers KA. Treatment of premenstrual dysphoric disorder with a new drospirenone-containing oral contraceptive formulation. *Contraception*. Dec 2005;72(6):414-21. doi:10.1016/j.contraception.2005.08.021
437. Eisenlohr-Moul TA, Girdler SS, Johnson JL, Schmidt PJ, Rubinow DR. Treatment of premenstrual dysphoria with continuous versus intermittent dosing of oral contraceptives: Results of a three-arm randomized controlled trial. *Depress Anxiety*. 10 2017;34(10):908-917. doi:10.1002/da.22673
438. Freeman EW, Kroll R, Rapkin A, et al. Evaluation of a unique oral contraceptive in the treatment of premenstrual dysphoric disorder. *J Womens Health Gend Based Med*. 2001 Jul-Aug 2001;10(6):561-9. doi:10.1089/15246090152543148
439. Freeman EW, Halbreich U, Grubb GS, et al. An overview of four studies of a continuous oral contraceptive (levonorgestrel 90 mcg/ethinyl estradiol 20 mcg) on premenstrual dysphoric disorder and premenstrual syndrome. *Contraception*. May 2012;85(5):437-45. doi:10.1016/j.contraception.2011.09.010
440. Halbreich U, Freeman EW, Rapkin AJ, et al. Continuous oral levonorgestrel/ethinyl estradiol for treating premenstrual dysphoric disorder. *Contraception*. Jan 2012;85(1):19-27. doi:10.1016/j.contraception.2011.05.008
441. Higgins J, Thomas J, Chandler J, et al. *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0 (updated July 2019). 6 ed: Cochrane; 2019.
442. Association AP. *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.). 2000.
443. Graham CA, Sherwin BB. The relationship between mood and sexuality in women using an oral contraceptive as a treatment for premenstrual symptoms. *Psychoneuroendocrinology*. 1993;18(4):273-81. doi:10.1016/0306-4530(93)90024-f
444. Wichianpitaya J, Taneepanichskul S. A comparative efficacy of low-dose combined oral contraceptives containing desogestrel and drospirenone in premenstrual symptoms. *Obstet Gynecol Int*. 2013;2013:487143. doi:10.1155/2013/487143
445. Shehata NAA, Moety GAFA, El Wahed HAA, Fahim AS, Katta MA, Hussein GK. Does Adding Fluoxetine to Combined Oral Contraceptives Containing Drospirenone Improve the Management of Severe Premenstrual Syndrome? A 6-Month Randomized Double-Blind Placebo-Controlled Three-Arm Trial. *Reprod Sci*. 02 2020;27(2):743-750. doi:10.1007/s43032-019-00080-x
446. Yonkers KA, Simoni MK. Premenstrual disorders. *Am J Obstet Gynecol*. 01 2018;218(1):68-74. doi:10.1016/j.ajog.2017.05.045

447. Lopez L, Kaptein A, Helmerhorst F. Oral contraceptives containing drospirenone for premenstrual syndrome. 2012
448. Marjoibanks J, Brown J, O'Brien P, Wyatt K. Selective serotonin reuptake inhibitors for premenstrual syndrome. *Cochrane Database of Systematic Reviews*. 2013;doi:<https://doi.org/10.1002/14651858.CD001396.pub3>
449. Busse JW, Montori VM, Krasnik C, Patelis-Siotis I, Guyatt GH. Psychological intervention for premenstrual syndrome: a meta-analysis of randomized controlled trials. *Psychother Psychosom*. 2009;78(1):6-15. doi:10.1159/000162296
450. FDA. FDA Drug Safety Communication: Safety review update on the possible increased risk of blood clots with birth control pills containing drospirenone FDA. Accessed January 5, 2021. <https://www.fda.gov/drugs/drug-safety-and-availability/fda-drug-safety-communication-safety-review-update-possible-increased-risk-blood-clots-birth-control>
451. Nieuws R. Meisjes van 16 die aan de pil zijn hebben vaker depressieve gevoelens. RTL Nieuws. Accessed August 28, 2020. <https://www.rtlnieuws.nl/nieuws/artikel/4869961/anticonceptiepil-depressie-meisjes-bijwerkingen>
452. Broersen S. Adolescenten aan de pil hebben meer depressieve symptomen. Medisch Contact. Accessed August 28, 2020. <https://www.medischcontact.nl/nieuws/laatste-nieuws/nieuwsartikel/adolescenten-aan-de-pil-hebben-meer-depressieve-symptomen.htm>
453. Huhn M. Birth control might be making teens depressed. New York Post. Accessed August 28, 2020. <https://nypost.com/2019/10/02/birth-control-might-be-making-teens-depressed/>
454. Radar. Meisjes aan de anticonceptiepil vaker depressief. Avro Tros. Accessed August 28, 2020. <https://radar.avrotros.nl/nieuws/item/meisjes-aan-de-anticonceptiepil-vaker-depressief/>
455. van Akkeren G. Meiden van 16 aan de pil hebben vaker depressieve gevoelens. OogTV; 2019.
456. de Wit A. Meisjes van 16 die aan de pil zijn hebben vaker depressieve gevoelens. Accessed August 28, 2020. <https://doktermedia.nl/reviews/alle-berichten/2019/10/07/meisjes-van-16-die-aan-de-pil-zijn-hebben-vaker-depressieve-gevoelens/>
457. Leeman F. Maakt de pil depressief? <https://www.bnr.nl/podcast/factgurus/10393426/maakt-de-pil-depressief>
458. Kancheva R, Hill M, Novák Z, Chrastina J, Kancheva L, Stárka L. Neuroactive steroids in periphery and cerebrospinal fluid. *Neuroscience*. Sep 2011;191:22-7. doi:10.1016/j.neuroscience.2011.05.054
459. Gordon JL, Rubinow DR, Eisenlohr-Moul TA, Xia K, Schmidt PJ, Girdler SS. Efficacy of Transdermal Estradiol and Micronized Progesterone in the Prevention of Depressive Symptoms in the Menopause Transition: A Randomized Clinical Trial. *JAMA Psychiatry*. Feb 2018;75(2):149-157. doi:10.1001/jamapsychiatry.2017.3998
460. Soares CN, Almeida OP, Joffe H, Cohen LS. Efficacy of estradiol for the treatment of depressive disorders in perimenopausal women: a double-blind, randomized, placebo-controlled trial. *Arch Gen Psychiatry*. Jun 2001;58(6):529-34. doi:10.1001/archpsyc.58.6.529

461. Maney DL. Perils and pitfalls of reporting sex differences. *Philos Trans R Soc Lond B Biol Sci.* Feb 2016;371(1688):20150119. doi:10.1098/rstb.2015.0119
462. Durdiakova J, Ostatnikova D, Celec P. Testosterone and its metabolites--modulators of brain functions. *Acta Neurobiol Exp (Wars).* 2011;71(4):434-54.
463. Hendricks V, Vestergaard M. Reality Lost. Springer, Cham; 2019:chap The Attention Economy.
464. FUNK C, GOTTFRIED J, MITCHELL A. Science News and Information Today. Pew Research Center. Accessed February 15, 2021. <https://www.journalism.org/2017/09/20/science-news-and-information-today/>
465. Jameson JL, Kasper DL, Longo DL, Fauci AS, Hauser SL, Loscalzo J. *Harrison's principles of internal medicine*. 20th edition. ed. McGraw-Hill Education; 2018.
466. Lood Y, Aardal E, Ahlner J, et al. Determination of testosterone in serum and saliva by liquid chromatography-tandem mass spectrometry: An accurate and sensitive method applied on clinical and forensic samples. *J Pharm Biomed Anal.* Dec 2020;113823. doi:10.1016/j.jpba.2020.113823
467. Walther A, Tsao C, Pande R, Kirschbaum C, Field E, Berkman L. Do dehydroepiandrosterone, progesterone, and testosterone influence women's depression and anxiety levels? Evidence from hair-based hormonal measures of 2105 rural Indian women. *Psychoneuroendocrinology.* Jul 2019;109:104382. doi:10.1016/j.psyneuen.2019.104382
468. Sakkas D, Howles CM, Atkinson L, et al. A multi-centre international study of salivary hormone oestradiol and progesterone measurements in ART monitoring. *Reprod Biomed Online.* Oct 2020;doi:10.1016/j.rbmo.2020.10.012
469. Santoro N, El Khoudary SR, Nasr A, et al. Daily luteal serum and urinary hormone profiles in the menopause transition: Study of Women's Health Across the Nation. *Menopause.* 02 2020;27(2):127-133. doi:10.1097/GME.0000000000001453
470. Fudulu DP, Angelini GD, Papadopolou FF, et al. The Peacock study: feasibility of the dynamic characterisation of the paediatric hypothalamic-pituitary-adrenal function during and after cardiac surgery. *BMC Cardiovasc Disord.* 05 2020;20(1):245. doi:10.1186/s12872-020-01516-y
471. Bhake R, Russell GM, Kershaw Y, et al. Continuous Free Cortisol Profiles in Healthy Men. *J Clin Endocrinol Metab.* 04 2020;105(4)doi:10.1210/clinem/dgz002
472. Callahan RL, Mehta NJ, Nanda K, Kopf GS. The new contraceptive revolution: developing innovative products outside of industry†,‡. *Biol Reprod.* 08 2020;103(2):157-166. doi:10.1093/biolre/ioaa067
473. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ.* May 2003;326(7400):1167-70. doi:10.1136/bmj.326.7400.1167
474. Lundh A, Lexchin J, Mintzes B, Schroll JB, Bero L. Industry sponsorship and research outcome. *Cochrane Database Syst Rev.* 02 2017;2:MR000033. doi:10.1002/14651858.MR000033.pub3
475. Nieuwsgierig en betrokken, de waarde van wetenschap (2018).
476. Shi Z, Araujo AB, Martin S, O'Loughlin P, Wittert GA. Longitudinal changes in testosterone over five years in community-dwelling men. *J Clin Endocrinol Metab.* Aug 2013;98(8):3289-97. doi:10.1210/jc.2012-3842

477. Boschloo L, Vogelzangs N, Smit JH, et al. Comorbidity and risk indicators for alcohol use disorders among persons with anxiety and/or depressive disorders: findings from the Netherlands Study of Depression and Anxiety (NESDA). *J Affect Disord.* Jun 2011;131(1-3):233-42. doi:10.1016/j.jad.2010.12.014
478. Rachdaoui N, Sarkar DK. Effects of alcohol on the endocrine system. *Endocrinol Metab Clin North Am.* Sep 2013;42(3):593-615. doi:10.1016/j.ecl.2013.05.008
479. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption--II. *Addiction.* Jun 1993;88(6):791-804. doi:10.1111/j.1360-0443.1993.tb02093.x
480. Richtlijn goede voeding 2015 (Gezondheidsraad) (2015).
481. de Wit AE, Giltay EJ, de Boer MK, et al. Associations between testosterone and metabolic syndrome in depressed and non-depressed older men and women. *Int J Geriatr Psychiatry.* 03 2019;34(3):463-471. doi:10.1002/gps.5040
482. Simon GE, Ludman EJ, Linde JA, et al. Association between obesity and depression in middle-aged women. *Gen Hosp Psychiatry.* 2008 Jan-Feb 2008;30(1):32-9. doi:10.1016/j.genhosppsych.2007.09.001
483. Jokela M, García-Velázquez R, Airaksinen J, Gluschkoff K, Kivimäki M, Rosenström T. Chronic diseases and social risk factors in relation to specific symptoms of depression: Evidence from the U.S. national health and nutrition examination surveys. *J Affect Disord.* 05 2019;251:242-247. doi:10.1016/j.jad.2019.03.074
484. Casson PR, Elkind-Hirsch KE, Buster JE, Hornsby PJ, Carson SA, Snabes MC. Effect of postmenopausal estrogen replacement on circulating androgens. *Obstet Gynecol.* Dec 1997;90(6):995-8. doi:10.1016/s0029-7844(97)00538-3
485. Epstein NB, Baldwin LM, Bishop DS. The McMaster family assessment device. *Journal of Marital and Family Therapy*; 1983. p. 171-180.

Supplements

Chapter 3

Supplement 1. Characteristics of 1432 participants with follow-up and 227 participants without follow-up at baseline.

	Sample with follow-up n = 1432	Sample without follow-up n = 227	p ^a
Age , mean (SD)	41.3 ± 13.0	39.6 ± 12.8	.07
Education , mean (SD)	12.4 ± 3.2	11.3 ± 3.1	<.0001
Body mass index , mean (SD)	25.2 ± 5.4	25.6 ± 5.1	.31
Current smoking , no. (%)	504 (35.1)	117 (51.5)	<.0001
>1 alcohol unit/day , no. (%)	185 (12.9)	28 (12.3)	.91
Treated chronic diseases , geo. mean (95%CI)	0.5 (0.5-0.5)	0.5 (0.4-0.6)	0.49
Oral contraceptive use , no. (%)	390 (27.2)	62 (27.3)	1.00
Hormonal replacement therapy use , no. (%)	25 (1.7)	3 (1.3)	1.00
Menopausal , no. (%)	449 (34.2)	57 (27.5)	.07
Antidepressant use , no. (%)	369 (25.8)	73 (32.2)	.05
Current MDD , no. (%)	405 (28.3)	108 (47.6)	<.0001
Current anxiety disorder , no. (%)	499 (34.8)	124 (54.6)	<.0001
<i>Predictors</i>			
Total testosterone (nmol/L), geo. mean (95%CI)	0.73 (0.71-0.74)	0.73 (0.69-0.77)	.84
SHBG (nmol/L), geo. mean (95%CI)	64.2 (62.2-66.2)	61.4 (56.8-66.4)	.29
Free testosterone (pmol/L), geo. mean (95%CI)	8.08 (7.84-8.33)	8.35 (7.77-8.96)	.44
5α-DHT (nmol/L), geo. mean (95%CI)	0.32 (0.31-0.33)	0.34 (0.31-0.36)	.16
Androstenedione (nmol/L), geo. mean (95%CI)	3.07 (2.99-3.16)	3.27 (3.06-3.49)	.10
DHEAS (μmol/L), geo. mean (95%CI)	4.40 (4.25-4.55)	4.53 (4.22-4.86)	.49

Abbreviations: 5α-DHT, 5α-dihydrotestosterone; BMI, body mass index; CI, confidence interval; DHEAS, dehydroepiandrosterone-sulphate; MDD, major depressive disorder; SD, standard deviation; SHBG, sex hormone binding globulin. To convert nmol/L to ng/dl for total testosterone, 5α-DHT, and androstenedione multiply by 28.84, 29.07, and 28.64, respectively. To convert pmol/L to pg/ml for free testosterone divide by 3.47. To convert μmol/L to μg/dl for DHEAS, multiply by 36.85. To convert to nmol/L to mg/ml for SHBG, multiply by 0.095.^a Based on one-way analysis of variance for independent samples or χ² tests.

Supplement 2. Motivation for including covariates and their definitions.

Below, we describe the motivation and the definition of the covariates that were used in the analyses.

- *Sociodemographic*

Motivation: Age and education were used as covariates to control for basic sociodemographic differences.

Definition: Both age and education were defined in years.

- *Current smoking status*

Motivation: Current smoking status was added to the models as smoking affects testosterone levels,⁴⁷⁶ and smokers are more likely to be depressed (see *Table 3.1*).

Definition: Smoking status was ascertained with a question whether participants currently smoked or not.

- *Alcohol use ($\leq$ or >7 units a week)*

Motivation: Alcohol use was considered as a covariate as it may increase the risk for depression,⁴⁷⁷ and chronic use may cause hormonal disturbances.⁴⁷⁸

Definition: Alcohol use was ascertained with two questions from the Alcohol Use Disorders Identification Test (AUDIT).⁴⁷⁹ By combining the answers from the question on how often a participant drinks, and the question that assessed the typical quantity when a participant drinks, the number of alcohol units per week were calculated. We dichotomized the number of drinks in seven or less and more than seven as use from the first mentioned group would be considered responsible alcohol use following Dutch guidelines.⁴⁸⁰

- *Body mass index (BMI)*

Motivation: BMI was considered as a covariate as higher testosterone levels are associated with higher measures of BMI,⁴⁸¹ and higher measures of BMI are associated with more depressive symptoms.⁴⁸²

Definition: BMI was defined as weight in kg divided by square of the body height.

- *Number of treated chronic diseases*

Motivation: In men, testosterone levels are affected by health status.²⁰³ Although that is unknown in women, we assumed that testosterone levels might be affected by other somatic diseases as well. And as the number of chronic diseases is associated with an increased risk of depressive symptoms,⁴⁸³ this variable was considered a covariate.

Definition: The number of chronic diseases were determined using a 21-item face to face interview instrument that is designed for the Netherlands Study on Depression and Anxiety (NESDA). The respondents were asked if they have a certain specified chronic disease. If a disease is present, respondents were further asked if they used medication and if they were under treatment by a physician for their disease. The following diseases are asked; asthma, chronic bronchitis or pulmonary emphysema; heart diseases or infarct; diabetes; cerebrovascular infarction or hemorrhage; arthritis or arthrosis; rheumatic complaints; benign or malignant tumor & metastasis; high blood pressure; stomach or intestinal ulcer; intestinal disorders; liver disease or liver cirrhosis; epilepsy; chronic fatigue syndrome; allergies; thyroid gland disease; injury (last year); head injury (ever); another chronic disease. In the last item of the questionnaire (item 21) respondents were asked about any other chronic disease that had not previously been mentioned. Respondents could mention up to five additional chronic diseases. For neurologic disorders, a separate variable was constructed including migraine & cluster headache, multiple sclerosis, neuropathy and other neurologic diseases. A count of number of treated chronic diseases was made. All diseases were included for which medications were used or respondents were treated, except allergies, injuries, head injuries, chronic fatigue syndrome, high blood pressure, and angina. Osteoarthritis and rheumatism were combined and counted as one, since people often confuse these diseases. Regardless of how many other diseases were mentioned in the last question (question 21 - Another chronic disease), this item was counted as one disease. For intestinal disorders some categories were excluded from the count as they were considered to be symptoms instead of pathologies. These categories were diarrhea, obstipation, stress-induced complaints, cramps and irregular defecation. The number of participants (%) of five prevalent groups of disorders were the following:

- Lifetime prevalence heart condition or heart attack n = 64 (3.9)
- Diabetes n = 46 (2.8)
- Lifetime prevalence stroke n = 23 (1.4)
- Rheumatoid arthritis n = 21 (1.2)
- Lifetime prevalence cancer n = 136 (8.2)

- *Current oral contraceptive pill (OC) use*

Motivation: As OC use lowers androgen levels by increasing sex hormone binding globulin (SHBG) levels (see *Supplement 3*) and is associated with an increased risk for depressive symptoms and depression,^{124,387} OC use was included as a covariate.

Definition: Current OC use was ascertained with the question, “Do you take a contraceptive pill?”.

- *Hormonal replacement therapy use*

Motivation: As for OC use, hormonal replacement therapy use lowers androgen levels and hence should be considered as a covariate.⁴⁸⁴

Definition: Current hormonal replacement use was ascertained when women used medications with World Health Organization Anatomical Therapeutic Chemical (ATC) G03CA or G03F.

- *Menopausal status*

Motivation: As androgen production by the ovaries ceases after menopause, being post-menopausal affects your androgen levels.¹⁹⁸

Also, the risk for depression is most prominent during the reproductive years. Hence, menopausal status was included as a covariate.¹⁴

Definition: Postmenopausal status was determined with the question, “Do you (still) have a menstrual cycle?”. Women who answered the question with “Yes” (n = 1082; 65.2%) were defined as being premenopausal, women who answered with “No, menopause” (n = 462; 27.8%) were considered postmenopausal, and women who answered with “No, hysterectomy” were categorized as premenopausal when they were aged ≤ 50 (n = 36; 2.2%), or as postmenopausal when they were aged ≥ 51 (n = 58; 3.5%). The latter classification was done, as it was unknown whether the hysterectomy had also included a bilateral oophorectomy.

Table. F-statistics with p-value for each of the covariates with each of the androgen and SHBG levels in the fully adjusted analysis of covariance analyses.

<i>Covariate</i>	Total testosterone		SHBG		Free testosterone		5 α -DHT		Androstenedione		DHEAS	
	F	p	F	p	F	p	F	p	F	p	F	p
Age	52.7	<.001	78.3	<.001	44.4	<.001	146.0	<.001	154.4	<.001	5.8	.02
Education	0.4	.54	.2	.63	1.2	.27	0.3	.59	1.5	.22	.1	.76
BMI	0.7	.40	81.5	<.001	24.8	<.001	8.6	.003	0.2	.67	220.2	<.001
Smoking status	27.9	<.001	9.8	.002	33.2	<.001	54.3	<.001	18.9	<.001	7.9	.005
Alcohol use	0.1	.71	1.8	.18	0.0	.91	1.6	.21	17.7	<.001	6.0	.02
Number of treated chronic diseases	28.3	<.001	10.0	.002	26.6	<.001	14.4	<.001	63.0	<.001	7.3	.007
Postmenopausal status	0.2	.64	12.0	.001	20.5	<.001	8.4	.004	3.3	.07	17.1	<.001
Oral contraceptive use	67.5	<.001	527.9	<.001	36.3	<.001	198.4	<.001	19.0	<.001	398.6	<.001
Hormonal replacement therapy use	0.0	.90	4.4	.04	0.0	.94	1.3	.25	4.4	.04	8.7	.003

Abbreviations: 5 α -DHT, 5 α -dihydrotestosterone; BMI, body mass index; DHEAS, dehydroepiandrosterone-sulphate; SHBG, sex hormone binding globulin.

Supplement 3. Mean plasma androgen and SHBG levels in 1659 women according to oral contraceptive pill use.

	Oral contraceptive pill use Geo. mean (95%CI) n = 452	No oral contraceptive pill use Geo. mean (95%CI) n = 1207	p
Total testosterone (nmol/L)			
Unadjusted	0.69 (0.66-0.72)	0.74 (0.72-0.76)	.0060
Adjusted ^a	0.61 (0.58-0.64)	0.78 (0.76-0.80)	<.0001
SHBG (nmol/L)			
Unadjusted	104.7 (100.0-109.8)	53.0 (51.5-54.6)	<.0001
Adjusted ^a	98.8 (94.1-103.7)	54.2 (52.7-55.8)	<.0001
Free testosterone (pmol/L)			
Unadjusted	5.26 (5.02-5.52)	9.54 (9.27-9.83)	<.0001
Adjusted ^a	4.83 (4.60-5.08)	9.86 (9.57-10.15)	<.0001
5α-DHT (nmol/L)			
Unadjusted	0.34 (0.32-0.35)	0.31 (0.30-0.32)	.01
Adjusted ^a	0.28 (0.26-0.29)	0.34 (0.33-0.35)	<.0001
Androstenedione (nmol/L)			
Unadjusted	2.90 (2.76-3.05)	3.17 (3.08-3.27)	.0026
Adjusted ^a	2.31 (2.20-2.42)	3.46 (3.37-3.55)	<.0001
DHEAS (μmol/L)			
Unadjusted	5.19 (4.89-5.51)	4.15 (4.01-4.31)	<.0001
Adjusted ^a	3.96 (3.75-4.19)	4.60 (4.45-4.75)	<.0001

Abbreviations: 5 α -DHT, 5 α -dihydrotestosterone; CI, confidence interval; DHEAS, dehydroepiandrosterone-sulphate; SHBG, sex hormone binding globulin. To convert nmol/L to ng/dl for total testosterone, 5 α -DHT, and androstenedione multiply by 28.84, 29.07, and 28.64, respectively. To convert pmol/L to pg/ml for free testosterone divide by 3.47. To convert μ mol/L to μ g/dl for DHEAS, multiply by 36.85. To convert nmol/L to mg/ml for SHBG, multiply by 0.095. Data are geometric means based on two-way analysis of covariance. ^a Adjusted for age, education, body mass index, smoking, alcohol use, no. of treated chronic diseases, menopausal status, hormonal replacement therapy use, and psychopathology.

Supplement 4. Mean plasma androgen and SHBG levels in 1659 women according to their psychopathology.

	Never MDD Geo. mean (95%CI) n = 392	Remitted MDD Geo. mean (95%CI) n = 754	p	Current MDD Geo. mean (95%CI) n = 513	p
Total testosterone (nmol/L)					
Unadjusted	0.68 (0.65-0.71)	0.74 (0.71-0.77)	.005	0.74 (0.71-0.78)	.007
Adjusted ^a	0.69 (0.66-0.72)	0.74 (0.71-0.76)	.02	0.74 (0.71-0.77)	.02
SHBG (nmol/L)					
Unadjusted	67.7 (63.8-71.8)	64.3 (61.6-67.2)	.18	60.3 (57.3-63.6)	.004
Adjusted ^a	65.4 (62.3-68.9)	64.5 (62.3-66.9)	.66	61.7 (59.2-64.4)	.09
Free testosterone (pmol/L)					
Unadjusted	7.26 (6.86-7.69)	8.23 (7.89-8.58)	.0006	8.66 (8.23-9.12)	<.0001
Adjusted ^a	7.55 (7.19-7.92)	8.17 (7.89-8.47)	.01	8.50 (8.13-8.87)	.0005
5α-DHT (nmol/L)					
Unadjusted	0.31 (0.30-0.33)	0.32 (0.31-0.33)	.60	0.33 (0.31-0.34)	.26
Adjusted ^a	0.32 (0.30-0.33)	0.32 (0.31-0.33)	.90	0.32 (0.31-0.34)	.48
Androstenedione (nmol/L)					
Unadjusted	2.88 (2.73-3.04)	3.11 (2.99-3.23)	.03	3.26 (3.11-3.41)	.0007
Adjusted ^a	2.97 (2.83-3.10)	3.10 (3.00-3.20)	.13	3.20 (3.07-3.33)	.02
DHEAS (μmol/L)					
Unadjusted	4.55 (4.27-4.85)	4.36 (4.16-4.57)	.29	4.39 (4.15-4.65)	.41
Adjusted ^a	4.56 (4.32-4.82)	4.38 (4.22-4.56)	.23	4.35 (4.15-4.56)	.20

Abbreviations: 5α-DHT, 5α-dihydrotestosterone; CI, confidence interval; DHEAS, dehydroepiandrosterone-sulphate; MDD, major depressive disorder; SHBG, sex hormone binding globulin. To convert nmol/L to ng/dl for total testosterone, 5α-DHT, and androstenedione multiply by 28.84, 29.07, and 28.64, respectively. To convert pmol/L to pg/ml for free testosterone divide by 3.47. To convert μmol/L to μg/dl for DHEAS, multiply by 36.85. To convert nmol/L to mg/ml for SHBG, multiply by 0.095. Data are geometric means based on two-way analysis of covariance with women who never experienced a MDD as the reference group. ^a Adjusted for age, education, body mass index, smoking, alcohol use, no. of treated chronic diseases, menopausal status, oral contraceptive pill use, and hormonal replacement therapy use.

Supplement 5. Data distribution of androgen and SHBG levels in women according to their diagnosis at baseline.

Abbreviations: 5α-DHT, 5α-dihydrotestosterone; DHEAS, dehydroepiandrosterone-sulphate. Figures are raincloud plots that combines an illustration of the data distribution with jittered raw data. The black lines in the jittered raw data are the 95% confidence intervals of the geometric means (back transformed means of log transformed data). P-values are based on adjusted mean analyses by two-way analysis of covariance models. Models were adjusted for age, education, body mass index, smoking, alcohol use, number of treated chronic diseases, menopausal status, oral contraceptive pill use, and hormonal replacement therapy use.

Supplement 6. Cross-sectional associations of plasma androgen and SHBG levels and severity of depressive symptoms in 1639 women.

	Severity of depressive symptoms	
	β (95%CI)	p
Total testosterone		
Unadjusted	0.006 (-0.003 to 0.014)	.18
Adjusted ^a	0.008 (-0.001 to 0.016)	.08
SHBG		
Unadjusted	-0.012 (-0.020 to -0.003)	.007
Adjusted ^a	-0.004 (-0.015 to 0.006)	.38
Free testosterone		
Unadjusted	0.013 (0.005 to 0.022)	.002
Adjusted ^a	0.011 (0.001 to 0.020)	.03
5α-DHT		
Unadjusted	0.002 (-0.006 to 0.010)	.63
Adjusted ^a	0.006 (-0.003 to 0.015)	.17
Androstenedione		
Unadjusted	0.006 (-0.003 to 0.014)	.18
Adjusted ^a	0.006 (-0.003 to 0.016)	.21
DHEAS		
Unadjusted	-0.009 (-0.018 to -0.001)	.03
Adjusted ^a	-0.008 (-0.017 to 0.002)	.12

Abbreviations: 5 α -DHT, 5 α -dihydrotestosterone; β , standardized beta-coefficient; CI, confidence interval; DHEAS, dehydroepiandrosterone-sulphate; SHBG, sex hormone binding globulin. Data are β s for 1 standard deviation change in each biomarker based on linear regression models. ^a Adjusted for age, level of education, body mass index, smoking, alcohol use, number of treated chronic diseases, menopausal status, oral contraceptive pill use, and hormonal replacement therapy use.

Supplement 7. Equivalence testing for mean plasma androgen and SHBG levels in 1659 women according to their psychopathology.

	Remitted MDD vs. Never MDD			Current MDD vs. Never MDD		
	mean difference (99.25%CI)	smallest effect size of interest	p for equivalence	mean difference (99.25%CI)	smallest effect size of interest	p for equivalence
Total testosterone (nmol/L)	-0.05 (-0.10, 0.00)	0.15	<.001	-0.05 (-0.11, 0.01)	0.17	<.001
SHBG (nmol/L)	1.00 (-4.50, 6.50)	16.3	<.001	3.10 (-2.60, 8.80)	14.0	<.001
Free testosterone (pmol/L)	-0.62 (-1.26, 0.02)	1.97	<.001	-0.95 (-1.66, -0.24)	1.80	<.001
5α-DHT (nmol/L)	0.00 (-0.02, 0.02)	0.07	<.001	0.00 (-0.03, 0.03)	0.07	<.001
Androstenedione (nmol/L)	-0.13 (-0.36, 0.10)	0.70	<.001	-0.23 (-0.49, 0.03)	0.64	<.001
DHEAS (μ mol/L)	0.18 (-0.25, 0.61)	1.23	<.001	0.21 (-0.24, 0.66)	1.10	<.001

Abbreviations: 5 α -DHT, 5 α -dihydrotestosterone; CI, confidence interval; DHEAS, dehydroepiandrosterone-sulphate; MDD, major depressive disorder; SHBG, sex hormone binding globulin. The mean difference with 99.25%CI was based on the difference between adjusted mean of the two groups that were compared. The 99.25% range was used since cut-off for significance $p < 0.0075$. For the actual means per group we refer to *Supplement 4*. The smallest effect size of interest was calculated on $0.5 \times$ the pooled standard deviation of the baseline levels of the two groups analyzed.

Chapter 4

Supplement 8. Baseline characteristics of men with and without follow-up assessments.

	MDD assessment at 2 years of follow-up		p ^a	Individual depressive symptom assessments at 1 and 2 years of follow-up		
	Sample with follow-up n = 504	Sample without follow-up n = 60		Sample with follow-up n = 751	Sample without follow-up n = 72	p ^a
Age , mean (SD)	43.4 (13.5)	44.2 (13.1)	.63	43.8 (12.7)	41.4 (13.4)	.16
Education , mean (SD)	12.5 (3.3)	11.8 (3.2)	.09	12.1 (3.3)	11.3 (3.3)	.06
Body Mass Index , mean (SD)	26.0 (4.4)	26.8 (5.1)	.26	26.2 (4.6)	26.7 (4.0)	.40
Current smoking , no. (%)	194 (38.5)	27 (45.0)	.40	297 (39.5)	45 (62.5)	.0003
>1 alcohol unit/day , no. (%)	156 (31.0)	16 (26.7)	.59	225 (30.0)	22 (30.6)	1.00
≥1 treated chronic diseases , no. (%)	188 (37.3)	32 (53.3)	.31	311 (41.4)	28 (38.9)	.91
Current antidepressant use , no. (%)	83 (16.5)	12 (20.0)	.61	188 (25.0)	17 (23.6)	.90
Current psychopathology , no. (%)						
Never MDD	232 (46.0%)	14 (23.3%)	.001	233 (31.0%)	13 (18.1%)	.06
Remitted MDD	272 (54.0%)	46 (76.7%)		284 (37.8%)	34 (47.2%)	
Current MDD	-	-		234 (31.2%)	25 (34.7%)	
Current anxiety disorder , no. (%)	90 (17.9%)	17 (28.3%)	.07	232 (30.9%)	28 (38.9%)	.21
<i>Predictors - Androgens</i>						
Low total testosterone (<11 nmol/L), no. (%)	66 (13.1)	12 (20.0)	.21	113 (15.0)	13 (18.1)	.61
Low free testosterone (<220 pmol/L), no. (%)	45 (8.9)	11 (18.3)	.04	77 (10.3)	10 (13.9)	.44
Total testosterone (nmol/L), mean (SD)	16.6 (5.7)	15.4 (5.0)	.09	16.5 (5.8)	15.4 (5.4)	.10
SHBG (nmol/L), mean (SD)	36.4 (15.7)	33.9 (13.1)	.18	36.3 (15.8)	32.2 (11.8)	.008
Free testosterone (pmol/L), mean (SD)	330.0 (98.4)	318.5 (104.6)	.42	328.0 (101.1)	324.8 (103.7)	.81
5α-DHT (nmol/L), mean (SD)	1.5 (0.6)	1.4 (0.6)	.35	1.5 (0.6)	1.4 (0.6)	.16
Androstenedione (nmol/L), mean (SD)	3.8 (1.4)	3.6 (1.5)	.43	3.8 (1.5)	4.2 (1.8)	.11
DHEAS (μmol/L), mean (SD)	8.0 (4.3)	7.7 (4.4)	.57	7.9 (4.1)	8.9 (4.8)	.09

Abbreviations: 5α-DHT, 5α-dihydrotestosterone; BMI, body mass index; DHEAS, dehydroepiandrosterone-sulphate; MDD, major depressive disorder; SD, standard deviation; SHBG, sex hormone binding globulin. To convert nmol/L to ng/dl for total t, 5α-DHT, and androstenedione multiply by 28.84, 29.07, and 28.64, respectively. To convert pmol/L to pg/ml for free testosterone divide by 3.47. To convert μmol/L to μg/dl for DHEAS, multiply by 36.85. To convert to nmol/L to mg/ml for SHBG, multiply by 0.095. ^a Based on two-way analyses of variance for independent samples or χ² tests for normally distributed variables and by Kruskal Wallis testing for non-normally distributed variables (severity of MDD symptoms).

Supplement 9. Flow diagram of men included in the analyses to test each of the three hypotheses.

Abbreviations: MDD, major depressive disorder; NESDA, Netherlands Study on Depression and Anxiety. First hypothesis: lower androgen levels are equally present in men with and without MDD. Second hypothesis: lower androgen levels are not associated with an elevated risk for new onset of MDD. Third hypothesis: lower androgen levels are only associated with some overlapping symptoms of MDD and hypogonadism such as less sexual interest and more fatigue.

Supplement 10. Mean plasma androgen and SHBG levels in 823 men according to their psychopathology.

	Without lifetime MDD mean (95%CI) (n = 246)	Remitted MDD mean (95%CI) (n = 318)	p	Current MDD mean (95%CI) (n = 259)	p
Total testosterone (nmol/L)					
Unadjusted	16.8 (16.1; 17.5)	16.2 (15.6; 16.9)	.25	16.2 (15.5; 16.9)	.28
Adjusted ^a	16.6 (15.8; 17.4)	16.6 (16.0; 17.2)	.98	16.7 (16.0; 17.4)	.84
SHBG (nmol/L)					
Unadjusted	36.2 (34.3; 38.2)	36.1 (34.4; 37.8)	.89	35.6 (33.7; 37.5)	.62
Adjusted ^a	36.5 (34.3; 38.7)	36.0 (34.4; 37.7)	.73	35.9 (34.1; 37.7)	.69
Free testosterone (pmol/L)					
Unadjusted	336 (323; 348)	323 (312; 335)	.15	325 (313; 338)	.25
Adjusted ^a	331 (319; 344)	333 (322; 342)	.87	334 (323; 345)	.74
5α-DHT (nmol/L)					
Unadjusted	1.46 (1.38; 1.54)	1.54 (1.47; 1.60)	.15	1.53 (1.45; 1.60)	.23
Adjusted ^a	1.46 (1.37; 1.55)	1.56 (1.49; 1.63)	.07	1.55 (1.48; 1.63)	.13
Androstenedione (nmol/L)					
Unadjusted	3.8 (3.6; 4.0)	3.8 (3.6; 4.0)	.99	4.0 (3.8; 4.2)	.14
Adjusted ^a	4.0 (3.8; 4.2)	3.9 (3.8; 4.1)	.65	4.1 (3.9; 4.2)	.59
DHEAS (μmol/L)					
Unadjusted	8.1 (7.6; 8.6)	7.9 (7.5; 8.4)	.64	7.9 (7.4; 8.4)	.65
Adjusted ^a	8.1 (7.6; 8.7)	8.4 (8.0; 8.8)	.46	8.2 (7.8; 8.7)	.77

Abbreviations: 5 α -DHT, 5 α -dihydrotestosterone; CI, Confidence interval; DHEAS, dehydroepiandrosterone-sulphate; MDD, major depressive disorder; SHBG, sex hormone binding globulin. Data are estimated means based on two-way analyses of covariance analysis with men who never experienced a MDD as the reference group. To convert nmol/L to ng/dl for total testosterone, 5 α -DHT, and androstenedione multiply by 28.84, 29.07, and 28.64, respectively. To convert pmol/L to pg/ml for free testosterone divide by 3.47. To convert μ mol/L to μ g/dl for DHEAS, multiply by 36.85. To convert to nmol/L to mg/ml for SHBG, multiply by 0.095. ^a Adjusted for age, waist circumference, number of treated chronic somatic disorders, smoking status, alcohol use, physical activity, the presence of a current anxiety disorder.

Supplement 11. Equivalence testing for mean plasma androgen and SHBG levels in 823 men according to their psychopathology.

Remitted MDD vs. without lifetime MDD						Current MDD vs. without lifetime MDD					
	mean difference (95%CI)	smallest effect size of interest	p	p ^a	conclusion	mean difference (95%CI)	smallest effect size of interest	p	p ^a	conclusion	
Total testosterone (nmol/L)	0.01 (-1.00; 1.02)	1.20	0.01	.98	not different from 0 & equivalent to 0	-0.11 (-1.15; 0.93)	1.19	0.02	.84	not different from 0 & equivalent to 0	
SHBG (nmol/L)	-0.11 (-1.15; 0.93)	1.19	0.02	.73	not different from 0 & equivalent to 0	-0.11 (-1.15; 0.93)	1.19	0.02	.69	not different from 0 & equivalent to 0	
Free testosterone (pmol/L)	-1.28 (-17.68; 15.12)	19.48	0.02	.87	not different from 0 & equivalent to 0	-2.88 (-19.79; 14.03)	19.29	0.03	.74	not different from 0 & equivalent to 0	
5α-DHT (nmol/L)	-0.10 (-0.21; 0.01)	0.13	0.28	.07	Inconclusive: not different from 0 but also not equivalent to 0	-0.09 (-0.21; 0.03)	0.13	0.23	.13	Inconclusive: not different from 0 but also not equivalent to 0	
Androstenedione (nmol/L)	0.05 (-0.22; 0.32)	0.32	0.004	.65	not different from 0 & equivalent to 0	-0.08 (-0.35; 0.19)	0.31	0.048	.59	not different from 0 & equivalent to 0	
DHEAS (μmol/L)	-0.23 (-0.89; 0.43)	0.79	0.049	.46	not different from 0 & equivalent to 0	-0.10 (-0.79; 0.59)	0.78	0.03	.77	not different from 0 & equivalent to 0	

Abbreviations: 5α-DHT, 5α-dihydrotestosterone; CI, confidence interval; DHEAS, dehydroepiandrosterone-sulphate; MDD, major depressive disorder; SHBG, sex hormone binding globulin. The mean difference with 95%CI was based on the difference between adjusted mean of the two groups that were compared. The smallest effect size of interest was calculated on $0.5 \times$ the pooled standard deviation of the baseline levels of the two groups analyzed. ^a P-values based on the classical two-way analyses of covariance testing as reported in *Supplement 10*.

Supplement 12. Associations of total testosterone and SHBG with individual MDD symptoms.

	Total testosterone (nmol/L)				SHBG (nmol/L)			
	β (95%CI)	p	β (95%CI) ^a	p	β (95%CI)	p	β (95%CI) ^a	p
Total score	-0.061 (-0.123; 0.001)	.05	-0.006 (-0.065; 0.053)	.84	-0.008 (-0.070; 0.054)	.81	-0.000 (-0.058; 0.058)	1.00
1. Difficulty falling asleep	0.024 (-0.037; 0.084)	.44	0.024 (-0.04; 0.088)	.45	0.000 (-0.061; 0.060)	.99	0.014 (-0.049; 0.077)	.67
2. Difficulty staying asleep	-0.102 (-0.159; -0.045)	.0005	-0.005 (-0.064; 0.055)	.88	0.056 (-0.002; 0.113)	.06	0.027 (-0.032; 0.085)	.37
3. Waking up early	-0.042 (-0.098; 0.015)	.15	0.017 (-0.044; 0.079)	.57	0.061 (0.004; 0.118)	.04	0.032 (-0.028; 0.092)	.29
4. Hypersomnia	-0.024 (-0.082; 0.034)	.42	0.005 (-0.058; 0.068)	.87	0.035 (-0.023; 0.093)	.23	0.034 (-0.028; 0.095)	.28
5. Sad mood	-0.024 (-0.082; 0.035)	.43	-0.002 (-0.061; 0.057)	.94	-0.021 (-0.080; 0.037)	.47	0.001 (-0.057; 0.059)	.97
6. Irritability	-0.052 (-0.109; 0.006)	.08	-0.031 (-0.091; 0.029)	.30	-0.067 (-0.125; -0.009)	.02	-0.046 (-0.104; 0.013)	.12
7. Anxious or tense	-0.033 (-0.092; 0.026)	.27	-0.003 (-0.059; 0.053)	.91	-0.040 (-0.099; 0.019)	.18	-0.022 (-0.077; 0.033)	.43
8. Responsiveness of mood	-0.036 (-0.091; 0.018)	.19	-0.012 (-0.07; 0.046)	.67	-0.009 (-0.064; 0.046)	.74	-0.023 (-0.08; 0.034)	.44
9. Diurnal mood variation	-0.028 (-0.083; 0.027)	.32	-0.055 (-0.114; 0.005)	.07	-0.027 (-0.082; 0.028)	.33	-0.025 (-0.084; 0.033)	.39
10. Quality mood	-0.042 (-0.098; 0.014)	.14	-0.021 (-0.078; 0.036)	.46	-0.027 (-0.083; 0.029)	.34	-0.013 (-0.069; 0.043)	.65
11. Decrease in appetite	0.039 (-0.012; 0.089)	.79	0.026 (-0.03; 0.081)	.37	0.039 (-0.012; 0.090)	.13	0.039 (-0.016; 0.093)	.16
12. Increase in appetite	-0.097 (-0.148; -0.045)	.64	-0.027 (-0.083; 0.029)	.35	-0.086 (-0.138; -0.034)	.001	-0.035 (-0.09; 0.02)	.21
13. Decrease in weight	-0.045 (-0.091; 0.001)	.02	-0.024 (-0.075; 0.027)	.36	-0.023 (-0.069; 0.023)	.33	-0.008 (-0.058; 0.042)	.75
14. Increase in weight	-0.055 (-0.102; -0.008)	.85	-0.010 (-0.062; 0.042)	.70	-0.041 (-0.088; 0.006)	.09	-0.014 (-0.065; 0.038)	.60
15. Poor concentration	-0.008 (-0.065; 0.049)	.43	0.009 (-0.049; 0.067)	.76	0.006 (-0.051; 0.063)	.83	0.019 (-0.038; 0.077)	.50
16. Low self-esteem	0.014 (-0.045; 0.073)	.34	0.009 (-0.052; 0.067)	.76	-0.015 (-0.074; 0.043)	.61	0.000 (-0.06; 0.06)	1.00
17. Pessimism	-0.073 (-0.133; -0.012)	.23	-0.018 (-0.079; 0.044)	.57	0.000 (-0.060; 0.061)	.99	0.007 (-0.054; 0.067)	.83
18. Suicidality	-0.006 (-0.062; 0.051)	.005	0.001 (-0.06; 0.062)	.98	0.013 (-0.043; 0.070)	.64	0.000 (-0.061; 0.06)	.99
19. Low general interest	-0.023 (-0.080; 0.034)	.27	0.004 (-0.056; 0.064)	.89	-0.018 (-0.075; 0.039)	.54	-0.002 (-0.061; 0.056)	.94
20. Energy level	-0.028 (-0.085; 0.029)	.55	0.028 (-0.03; 0.086)	.35	-0.005 (-0.062; 0.052)	.86	0.010 (-0.048; 0.067)	.74
21. Low capacity for pleasure	-0.035 (-0.093; 0.022)	.0001	-0.002 (-0.062; 0.058)	.95	0.007 (-0.050; 0.065)	.80	0.012 (-0.047; 0.071)	.69
22. Low interest in sex	-0.086 (-0.146; -0.025)	.0004	-0.019 (-0.082; 0.044)	.56	0.048 (-0.013; 0.108)	.12	0.044 (-0.018; 0.106)	.16
23. Psychomotor retardation	-0.032 (-0.088; 0.025)	.30	0.030 (-0.03; 0.089)	.32	0.015 (-0.042; 0.071)	.61	0.030 (-0.028; 0.089)	.30
24. Psychomotor agitation	-0.018 (-0.075; 0.040)	.14	-0.014 (-0.074; 0.047)	.65	-0.008 (-0.065; 0.049)	.78	-0.010 (-0.069; 0.049)	.73
25. Aches and pains	-0.119 (-0.177; -0.06)	.50	-0.018 (-0.076; 0.04)	.54	-0.018 (-0.077; 0.041)	.56	-0.018 (-0.075; 0.040)	.55

	Total testosterone (nmol/L)				SHBG (nmol/L)			
	β (95%CI)	p	β (95%CI) ^a	p	β (95%CI)	p	β (95%CI) ^a	p
26. Sympathetic arousal	-0.106 (-0.164; -0.048)	.11	-0.026 (-0.083; 0.031)	.37	-0.025 (-0.083; 0.034)	.41	-0.013 (-0.070; 0.043)	.64
27. Panic or phobia	-0.031 (-0.090; 0.028)	.14	-0.017 (-0.072; 0.038)	.54	-0.044 (-0.103; 0.014)	.14	-0.025 (-0.079; 0.029)	.36
28. Constipation or diarrhea	-0.045 (-0.104; 0.014)	.0003	-0.008 (-0.07; 0.053)	.79	0.005 (-0.055; 0.064)	.88	-0.002 (-0.062; 0.059)	.95
29. Interpersonal sensitivity	-0.020 (-0.080; 0.039)	.06	-0.019 (-0.081; 0.043)	.54	-0.070 (-0.130; -0.011)	.02	-0.044 (-0.105; 0.017)	.16
30. Leaden paralysis	-0.047 (-0.106; 0.011)	.02	0.007 (-0.051; 0.064)	.82	-0.012 (-0.070; 0.047)	.69	-0.007 (-0.064; 0.049)	.80

Abbreviations: CI, Confidence interval; SHBG, sex hormone binding globulin; MDD, major depressive disorder. Table shows unadjusted and adjusted standardized β 's with 95%CI for the association between baseline total testosterone and SHBG levels and the MDD symptom severity item score over two years examined with a linear mixed model. ^aAdjusted for age, waist circumference, number of treated chronic somatic disorders, smoking status, alcohol use, physical activity, the presence of a current anxiety disorder.

Supplement 13. Associations of free testosterone and 5 α -DHT with individual MDD symptoms.

	Free testosterone (pmol/L)				5 α -DHT (nmol/L)			
	β (95%CI)	p	β (95%CI) ^a	p	β (95%CI)	p	β (95%CI) ^a	p
Total score	-0.079 (-0.140; -0.017)	.01	-0.019 (-0.082; 0.044)	.55	-0.010 (-0.072; 0.051)	.75	0.013 (-0.043; 0.069)	.65
1. Difficulty falling asleep	0.038 (-0.022; 0.098)	.22	0.027 (-0.041; 0.096)	.44	0.021 (-0.039; 0.081)	.50	0.013 (-0.048; 0.075)	.67
2. Difficulty staying asleep	-0.190 (-0.246; -0.135)	<.0001	-0.039 (-0.103; 0.025)	.23	-0.048 (-0.105; 0.009)	.10	0.000 (-0.057; 0.057)	1.00
3. Waking up early	-0.112 (-0.168; -0.056)	<.0001	-0.009 (-0.075; 0.057)	.79	0.006 (-0.051; 0.062)	.85	0.035 (-0.023; 0.094)	.24
4. Hypersomnia	-0.064 (-0.121; -0.006)	.03	-0.030 (-0.097; 0.037)	.38	-0.008 (-0.066; 0.049)	.78	0.000 (-0.061; 0.06)	.99
5. Sad mood	-0.014 (-0.073; 0.044)	.63	-0.010 (-0.073; 0.054)	.76	-0.008 (-0.067; 0.050)	.78	-0.005 (-0.062; 0.051)	.85
6. Irritability	-0.014 (-0.072; 0.044)	.64	-0.008 (-0.072; 0.056)	.79	-0.028 (-0.086; 0.030)	.34	-0.017 (-0.074; 0.04)	.56
7. Anxious or tense	-0.016 (-0.074; 0.043)	.60	0.005 (-0.055; 0.064)	.88	0.008 (-0.051; 0.066)	.80	0.011 (-0.042; 0.064)	.69
8. Responsiveness of mood	-0.047 (-0.102; 0.008)	.09	-0.006 (-0.069; 0.056)	.84	-0.020 (-0.074; 0.035)	.49	-0.021 (-0.077; 0.034)	.45
9. Diurnal mood variation	-0.006 (-0.060; 0.049)	.84	-0.049 (-0.113; 0.015)	.13	-0.025 (-0.08; 0.030)	.37	-0.047 (-0.105; 0.01)	.10
10. Quality mood	-0.032 (-0.088; 0.023)	.25	-0.022 (-0.083; 0.039)	.49	-0.008 (-0.064; 0.047)	.77	-0.001 (-0.056; 0.053)	.97
11. Decrease in appetite	0.018 (-0.033; 0.068)	.50	-0.004 (-0.064; 0.056)	.88	0.047 (-0.004; 0.097)	.07	0.029 (-0.025; 0.083)	.29
12. Increase in appetite	-0.056 (-0.108; -0.004)	.03	-0.005 (-0.066; 0.055)	.86	-0.090 (-0.142; -0.038)	.001	-0.035 (-0.088; 0.019)	.21
13. Decrease in weight	-0.044 (-0.090; 0.002)	.06	-0.035 (-0.09; 0.019)	.20	-0.025 (-0.071; 0.021)	.29	-0.005 (-0.054; 0.044)	.84

	Free testosterone (pmol/L)				5 α -DHT (nmol/L)			
	β (95%CI)	p	β (95%CI) ^a	p	β (95%CI)	p	β (95%CI) ^a	p
14. Increase in weight	-0.029 (-0.075; 0.018)	.23	0.014 (-0.042; 0.07)	.61	-0.041 (-0.088; 0.006)	.09	-0.006 (-0.056; 0.044)	.81
15. Poor concentration	-0.018 (-0.075; 0.039)	.53	-0.014 (-0.077; 0.048)	.50	0.047 (-0.010; 0.104)	.10	0.053 (-0.002; 0.109)	.06
16. Low self-esteem	0.029 (-0.029; 0.087)	.33	0.006 (-0.059; 0.071)	.85	0.040 (-0.018; 0.099)	.18	0.028 (-0.03; 0.087)	.34
17. Pessimism	-0.101 (-0.161; -0.041)	.001	-0.040 (-0.106; 0.026)	.23	-0.024 (-0.084; 0.037)	.44	-0.001 (-0.06; 0.058)	.97
18. Suicidality	-0.022 (-0.079; 0.034)	.44	-0.005 (-0.071; 0.061)	.88	-0.002 (-0.059; 0.054)	.94	-0.012 (-0.071; 0.047)	.69
19. Low general interest	-0.017 (-0.074; 0.040)	.55	0.003 (-0.061; 0.067)	.93	-0.008 (-0.064; 0.049)	.79	0.003 (-0.055; 0.06)	.93
20. Energy level	-0.045 (-0.102; 0.011)	.12	0.014 (-0.049; 0.076)	.67	0.012 (-0.044; 0.069)	.67	0.040 (-0.016; 0.096)	.16
21. Low capacity for pleasure	-0.057 (-0.115; 0.000)	.05	-0.024 (-0.088; 0.040)	.46	0.013 (-0.045; 0.070)	.66	0.024 (-0.033; 0.082)	.41
22. Low interest in sex	-0.157 (-0.216; -0.098)	<.0001	-0.077 (-0.144; -0.010)	.02	-0.016 (-0.077; 0.044)	.59	0.01 (-0.049; 0.071)	.73
23. Psychomotor retardation	-0.065 (-0.121; -0.009)	.02	-0.001 (-0.064; 0.063)	.98	0.011 (-0.045; 0.067)	.70	0.044 (-0.012; 0.101)	.12
24. Psychomotor agitation	-0.015 (-0.072; 0.042)	.61	-0.012 (-0.077; 0.052)	.71	0.015 (-0.042; 0.072)	.60	0.014 (-0.044; 0.071)	.65
25. Aches and pains	-0.149 (-0.207; -0.091)	<.0001	-0.017 (-0.08; 0.045)	.58	-0.051 (-0.110; 0.007)	.09	0.015 (-0.041; 0.07)	.60
26. Sympathetic arousal	-0.125 (-0.183; -0.068)	<.0001	-0.035 (-0.096; 0.026)	.26	-0.028 (-0.086; 0.031)	.35	0.025 (-0.03; 0.08)	.37
27. Panic or phobia	-0.001 (-0.059; 0.058)	.98	-0.003 (-0.062; 0.056)	.92	-0.002 (-0.061; 0.056)	.93	-0.010 (-0.062; 0.043)	.71
28. Constipation or diarrhea	-0.063 (-0.122; -0.004)	.04	-0.013 (-0.079; 0.053)	.70	0.003 (-0.056; 0.062)	.92	0.028 (-0.031; 0.087)	.35
29. Interpersonal sensitivity	0.033 (-0.026; 0.092)	.27	0.010 (-0.056; 0.077)	.76	-0.032 (-0.091; 0.027)	.29	-0.039 (-0.099; 0.02)	.19
30. Leadon paralysis	-0.060 (-0.118; -0.002)	.04	0.005 (-0.056; 0.067)	.86	0.009 (-0.049; 0.067)	.76	0.030 (-0.026; 0.085)	.29

Abbreviations: CI, Confidence interval; 5 α -DHT, 5 α -dihydrotestosterone; MDD, major depressive disorder. Table shows unadjusted and adjusted standardized β 's with 95%CI for the association between baseline free testosterone and 5 α -DHT levels and the MDD symptom severity item score over two years examined with a linear mixed model. ^a Adjusted for age, waist circumference, number of treated chronic somatic disorders, smoking status, alcohol use, physical activity, the presence of a current anxiety disorder.

Supplement 14. Associations of androstenedione and DHEAS with individual depressive symptoms.

	Androstenedione (nmol/L)				DHEAS (μmol/L)			
	β (95%CI)	p	β (95%CI) ^a	p	β (95%CI)	p	β (95%CI) ^a	p
Total score	0.043 (-0.018; 0.104)	.17	0.042 (-0.015; 0.099)	.15	-0.032 (-0.094; 0.029)	.30	0.024 (-0.040; 0.088)	.45
1. Difficulty falling asleep	0.091 (0.031; 0.151)	.003	0.045 (-0.017; 0.107)	.15	0.012 (-0.048; 0.072)	.69	-0.024 (-0.094; 0.045)	.49
2. Difficulty staying asleep	-0.085 (-0.142; -0.029)	.003	0.002 (-0.056; 0.06)	.95	-0.181 (-0.237; -0.125)	<.0001	-0.028 (-0.093; 0.037)	.40
3. Waking up early	-0.061 (-0.117; -0.004)	.04	-0.006 (-0.065; 0.054)	.85	-0.100 (-0.156; -0.043)	.001	0.013 (-0.054; 0.08)	.70
4. Hypersomnia	0.011 (-0.047; 0.068)	.72	0.013 (-0.048; 0.074)	.66	-0.061 (-0.118; -0.004)	.04	-0.025 (-0.093; 0.044)	.48
5. Sad mood	0.077 (0.019; 0.135)	.01	0.054 (-0.004; 0.111)	.07	0.008 (-0.050; 0.066)	.79	0.002 (-0.062; 0.067)	.95
6. Irritability	0.038 (-0.020; 0.095)	.20	0.015 (-0.043; 0.073)	.60	0.044 (-0.014; 0.101)	.14	0.042 (-0.023; 0.107)	.20
7. Anxious or tense	0.084 (0.026; 0.142)	.005	0.058 (0.004; 0.111)	.04	0.042 (-0.016; 0.100)	.16	0.053 (-0.007; 0.114)	.08
8. Responsiveness of mood	0.034 (-0.021; 0.089)	.22	0.033 (-0.024; 0.09)	.25	-0.020 (-0.075; 0.035)	.48	0.021 (-0.043; 0.084)	.52
9. Diurnal mood variation	0.081 (0.027; 0.135)	.004	0.042 (-0.016; 0.1)	.15	0.060 (0.006; 0.115)	.03	0.036 (-0.029; 0.101)	.28
10. Quality mood	0.060 (0.005; 0.115)	.03	0.038 (-0.017; 0.093)	.17	0.019 (-0.037; 0.074)	.50	0.023 (-0.038; 0.085)	.46
11. Decrease in appetite	0.067 (0.016; 0.118)	.01	0.026 (-0.029; 0.08)	.36	0.025 (-0.026; 0.075)	.34	0.006 (-0.055; 0.066)	.86
12. Increase in appetite	-0.032 (-0.084; 0.021)	.23	-0.010 (-0.064; 0.045)	.73	-0.034 (-0.086; 0.018)	.20	-0.012 (-0.073; 0.049)	.70
13. Decrease in weight	0.009 (-0.037; 0.055)	.70	0.009 (-0.040; 0.059)	.71	0.027 (-0.019; 0.073)	.25	0.056 (0.001; 0.112)	.046
14. Increase in weight	-0.062 (-0.109; -0.016)	.01	-0.052 (-0.102; -0.001)	.045	-0.047 (-0.094; 0.000)	.05	-0.030 (-0.087; 0.027)	.30
15. Poor concentration	0.045 (-0.011; 0.102)	.12	0.017 (-0.039; 0.074)	.54	0.020 (-0.036; 0.077)	.48	0.021 (-0.043; 0.084)	.52
16. Low self-esteem	0.071 (0.013; 0.129)	.02	0.036 (-0.023; 0.095)	.23	0.111 (0.053; 0.169)	.0002	0.102 (0.036; 0.168)	.003
17. Pessimism	0.061 (0.001; 0.121)	.05	0.085 (0.026; 0.145)	.005	-0.014 (-0.074; 0.046)	.65	0.069 (0.002; 0.136)	.04
18. Suicidality	0.064 (0.008; 0.120)	.03	0.063 (0.004; 0.123)	.04	0.030 (-0.026; 0.087)	.29	0.068 (0.001; 0.134)	.046
19. Low general interest	0.017 (-0.039; 0.074)	.55	-0.001 (-0.059; 0.057)	.97	0.017 (-0.040; 0.074)	.56	0.030 (-0.035; 0.095)	.36
20. Energy level	0.033 (-0.023; 0.089)	.25	0.030 (-0.026; 0.086)	.29	-0.059 (-0.115; -0.002)	.04	-0.026 (-0.089; 0.037)	.41
21. Low capacity for pleasure	0.019 (-0.038; 0.077)	.50	0.007 (-0.051; 0.065)	.81	-0.013 (-0.070; 0.045)	.67	0.022 (-0.043; 0.087)	.51
22. Low interest in sex	-0.012 (-0.072; 0.048)	.69	0.005 (-0.056; 0.066)	.87	-0.100 (-0.160; -0.041)	.001	-0.008 (-0.076; 0.06)	.82
23. Psychomotor retardation	0.021 (-0.035; 0.077)	.47	0.036 (-0.021; 0.094)	.21	-0.040 (-0.096; 0.016)	.16	0.010 (-0.054; 0.074)	.76
24. Psychomotor agitation	0.003 (-0.053; 0.060)	.91	-0.031 (-0.090; 0.027)	.29	0.002 (-0.055; 0.059)	.94	-0.008 (-0.074; 0.057)	.80

	Androstenedione (nmol/L)				DHEAS (μmol/L)			
	β (95%CI)	p	β (95%CI) ^a	p	β (95%CI)	p	β (95%CI) ^a	p
25. Aches and pains	-0.022 (-0.081; 0.036)	.46	0.028 (-0.029; 0.084)	.33	-0.160 (-0.218; -0.103)	<.0001	-0.056 (-0.119; 0.006)	.08
26. Sympathetic arousal	0.000 (-0.058; 0.058)	1.00	0.023 (-0.033; 0.078)	.42	-0.087 (-0.145; -0.029)	.004	-0.005 (-0.067; 0.058)	.89
27. Panic or phobia	0.064 (0.006; 0.122)	.03	0.033 (-0.020; 0.086)	.23	0.022 (-0.037; 0.080)	.46	-0.002 (-0.062; 0.057)	.94
28. Constipation or diarrhea	-0.019 (-0.077; 0.040)	.54	-0.028 (-0.087; 0.032)	.36	-0.046 (-0.105; 0.013)	.12	0.000 (-0.067; 0.067)	.99
29. Interpersonal sensitivity	0.064 (0.005; 0.123)	.03	0.032 (-0.028; 0.092)	.30	0.091 (0.033; 0.150)	.002	0.063 (-0.004; 0.130)	.07
30. Leadon paralysis	0.045 (-0.013; 0.103)	.12	0.043 (-0.013; 0.099)	.13	-0.055 (-0.114; 0.003)	.06	-0.013 (-0.075; 0.050)	.69

Abbreviations: CI, Confidence interval; DHEAS, dehydroepiandrosterone-sulphate; MDD, major depressive disorder. Table shows unadjusted and adjusted standardized β's with 95%CI for the association between baseline androstenedione and DHEAS levels and the MDD symptom severity item score over two years examined with a linear mixed model. ^a Adjusted for age, waist circumference, number of treated chronic somatic disorders, smoking status, alcohol use, physical activity, the presence of a current anxiety disorder.

Supplement 15. Associations of low total and free testosterone with individual MDD symptoms.

	Low total testosterone (<11 nmol/L)				Low free testosterone (<220 pmol/L)			
	β (95%CI)	p	β (95%CI) ^a	p	β (95%CI)	p	βb (95%CI) ^a	p
Total score	0.223 (0.051; 0.394)	.01	0.083 (-0.074; 0.239)	.30	0.232 (0.030; 0.434)	.02	0.045 (-0.138; 0.227)	.63
1. Difficulty falling asleep	-0.009 (-0.177; 0.16)	.92	-0.045 (-0.215; 0.125)	.60	0.003 (-0.195; 0.202)	.97	-0.024 (-0.222; 0.174)	.81
2. Difficulty staying asleep	0.338 (0.179; 0.497)	<.0001	0.165 (0.007; 0.323)	.04	0.334 (0.147; 0.521)	.0005	0.054 (-0.13; 0.238)	.56
3. Waking up early	0.108 (-0.051; 0.266)	.18	-0.016 (-0.179; 0.147)	.85	0.143 (-0.043; 0.330)	.13	-0.057 (-0.247; 0.132)	.55
4. Hypersomnia	0.122 (-0.038; 0.283)	.14	0.039 (-0.128; 0.206)	.65	0.332 (0.144; 0.520)	.0006	0.249 (0.055; 0.442)	.01
5. Sad mood	0.094 (-0.069; 0.257)	.26	0.029 (-0.128; 0.187)	.72	0.088 (-0.104; 0.280)	.37	0.031 (-0.153; 0.214)	.74
6. Irritability	0.177 (0.016; 0.338)	.03	0.103 (-0.056; 0.262)	.20	0.023 (-0.167; 0.213)	.81	-0.058 (-0.242; 0.127)	.54
7. Anxious or tense	0.128 (-0.036; 0.291)	.13	0.030 (-0.118; 0.178)	.69	0.063 (-0.129; 0.256)	.52	-0.040 (-0.212; 0.132)	.65
8. Responsiveness of mood	0.106 (-0.047; 0.259)	.17	0.049 (-0.105; 0.204)	.53	0.093 (-0.088; 0.274)	.31	0.014 (-0.167; 0.194)	.88
9. Diurnal mood variation	0.140 (-0.012; 0.292)	.07	0.192 (0.033; 0.351)	.02	0.105 (-0.074; 0.284)	.25	0.167 (-0.018; 0.353)	.08
10. Quality mood	0.139 (-0.016; 0.294)	.08	0.072 (-0.08; 0.223)	.35	0.166 (-0.017; 0.349)	.07	0.107 (-0.069; 0.283)	.23
11. Decrease in appetite	0.009 (-0.133; 0.151)	.16	0.045(-0.103; 0.194)	.55	-0.017 (-0.185; 0.150)	.84	0.011 (-0.162; 0.184)	.9
12. Increase in appetite	0.224 (0.079; 0.369)	.80	0.055 (-0.095; 0.204)	.47	0.035 (-0.137; 0.207)	.69	-0.145 (-0.32; 0.029)	.10
13. Decrease in weight	0.052 (-0.077; 0.180)	.002	-0.010 (-0.145; 0.125)	.88	-0.023 (-0.174; 0.128)	.77	-0.094 (-0.251; 0.062)	.40

	Low total testosterone (<11 nmol/L)				Low free testosterone (<220 pmol/L)			
	β (95%CI)	p	β (95%CI) ^a	p	β (95%CI)	p	β (95%CI) ^a	p
14. Increase in weight	0.050 (-0.081; 0.180)	.43	-0.063 (-0.202; 0.076)	.37	0.041 (-0.113; 0.194)	.60	-0.070 (-0.231; 0.091)	.39
15. Poor concentration	0.113 (-0.046; 0.271)	.07	0.055 (-0.1; 0.21)	.49	0.072 (-0.115; 0.259)	.45	0.013 (-0.167; 0.194)	.88
16. Low self-esteem	-0.022 (-0.185; 0.142)	.07	-0.033 (-0.195; 0.129)	.69	-0.032 (-0.224; 0.159)	.74	-0.016 (-0.204; 0.172)	.86
17. Pessimism	0.271 (0.103; 0.438)	.07	0.137 (-0.027; 0.301)	.10	0.300 (0.103; 0.497)	.003	0.136 (-0.055; 0.326)	.16
18. Suicidality	0.064 (-0.094; 0.222)	.007	0.033 (-0.13; 0.196)	.69	0.054 (-0.132; 0.240)	.57	0.000 (-0.19; 0.19)	1.00
19. Low general interest	0.146 (-0.012; 0.305)	.11	0.081 (-0.078; 0.24)	.32	0.165 (-0.021; 0.352)	.08	0.100 (-0.085; 0.285)	.29
20. Energy level	0.146 (-0.012; 0.304)	.54	0.013 (-0.141; 0.168)	.87	0.143 (-0.043; 0.330)	.13	-0.007 (-0.187; 0.173)	.94
21. Low capacity for pleasure	0.148 (-0.012; 0.308)	.0003	0.072 (-0.087; 0.231)	.37	0.172 (-0.017; 0.361)	.07	0.072 (-0.113; 0.257)	.44
22. Low interest in sex	0.231 (0.065; 0.398)	.001	0.104 (-0.064; 0.271)	.22	0.341 (0.146; 0.537)	.0006	0.165 (-0.03; 0.359)	.09
23. Psychomotor retardation	0.129 (-0.028; 0.285)	.05	-0.001 (-0.158; 0.157)	.99	0.185 (0.001; 0.369)	.05	0.045 (-0.138; 0.228)	.63
24. Psychomotor agitation	0.049 (-0.110; 0.209)	.02	0.014 (-0.147; 0.174)	.87	-0.034 (-0.221; 0.153)	.72	-0.089 (-0.275; 0.097)	.35
25. Aches and pains	0.303 (0.140; 0.466)	.97	0.040 (-0.114; 0.194)	.61	0.383 (0.192; 0.575)	.0001	0.045 (-0.134; 0.224)	.62
26. Sympathetic arousal	0.271 (0.109; 0.433)	.04	0.057 (-0.095; 0.209)	.46	0.259 (0.069; 0.450)	.01	-0.010 (-0.187; 0.167)	.91
27. Panic or phobia	0.165 (0.001; 0.328)	.90	0.102 (-0.043; 0.248)	.17	0.013 (-0.180; 0.206)	.89	-0.035 (-0.204; 0.135)	.69
28. Constipation or diarrhea	0.200 (0.035; 0.364)	.003	0.082 (-0.082; 0.246)	.33	0.260 (0.066; 0.454)	.01	0.099 (-0.092; 0.289)	.31
29. Interpersonal sensitivity	-0.003 (-0.168; 0.162)	.43	-0.036 (-0.201; 0.129)	.66	-0.059 (-0.253; 0.135)	.55	-0.059 (-0.251; 0.133)	.54
30. Leaden paralysis	0.172 (0.010; 0.334)	.46	0.033 (-0.121; 0.186)	.68	0.187 (-0.004; 0.378)	.06	0.018 (-0.161; 0.197)	.84

Abbreviations: CI, Confidence interval; MDD, major depressive disorder. Figure shows unadjusted and adjusted standardized β 's with 95%CI for the association of baseline hypo or eugonadal levels of total and free testosterone with the MDD symptom severity item score over two years examined with a linear mixed model. ^aAdjusted for age, waist circumference, number of treated chronic somatic disorders, smoking status, alcohol use, physical activity, the presence of a current anxiety disorder.

Supplement 16. Non-linear associations of androgens and SHBG with individual MDD symptoms.

	Total testosterone (nmol/L)		SHBG (nmol/L)		Free testosterone (pmol/L)		5 α -DHT (nmol/L)		Androstenedione (nmol/L)		DHEAS (mmol/L)	
	β (95%CI)	p	β (95%CI)	p	β (95%CI)	p	β (95%CI)	p	β (95%CI)	p	β (95%CI)	p
Total score	0.018 (-0.013; 0.048)	.25	0.020 (-0.011; 0.051)	.20	0.001 (-0.032; 0.034)	.94	0.040 (0.009; 0.071)	.01	0.015 (-0.017; 0.047)	.35	-0.020 (-0.055; 0.014)	.24
1. Difficulty falling asleep	0.004 (-0.029; 0.037)	.82	-0.004 (-0.038; 0.029)	.80	0.017 (-0.019; 0.053)	.34	0.020 (-0.014; 0.054)	.24	0.051 (0.017; 0.085)	.004	0.020 (-0.018; 0.057)	.30
2. Difficulty staying asleep	0.024 (-0.007; 0.055)	.12	0.023 (-0.008; 0.054)	.15	-0.006 (-0.039; 0.027)	.71	0.042 (0.010; 0.073)	.009	0.003 (-0.029; 0.035)	.86	0.003 (-0.032; 0.038)	.87
3. Waking up early	0.011 (-0.021; 0.042)	.51	0.023 (-0.009; 0.055)	.15	-0.019 (-0.053; 0.015)	.27	0.036 (0.004; 0.069)	.03	0.006 (-0.027; 0.039)	.73	-0.006 (-0.042; 0.03)	.74
4. Hypersomnia	0.011 (-0.021; 0.044)	.51	-0.002 (-0.035; 0.031)	.89	0.022 (-0.013; 0.057)	.21	-0.012 (-0.046; 0.021)	.46	-0.022 (-0.056; 0.011)	.19	-0.003 (-0.04; 0.034)	.88
5. Sad mood	0.003 (-0.028; 0.033)	.86	0.001 (-0.030; 0.032)	.94	-0.004 (-0.038; 0.029)	.79	0.025 (-0.007; 0.056)	.12	0.009 (-0.023; 0.040)	.59	-0.023 (-0.057; 0.012)	.20
6. Irritability	-0.007 (-0.038; 0.024)	.64	0.007 (-0.025; 0.038)	.68	-0.017 (-0.051; 0.016)	.3	0.018 (-0.014; 0.049)	.27	0.032 (0.000; 0.063)	.05	-0.008 (-0.043; 0.027)	.66
7. Anxious or tense	-0.004 (-0.033; 0.025)	.78	-0.008 (-0.037; 0.021)	.58	-0.011 (-0.042; 0.020)	.47	0.014 (-0.015; 0.043)	.35	0.011 (-0.019; 0.041)	.47	-0.010 (-0.043; 0.022)	.53
8. Responsiveness of mood	0.018 (-0.012; 0.048)	.24	0.038 (0.008; 0.068)	.01	-0.008 (-0.040; 0.025)	.64	0.032 (0.002; 0.063)	.04	-0.001 (-0.032; 0.030)	.93	-0.032 (-0.067; 0.002)	.07
9. Diurnal mood variation	0.000 (-0.031; 0.030)	.99	-0.007 (-0.038; 0.024)	.67	0.036 (0.003; 0.069)	.03	0.008 (-0.024; 0.039)	.63	0.040 (0.008; 0.072)	.01	0.007 (-0.028; 0.042)	.69
10. Quality mood	0.009 (-0.020; 0.038)	.55	0.010 (-0.020; 0.040)	.51	0.004 (-0.028; 0.035)	.82	0.023 (-0.007; 0.053)	.13	0.020 (-0.011; 0.050)	.21	-0.035 (-0.068; -0.002)	.04
11. Decrease in appetite	0.000 (-0.029; 0.029)	.99	-0.001 (-0.030; 0.028)	.95	0.006 (-0.025; 0.037)	.7	0.014 (-0.015; 0.044)	.34	-0.011 (-0.040; 0.019)	.48	-0.030 (-0.062; 0.003)	.07
12. Increase in appetite	0.003 (-0.026; 0.032)	.86	0.026 (-0.004; 0.055)	.08	-0.016 (-0.047; 0.015)	.31	-0.008 (-0.038; 0.022)	.6	-0.021 (-0.051; 0.009)	.16	-0.015 (-0.047; 0.018)	.38
13. Decrease in weight	-0.002 (-0.028; 0.025)	.91	0.002 (-0.025; 0.029)	.88	-0.015 (-0.043; 0.013)	.3	-0.014 (-0.04; 0.013)	.32	-0.021 (-0.048; 0.006)	.13	-0.018 (-0.048; 0.012)	.24
14. Increase in weight	-0.003 (-0.029; 0.024)	.85	0.017 (-0.010; 0.044)	.21	0.006 (-0.023; 0.035)	.68	0.012 (-0.015; 0.040)	.38	-0.011 (-0.039; 0.017)	.43	0.004 (-0.027; 0.035)	.81
15. Poor concentration	0.010 (-0.020; 0.040)	.51	0.023 (-0.008; 0.053)	.14	-0.002 (-0.035; 0.030)	.9	0.026 (-0.004; 0.057)	.09	0.016 (-0.015; 0.047)	.31	-0.014 (-0.048; 0.02)	.43
16. Low self-esteem	0.002 (-0.030; 0.033)	.91	0.023 (-0.009; 0.055)	.15	-0.015 (-0.049; 0.019)	.39	0.023 (-0.009; 0.055)	.15	-0.021 (-0.054; 0.012)	.21	-0.034 (-0.069; 0.002)	.06
17. Pessimism	0.032 (0.000; 0.064)	.047	0.026 (-0.006; 0.059)	.11	0.012 (-0.022; 0.047)	.49	0.040 (0.007; 0.072)	.02	0.014 (-0.019; 0.047)	.42	-0.010 (-0.046; 0.027)	.60
18. Suicidality	0.009 (-0.022; 0.041)	.56	0.008 (-0.024; 0.040)	.61	0.001 (-0.033; 0.035)	.96	0.006 (-0.027; 0.038)	.73	-0.005 (-0.038; 0.028)	.76	-0.031 (-0.066; 0.005)	.09

	Total testosterone (nmol/L)		SHBG (nmol/L)		Free testosterone (pmol/L)		5 α -DHT (nmol/L)		Androstenedione (nmol/L)		DHEAS (nmol/L)	
	β (95%CI)	p	β (95%CI)	p	β (95%CI)	p	β (95%CI)	p	β (95%CI)	p	β (95%CI)	p
19. Low general interest	0.010 (-0.021; 0.041)	.52	0.024 (-0.007; 0.056)	.13	-0.004 (-0.038; 0.029)	.79	0.030 (-0.002; 0.062)	.06	-0.004 (-0.036; 0.028)	.79	-0.009 (-0.044; 0.026)	.63
20. Energy level	0.015 (-0.016; 0.045)	.34	0.008 (-0.022; 0.039)	.58	-0.013 (-0.045; 0.019)	.42	0.031 (0.000; 0.061)	.047	-0.002 (-0.033; 0.029)	.90	-0.026 (-0.06; 0.008)	.13
21. Low capacity for pleasure	0.019 (-0.012; 0.050)	.23	0.029 (-0.002; 0.061)	.07	0.004 (-0.029; 0.037)	.82	0.036 (0.004; 0.068)	.02	0.003 (-0.029; 0.035)	.86	-0.021 (-0.056; 0.014)	.25
22. Low interest in sex	0.033 (0.001; 0.066)	.04	0.023 (-0.010; 0.056)	.16	0.026 (-0.009; 0.061)	.14	0.03 (-0.003; 0.063)	.07	-0.008 (-0.042; 0.026)	.64	-0.011 (-0.048; 0.025)	.54
23. Psychomotor retardation	0.034 (0.003; 0.064)	.03	0.029 (-0.001; 0.060)	.06	0.005 (-0.028; 0.038)	.77	0.052 (0.021; 0.083)	.0009	0.004 (-0.028; 0.036)	.81	-0.018 (-0.053; 0.016)	.30
24. Psychomotor agitation	-0.002 (-0.033; 0.029)	.91	0.012 (-0.019; 0.044)	.44	-0.003 (-0.037; 0.03)	.84	0.024 (-0.007; 0.056)	.13	0.023 (-0.009; 0.056)	.15	-0.002 (-0.037; 0.034)	.93
25. Aches and pains	0.008 (-0.022; 0.038)	.61	0.013 (-0.017; 0.043)	.40	-0.010 (-0.042; 0.023)	.56	0.027 (-0.003; 0.058)	.08	0.025 (-0.006; 0.056)	.11	0.022 (-0.012; 0.056)	.20
26. Sympathetic arousal	0.006 (-0.024; 0.036)	.69	0.002 (-0.028; 0.032)	.91	-0.008 (-0.040; 0.024)	.61	0.030 (0.000; 0.060)	.047	0.008 (-0.022; 0.039)	.60	0.011 (-0.023; 0.045)	.52
27. Panic or phobia	0.010 (-0.019; 0.038)	.50	0.012 (-0.016; 0.041)	.40	0.011 (-0.019; 0.042)	.46	0.008 (-0.021; 0.037)	.58	0.028 (-0.001; 0.057)	.06	-0.021 (-0.053; 0.011)	.20
28. Constipation or diarrhea	0.013 (-0.019; 0.045)	.41	-0.017 (-0.049; 0.015)	.30	0.025 (-0.010; 0.059)	.16	0.029 (-0.003; 0.062)	.08	0.016 (-0.017; 0.049)	.33	0.01 (-0.026; 0.046)	.58
29. Interpersonal sensitivity	0.001 (-0.031; 0.033)	.96	0.011 (-0.021; 0.044)	.50	-0.009 (-0.044; 0.025)	.59	0.018 (-0.014; 0.051)	.27	0.016 (-0.017; 0.05)	.34	-0.007 (-0.043; 0.030)	.72
30. Leaden paralysis	-0.006 (-0.036; 0.024)	.68	-0.001 (-0.031; 0.030)	.98	-0.021 (-0.054; 0.011)	.19	0.033 (0.003; 0.064)	.03	0.018 (-0.013; 0.049)	.26	-0.019 (-0.052; 0.015)	.28

Abbreviations: CI, Confidence interval; DHEAS, dehydroepiandrosterone-sulphate; 5 α -DHT, 5 α -dihydrotestosterone; MDD, major depressive disorder; SHBG, sex hormone binding globulin. Table shows and adjusted standardized non-linear β 's with 95%CI for the association of baseline androgens and SHBG levels with the MDD symptom severity item score over two years examined with a linear mixed model. The p-value is based on the comparison of the model only including a linear term and the model including both a linear and quadratic term using two-way analyses of covariance analysis. The p-value refers to the significance of the added quadratic term. All models were adjusted for age, waist circumference, number of treated chronic somatic disorders, smoking status, alcohol use, physical activity, the presence of a current anxiety disorder.

Chapter 6

Supplement 16. Characteristics at baseline of 2861 included and 120 excluded participants.

	Included n = 2861	Excluded n = 120	p
Female, no. (%)	1896 (65.9)	83 (69.2)	.56
Age, mean (SD), y	42.0 ± 13.1	39.7 ± 13.5	.06
Education, mean (SD), y	12.2 ± 3.1	11.9 ± 3.0	.31
Married or with partner, no. (%)	1981 (69.2)	85 (70.8)	.76
Body mass index, mean (SD), kg/m²	25.6 ± 5.0	25.9 ± 5.5	.47
Smoking, no. (%)	1105 (38.6)	44 (36.7)	.84
Alcohol intake >1 unit a day, no. (%)	569 (19.8)	15 (12.5)	.046
Treated chronic diseases, geo. mean (95%CI)	0.5 (0.0-1.5)	0.6 (0.0-1.7)	.33
Oral contraceptive use, no. (%)	508 (17.7)	18 (15.0)	.37
Antidepressant use, no. (%)	736 (25.7)	32 (26.7)	.83
MDD or anxiety disorder past month, no. (%)	1342 (46.9)	56 (46.7)	1.00
Suicide ideation in preceding week, no. (%)	340 (11.9)	11 (9.2)	.47
Suicide attempt in past, no. (%)	328 (18.8)	19 (15.8)	.04
Reactive aggression, geo. mean (95%CI)	3.5 (2.5-4.5)	3.1 (2.0-4.2)	.19

Abbreviations: CI, confidence interval; MDD, major depressive disorder; SD, standard deviation. Data is based on t-tests for independent samples or χ^2 -test.

Supplement 17. Associations between baseline plasma androgen levels and concurrent reactive aggression in men and women.

	Reactive aggression			
	β (95%CI)	p	β (95%CI)	p ^a
<i>Men</i>				
Total testosterone	-0.195 (-0.500-0.110)	.21	-0.248 (-0.579-0.084)	.14
5α-DHT	-0.136 (-0.444-0.172)	.39	-0.089 (-0.415-0.238)	.60
Androstenedione	0.259 (-0.050-0.568)	.10	0.099 (-0.234-0.431)	.56
DHEAS	0.214 (-0.097-0.525)	.18	0.067 (-0.309-0.443)	.73
<i>Women</i>				
Total testosterone	0.128 (-0.068-0.324)	.20	-0.086 (-0.292-0.119)	.41
5α-DHT	0.212 (0.015-0.409)	.04	-0.077 (-0.296-0.143)	.49
Androstenedione	0.362 (0.166-0.559)	.0003	0.012 (-0.219-0.243)	.92
DHEAS	0.149 (-0.044-0.342)	.13	-0.282 (-0.510- -0.055)	.02

Abbreviations: 5 α -DHT, 5 α -dihydrotestosterone; CI, confidence interval; DHEAS, dehydroepiandrosterone-sulphate; β , standardized beta-coefficient. Data are standardized β 's with 95%CI's based on linear regression analysis. To convert nmol/L to ng/dl for total testosterone, 5 α -DHT, and androstenedione multiply by 28.84, 29.07, and 28.64, respectively. To convert pmol/L to pg/ml for free testosterone divide by 3.47. To convert μ mol/L to μ g/dl for DHEAS, multiply by 36.85.^a Adjusted for age, education, married or with partner, body mass index, smoking, alcohol use, and number of treated chronic diseases, (and in women also for oral contraceptive pill use).

Chapter 9

Supplement 18. Sociodemographic and clinical characteristics of included and excluded girls per wave.

	16 years of age				19 years of age				22 years of age				25 years of age			
	Included	Excluded	n = 107	p	Included	Excluded	n = 903	p	Included	Excluded	n = 826	p	Included	Excluded	n = 743	p
Age , mean (SD)	16.2 (0.7)	16.5 (0.9)		.005	19.0 (0.6)	19.5 (0.6)		<.001	22.2 (0.7)	22.5 (0.6)		<.001	25.6 (0.6)	25.8 (0.6)		.04
Dutch ethnicity , no. (%)	745 (88.2)	93 (86.9)		.75	799 (88.5)	70 (87.5)		.72	736 (89.1)	93 (83.8)		.11	669 (90.0)	44 (89.8)		1.00
Socioeconomic status , mean (SD) ^a	0.06 (0.77)	-0.16 (0.72)		.005	0.05 (0.76)	-0.18 (0.76)		.008	0.09 (0.76)	-0.19 (0.77)		<.001	0.11 (0.76)	-0.01 (0.74)		.26
Body mass index , mean (SD)	21.7 (3.1)	21.8 (3.8)		.82	23.1 (3.9)	23.9 (5.9)		.26	23.7 (4.4)	23.4 (4.9)		.66	24.1 (4.7)	24.0 (5.2)		.86
Virgin , no. (%)	498 (59.6)	22/43 (51.2)		.34	242 (26.9)	2/8 (25.0)		1.00	98/825 (11.9)	1/19 (5.3)		.72	42/697 (5.7)	0/24 (0.0)		.64
Age menarche , mean (SD)	12.9 (1.2)	12.9 (1.2)		.98	12.8 (1.2)	12.6 (1.5)		.22	12.8 (1.2)	12.8 (1.3)		.82	12.8 (1.1)	12.9 (1.2)		.75
Antidepressant use , no. (%)	1/844 (0.1)	1/107 (0.9)		.21	10 (1.1)	0 (0.0)		1.00	22 (2.7)	1 (0.9)		.51	31 (4.2)	1 (2.0)		.72
Depressive symptoms , mean (SD)	0.36 (0.30)	0.31 (0.22)		.33	0.36 (0.33)	0.22 (0.19)		.001	0.38 (0.33)	0.25 (0.31)		.09	0.43 (0.36)	0.39 (0.33)		.44

Abbreviations: SD, standard deviation. Data are based on two-tailed χ^2 tests for dichotomous variables or t tests for continuous variables. All variables were assessed at every wave, except for socioeconomic status, ethnicity, and menarche; the first two were determined at study entry (T1) and the latter at age 16 years (T3). ^a Continuous scale based on the educational level, job, and income of the parents.

Supplement 19. Associations between OC use and depressive symptom scores during 9 years of follow-up.

	n users vs. nonusers	Depressive symptom score	
		B (95%CI)	p
<i>Users vs. nonusers</i>			
OC use	1806 vs. 1507	0.007 (-0.013; 0.026)	.50
Adjusted OC use^a	1782 vs. 1483	0.006 (-0.013; 0.025)	.52
16 year old users	300 vs. 530	0.075 (0.031; 0.120)	<.001
19 year old users	606 vs. 285	0.013 (-0.033; 0.060)	.57
22 year old users	534 vs. 278	-0.013 (-0.061; 0.034)	.59
25 year old users	342 vs. 390	-0.032 (-0.083; 0.020)	.23
<i>First-time users vs. never users</i>			
OC use	828 vs. 1071	0.026 (0.002; 0.050)	.03
Adjusted OC use^a	818 vs. 1054	0.021 (-0.005; 0.046)	.11

Abbreviations: B, beta-coefficient; CI, confidence interval; OC, oral contraceptive pill. Estimates were obtained through multilevel analysis (i.e., mixed models). ^aAdjusted for age, ethnicity, and socioeconomic status.

Supplement 20. Post Hoc comparison of characteristics of OC users and nonusers at age 16 years and at age 13 years when OC use was unlikely.

	13 years of age			16 years of age		
	No OC use age 16 n = 530	OC use age 16 n = 306	p	No OC use n = 536	OC use n = 309	p
Dutch ethnicity , no. (%)	460 (85.8)	285 (92.2)	.006	460 (85.8)	285 (92.2)	.006
Socioeconomic status^a , mean (SD)	0.17 (0.78)	-0.15 (0.71)	<.001	0.17 (0.78)	-0.15 (0.71)	<.001
Virgin , no. (%)	515 (98.5)	276 (91.1)	<.001	424 (79.5)	74 (24.4)	<.001
Family functioning^b , mean (SD)	1.62 (0.39)	1.65 (0.40)	.31	1.65 (0.39)	1.67 (0.41)	.10
Stressful experiences^c , median (IQR)	2.00 (1.00-4.00)	3.00 (1.00-5.75)	.003	3.00 (2.00-6.00)	5.00 (3.00-7.00)	<.001
Depressive symptoms , mean (SD)	0.30 (0.28)	0.37 (0.30)	.001	0.33 (0.30)	0.40 (0.30)	.001
Body mass index , mean (SD)	19.1 (3.3)	19.5 (3.2)	.12	21.5 (3.12)	22.0 (3.06)	.01
Age at menarche , mean (SD)	12.3 (0.90)	12.1 (0.96)	.11	13.0 (1.17)	12.6 (1.15)	<.001
Acne , no. (%)	150 (30.4)	127 (46.9)	<.001	203 (43.2)	107 (43.9)	.87

Abbreviations: IQR, interquartile range; OC, oral contraceptive pill; SD, standard deviation. All variables were assessed at both age 13 years (T2) and age 16 years (T3). Socioeconomic status and ethnicity are time-invariant variables (determined at study entry [T1]). Differences in these variables are therefore copied from *Table 9.1*. Data are based on two-tailed χ^2 tests for dichotomous variables or t tests or Mann Whitney U tests for continuous variables. ^a Continuous scale based on the educational level, job, and income of the parents. ^b Family functioning was assessed by a modified version of the General Functioning Scale of the McMaster Family Assessment Device.⁴⁸⁵ A higher scores indicate worse levels of family functioning. ^c Stressful experiences were assessed with the question "How many stressful experiences have you gone through in the past two years?". The answer was rated on a continuous scale between 0 and 10, where 0 indicated no stressful experiences and 10 indicated many stressful experiences.

Chapter 10

Supplement 21. Details of the 818 analytic specifications - Included outcomes, fixed covariates and time-variant/invariant covariates in all women and the subgroup of women with no history of MDD.

Decision level		Details of specification		n of specifications per sample and outcome				Total n
Sample								
Outcome								
Covariates								
Fixed								
Crude model								
One by one								
All combined								
Time-variant: baseline								
One by one								
Unique combinations of two								
Unique interactions of two								
Time-variant: age 16								
One by one								
Unique combinations of two								
Unique interactions of two								
Time-variant: age 19								
One by one								
Unique combinations of two								
Unique interactions of two								
Time-invariant								
One by one								
Unique combinations of two								
Unique interactions of two								
Unique combinations of two with time-variant covariates								
Unique interactions of two with time-variant covariates								
Total		205	205	204	204	205	204	818

Abbreviations: BMI, body mass index; MDD, major depressive disorder; OC, oral contraceptive pill.

Of the fixed covariates, ethnicity and socioeconomic status were determined at study entry (around age 11). Socioeconomic status was assessed on a continuous scale based on the education, job and income of the parents, ranging from -1.94 to 1.73 with higher values indicating a better social and economic position. Due to the small proportion of women having a different ancestry than Dutch (9.3%), Ethnicity was dichotomized as having a Dutch or non-Dutch background. Age was determined at baseline (around age 13). History of major depressive disorder (MDD) before age 19 was determined using Composite International Diagnostic Interview (CIDI) data, and was defined as having at least one episode of MDD by the time of assessment at age 19. In case such data was missing for the CIDI ($n = 41$), the Lifetime Depression Assessment Self-Report (LIDAS) (history of MDD data combined with age of onset data) at age 25 was used to determine whether an episode of MDD had occurred prior to or at age 19.

Body mass index, virginity, and smoking were measured at baseline, at age 16, and at age 19; stressful experiences, acne, and menstruation-related pain were assessed only at baseline and age 16. Virginity was ascertained with the question "Have you ever had sexual intercourse?", Smoking was assessed with the question "Do you ever smoke cigarettes?". The presence of acne was determined with the question "Have you had acne during the last year?", and the presence of menstruation-related pain with "Have you had menstrual pain during the last year?". These four variables had dichotomous response choices. Stress was assessed using a continuous measure ("How many stressful experiences have you gone through in the past two years?" 0 = no stressful experiences and 10 = many stressful experiences).

The time-invariant model-dependent covariates were ascertained at baseline (depressive symptoms), age 16 (age at menarche), age 19 (sexual orientation), or age 25 (current OC use). Depressive symptom scores were assessed with the Affective Problems scale of the Youth Self-Report (YSR).³⁶⁶ This scale matches closely the structure and item content of Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV depressive disorders (MDD and dysthymia), and reflects the depressive symptoms during the preceding six months. Age at menarche was determined at baseline and completed with data from age 16 in case data at baseline was missing. Sexual orientation was categorized as 'Heterosexual' or 'Homo/bisexual'.

Supplement 22. Lifetime presence of psychiatric comorbidities at age 25 depending on history of oral contraceptive pill use during adolescence of 725 women.

	Adolescent nonuser n = 191	Adolescent OC user n = 534	p
Bipolar disorder, no. (%)	4 (2.1)	4 (0.8)	0.22
Schizophrenia or psychosis, no. (%)	2 (1.0)	2 (0.4)	0.28
Eating disorder, no. (%)	5 (2.6)	17 (3.2)	0.88
Anxiety disorder, no. (%)	7 (3.7)	41 (7.7)	0.08
Panic disorder, no. (%)	2 (1.0)	24 (4.5)	0.05
Obsessive compulsive disorder, no. (%)	2 (1.0)	11 (2.1)	0.53
Posttraumatic stress disorder, no. (%)	7 (3.7)	31 (5.8)	0.34
Attention Deficit Disorder/Attention Deficit Hyperactivity Disorder, no. (%)	8 (4.2)	26 (4.9)	0.85
Personality disorder, no. (%)	4 (2.1)	19 (3.6)	0.45
Alcohol addiction, no. (%)	0 (0.0)	2 (0.4)	1.00
Drug addiction, no. (%)	0 (0.0)	5 (0.9)	0.33

Abbreviation: OC, oral contraceptive pill. Data are based on two-tailed χ^2 tests. ^a Lifetime prevalences were assessed with the question "Could please mark which of the following diseases you were ever diagnosed with by a doctor or aid worker?".

Chapter 11

Supplement 23. Search string of Chapter 11 and 12.

Search conducted in PubMed. The search includes searching terms for the intervention (hormonal contraceptive), the outcome (depressive symptoms), the design (randomized clinical trial), and the participants (female sex).

Please mind that “males” were excluded by the NOT command, but “females” were not because of the double negative use of the NOT command. This double negative was necessary so that articles which discuss both females and males are not excluded. We can translate this part of the search strategy as “exclude papers that mention men and do not mention women” i.e. NOT (men NOT women) [note the brackets]. A similar strategy was also applied to limit the search to adolescents and/or adults (not children) and humans (not animals).

((“synthetic estrogen”[tw] OR “synthetic estrogens”[tw] OR “synthetic oestrogen”[tw] OR “synthetic oestrogens”[tw] OR “Estradiol Congeners”[Mesh] OR “2- Methoxyestradiol”[tw] OR “Alkylated Estrogenic Steroids”[tw] OR “Catechol Estrogens”[tw] OR “Conjugated Estrogens”[tw] OR “Equilenin”[tw] OR “Equilin”[tw] OR “Esterified Estrogens”[tw] OR “Estetrol”[tw] OR “Estradiol”[tw] OR “Estriol”[tw] OR “Estrone”[tw] OR “Ethinyl Estradiol”[tw] OR “Ethinyl Estradiol-Norgestrel Combination”[tw] OR “Fulvestrant”[tw] OR “Hydroxyestrones”[tw] OR “Mestranol”[tw] OR “Quinestrol”[tw] OR “Catechol Estrogen”[tw] OR “Conjugated Estrogen”[tw] OR “Esterified Estrogen”[tw] OR “Hydroxyestrone”[tw] OR “Catechol oestrogen”[tw] OR “Conjugated oestrogen”[tw] OR “Esterified oestrogen”[tw] OR “Catechol oestrogen”[tw] OR “Conjugated oestrogen”[tw] OR “Esterified oestrogen”[tw] OR “Estrogen Replacement Therapy”[Mesh] OR “Estrogen Replacement”[tw] OR “Oestrogen Replacement”[tw] OR “Hormone Replacement”[tw] OR “Progestin Replacement”[tw] OR “Contraceptives, Oral, Hormonal”[Mesh] OR “Contraceptives, Oral, Hormonal”[Pharmacological Action] OR “Contraceptive Agents, Hormonal”[Mesh] OR “Contraceptive Agents, Hormonal”[Pharmacological Action] OR “Contraceptive Agents, Female” [mesh] OR “Contraceptive Agents, Female”[Pharmacological Action] OR “Algestone Acetophenide”[mesh] OR “Algestone Acetophenide”[mesh] OR “Algestone Acetophenide”[tw] OR “Algestone”[mesh] OR “Algestone”[tw] OR “anordiol”[Supplementary Concept] OR “anordiol”[tw] OR “anordrin”[Supplementary Concept] OR “anordrin”[tw] OR “Bisecurin I”[Supplementary Concept] OR “Bisecurin I”[tw] OR “Centchroman”[mesh] OR “Centchroman”[tw] OR “Chlormadinone Acetate”[mesh] OR “Chlormadinone Acetate”[tw] OR “Chlormadinone Acetate”[tw] OR “Cloprostenol”[mesh] OR “Cloprostenol”[tw] OR “CycloProvera”[Supplementary Concept] OR “CycloProvera”[tw] OR “Cyklabil”[Supplementary Concept] OR “Cyklabil”[tw] OR “Deposiston”[Supplementary Concept] OR “Deposiston”[tw] OR “Desogestrel”[mesh] OR “Desogestrel”[tw] OR “diallyl disulfide”[supplementary concept] OR “diallyl disulfide”[tw] OR “diallyl trisulfide”[supplementary concept] OR “diallyl trisulfide”[tw] OR “dienogest”[Supplementary Concept] OR “dienogest”[tw] OR “Dimethisterone”[mesh] OR “Dimethisterone”[tw] OR “estradiol 17 beta-cypionate”[Supplementary Concept] OR “estradiol 17 beta-cypionate”[tw] OR “estradiol 3-benzoate”[Supplementary Concept] OR “estradiol 3-benzoate”[tw] OR “estropipate”[Supplementary Concept] OR “estropipate”[tw] OR “Ethinyl Estradiol”[mesh] OR “Ethinyl Estradiol”[tw] OR “ethinyl estradiol, levonorgestrel drug combination”[Supplementary Concept] OR “ethinyl estradiol, levonorgestrel drug combination”[tw] OR “Ethinyl Estradiol-Norgestrel Combination”[mesh] OR “Ethinyl Estradiol-Norgestrel Combination”[tw] OR “Ethynodiol Diacetate”[mesh] OR “Ethynodiol Diacetate”[tw] OR “etonogestrel”[Supplementary

Concept] OR "etonogestrel"[tw] OR "Femovan"[supplementary concept] OR "Femovan"[tw] OR "fluprostenol"[Supplementary Concept] OR "fluprostenol"[tw] OR "Gestodene"[Supplementary Concept] OR "Gestodene"[tw] OR "Gestrinone"[mesh] OR "Gestrinone"[tw] OR "gossypol acetic acid"[supplementary concept] OR "gossypol acetic acid"[tw] OR "Infecundin"[Supplementary Concept] OR "Infecundin"[tw] OR "levonorgestrel butanoate"[Supplementary Concept] OR "levonorgestrel butanoate"[tw] OR "Levonorgestrel"[mesh] OR "Levonorgestrel"[tw] OR "Lynestrenol"[mesh] OR "Lynestrenol"[tw] OR "Medroxyprogesterone Acetate"[mesh] OR "Medroxyprogesterone Acetate"[tw] OR "Medroxyprogesterone"[mesh] OR "Medroxyprogesterone"[tw] OR "Megestrol"[mesh] OR "Megestrol"[tw] OR "Mestranol"[mesh] OR "Mestranol"[tw] OR "Mifepristone"[mesh] OR "Mifepristone"[tw] OR "neem oil"[supplementary concept] OR "neem oil"[tw] OR "Nonidet P-40"[supplementary concept] OR "Nonidet P-40"[tw] OR "Non- Ovlon"[Supplementary Concept] OR "Non-Ovlon"[tw] OR "Nonoxynol"[mesh] OR "Nonoxynol"[tw] OR "Noracycline"[Supplementary Concept] OR "Noracycline"[tw] OR "norelgestromin"[Supplementary Concept] OR "norelgestromin"[tw] OR "Norethindrone Acetate"[mesh] OR "Norethindrone Acetate"[tw] OR "norethindrone acetate, ethinyl estradiol, ferrous fumarate drug combination"[Supplementary Concept] OR "norethindrone acetate, ethinyl estradiol, ferrous fumarate drug combination"[tw] OR "norethindrone enanthate"[Supplementary Concept] OR "norethindrone enanthate"[tw] OR "Norethindrone"[mesh] OR "Norethindrone"[tw] OR "Norethynodrel"[mesh] OR "Norethynodrel"[tw] OR "norgestimate"[Supplementary Concept] OR "norgestimate"[tw] OR "norgestimate, ethinyl estradiol drug combination"[Supplementary Concept] OR "norgestimate, ethinyl estradiol drug combination"[tw] OR "Norgestrel"[mesh] OR "Norgestrel"[tw] OR "Norgestrienone"[mesh] OR "Norgestrienone"[tw] OR "Norinyl"[Supplementary Concept] OR "Norinyl"[tw] OR "NuvaRing"[Supplementary Concept] OR "NuvaRing"[tw] OR "Octoxynol"[mesh] OR "Octoxynol"[tw] OR "Ortho Evra"[Supplementary Concept] OR "Ortho Evra"[tw] OR "Oviol"[Supplementary Concept] OR "Oviol"[tw] OR "SC-11800 EE"[Supplementary Concept] OR "SC-11800 EE"[tw] OR "ST 1435"[Supplementary Concept] OR "ST 1435"[tw] OR "sulprostone"[Supplementary Concept] OR "sulprostone"[tw] OR "trestolone"[Supplementary Concept] OR "trestolone"[tw] OR "Triptorelin Pamoate"[mesh] OR "Triptorelin Pamoate"[tw] OR "Trisequens"[Supplementary Concept] OR "Trisequens"[tw] OR "ulipristal acetate"[Supplementary Concept] OR "ulipristal acetate"[tw] OR "Estrogens"[Mesh] OR "Estrogens"[Pharmacological Action] OR "estrogen"[tw] OR "estrogens"[tw] OR estrogen*[tw] OR "oestrogen"[tw] OR "oestrogens"[tw] OR oestrogen*[tw] OR "2,3-bis(3'-hydroxybenzyl)butane-1,4- diol"[Supplementary Concept] OR "2,3-bis(3'-hydroxybenzyl)butyrolactone"[Supplementary Concept] OR "4- octylphenol"[Supplementary Concept] OR "8-prenylnaringenin"[Supplementary Concept] OR "biochanin A"[Supplementary Concept] OR "bisphenol A"[Supplementary Concept] OR "Chlorotrianisene"[mesh] OR "Coumestrol"[mesh] OR "daidzein"[Supplementary Concept] OR "Dienestrol"[mesh] OR "Diethylstilbestrol"[mesh] OR "diethylstilbestrol dipropionate"[Supplementary Concept] OR "Epimestrol"[mesh] OR "Equol"[mesh] OR "Estradiol"[mesh] OR "estradiol enanthate"[Supplementary Concept] OR "Estrogenic Steroids, Alkylated"[mesh] OR "Estrogens, Conjugated (USP)"[mesh] OR "estrogens, conjugated synthetic A"[Supplementary Concept] OR "estrogens, conjugated synthetic B"[Supplementary Concept] OR "Estrogens, Esterified (USP)"[mesh] OR "Estrone"[mesh] OR "Ethinyl Estradiol"[mesh] OR "formononetin"[Supplementary Concept] OR "Genistein"[mesh] OR "glycitein"[Supplementary Concept] OR "glycitin"[Supplementary Concept] OR "Hexestrol"[mesh] OR "Infecundin"[Supplementary Concept] OR "Mestranol"[mesh] OR "nylestrol"[Supplementary Concept] OR "o,p'-DDT"[Supplementary Concept] OR "O-desmethylanogolensin"[Supplementary Concept] OR "polyestradiol phosphate"[Supplementary Concept] OR "Quinestrol"[mesh] OR "secoisolariciresinol"[Supplementary Concept] OR "Zearalenone"[mesh] OR "Zeranol"[mesh] OR "2,3-bis(3'-hydroxybenzyl)butane-1,4-diol"[tw] OR "2,3-bis(3'- hydroxybenzyl) butyrolactone"[tw] OR "4-octylphenol"[tw] OR "8-prenylnaringenin"[tw] OR "biochanin A"[tw] OR "bisphenol A"[tw] OR "Chlorotrianisene"[tw] OR "Coumestrol"[tw] OR "daidzein"[tw] OR "Dienestrol"[tw] OR "Diethylstilbestrol"[tw] OR "diethylstilbestrol dipropionate"[tw] OR "Epimestrol"[tw] OR "Equol"[tw] OR "Estradiol"[tw] OR "estradiol enanthate"[tw] OR "Estrogenic Steroids, Alkylated"[tw] OR "Estrogens, Conjugated (USP)"[tw] OR "estrogens, conjugated synthetic A"[tw] OR "estrogens, conjugated synthetic B"[tw] OR "Estrogens,

Esterified (USP)[tw] OR "Estrone"[tw] OR "Ethinyl Estradiol"[tw] OR "formononetin"[tw] OR "Genistein"[tw] OR "glycitein"[tw] OR "glycitin"[tw] OR "Hexestrol"[tw] OR "Infecundin"[tw] OR "Mestranol"[tw] OR "nylestriol"[tw] OR "o,p'-DDT"[tw] OR "O-desmethylangolensin"[tw] OR "polyestradiol phosphate"[tw] OR "Quinestrol"[tw] OR "secoisolariciresinol"[tw] OR "Zearalenone"[tw] OR "Zeranol"[tw]) AND ("Depression"[Mesh] OR "Depressive Disorder"[Mesh] OR "depression"[tw] OR "depressive"[tw] OR "Mood Disorders"[Mesh] OR "Affect"[Mesh] OR "mood"[tw] OR "Cyclothymic Disorder"[tw] OR "Dysthymic Disorder"[tw] OR "Premenstrual Dysphoric Disorder"[tw] OR "Seasonal Affective Disorder"[tw] OR "Cyclothymic Disorders"[tw] OR "Dysthymic Disorders"[tw] OR "Premenstrual Dysphoric Disorders"[tw] OR "Seasonal Affective Disorders"[tw] OR "Premenstrual Syndrome"[mesh] OR "premenstrual syndrome"[tw] OR premenstrual symptom*[tw] OR menstrual symptom*[tw] OR "premenstrual mood"[tw] OR "intermenstrual mood"[tw] OR "menstrual mood"[tw]) NOT (("Male"[mesh] OR "Men"[mesh] OR "male"[ti] OR "males"[ti] OR "men"[ti]) NOT ("Female"[mesh] OR "Women"[mesh] OR "Female"[ti] OR "Women"[ti] OR "girl"[ti] OR "girls"[ti] OR "girlhood"[ti])) NOT (("Infant"[mesh] OR "Child"[mesh]) NOT ("Adolescent"[mesh] OR "Adult"[mesh])) NOT ("Animals"[mesh] NOT "Humans"[mesh]) AND english[la] AND ("clinical trial"[pt] OR "clinical trial"[tiab] OR "clinical trials as topic"[mesh] OR "clinical trials"[tiab] OR "control groups"[mesh] OR "control group"[tiab] OR "control groups"[tiab] OR "controlled clinical trial"[pt] OR "controlled clinical trials as topic"[mesh] OR "cross-over studies"[mesh] OR "cross over study"[tiab] OR "cross over studies"[tiab] OR "double-blind method"[mesh] OR "double blind"[tiab] OR "evaluation studies as topic"[mesh] OR "follow-up studies"[mesh] OR "follow up study"[tiab] OR "follow up studies"[tiab] OR "placebos"[mesh] OR placebo*[tiab] OR placebo*[tiab] OR "pragmatic clinical trial"[pt] OR "prospective studies"[mesh] OR "prospective study"[tiab] OR "prospective studies"[tiab] OR "RaCT"[tiab] OR "RaCTs"[tiab] OR "random allocation"[mesh] OR "randomised"[tiab] OR "randomized controlled trial"[pt] OR "randomized controlled trials as topic"[mesh] OR "randomized"[tiab] OR random*[tiab] OR "RCT"[tiab] OR "RCTs"[tiab] OR "Research Design"[MeSH:noexp] OR "Research design"[tiab] OR "Research designs"[tiab] OR "single blind"[tiab] OR "single-blind method"[mesh] OR ((single*[tiab] OR double*[tiab] OR triple*[tiab]) AND (blind*[tiab] OR mask*[tiab])) OR volunteer*[tiab] OR "trial"[ti] OR "trials"[ti]))

Search conducted in Pubmed to identify related reviews.

("synthetic estrogen"[tw] OR "synthetic estrogens"[tw] OR "synthetic oestrogen"[tw] OR "synthetic oestrogens"[tw] OR "Estradiol Congeners"[Mesh] OR "2-Methoxyestradiol"[tw] OR "Alkylated Estrogenic Steroids"[tw] OR "Catechol Estrogens"[tw] OR "Conjugated Estrogens"[tw] OR "Equilenin"[tw] OR "Equilin"[tw] OR "Esterified Estrogens"[tw] OR "Estetrol"[tw] OR "Estradiol"[tw] OR "Estriol"[tw] OR "Estrone"[tw] OR "Ethinyl Estradiol"[tw] OR "Ethinyl Estradiol-Norgestrel Combination"[tw] OR "Fulvestrant"[tw] OR "Hydroxyestrones"[tw] OR "Mestranol"[tw] OR "Quinestrol"[tw] OR "Catechol Estrogen"[tw] OR "Conjugated Estrogen"[tw] OR "Esterified Estrogen"[tw] OR "Hydroxyestrone"[tw] OR "Catechol oestrogen"[tw] OR "Conjugated oestrogen"[tw] OR "Esterified oestrogen"[tw] OR "Catechol oestrogen"[tw] OR "Conjugated oestrogen"[tw] OR "Esterified oestrogen"[tw] OR "Estrogen Replacement Therapy"[Mesh] OR "Estrogen Replacement"[tw] OR "Oestrogen Replacement"[tw] OR "Hormone Replacement"[tw] OR "Progestin Replacement"[tw] OR "Contraceptives, Oral, Hormonal"[Mesh] OR "Contraceptives, Oral, Hormonal"[Pharmacological Action] OR "Contraceptive Agents, Hormonal"[Mesh] OR "Contraceptive Agents, Hormonal"[Pharmacological Action] OR "Contraceptive Agents, Female"[mesh] OR "Contraceptive Agents, Female"[Pharmacological Action] OR "Algestone Acetophenide"[mesh] OR "Algestone Acetophenide"[tw] OR "Algestone"[mesh] OR "Algestone"[tw] OR "anordiol"[Supplementary Concept] OR "anordiol"[tw] OR "anordrin"[Supplementary Concept] OR "anordrin"[tw] OR "Bisecurin I"[Supplementary Concept] OR "Bisecurin I"[tw] OR "Centchroman"[mesh] OR "Centchroman"[tw] OR "Chlormadinone Acetate"[mesh] OR "Chlormadinone Acetate"[tw] OR "Cloprostenol"[mesh] OR "Cloprostenol"[tw] OR "CycloProvera"[Supplementary Concept] OR "CycloProvera"[tw] OR "Cyklabil"[Supplementary Concept] OR "Cyklabil"[tw] OR "Deposiston"[Supplementary Concept] OR "Deposiston"[tw]

OR "Desogestrel"[mesh] OR "Desogestrel"[tw] OR "diallyl disulfide"[supplementary concept] OR "diallyl disulfide"[tw] OR "diallyl trisulfide"[supplementary concept] OR "diallyl trisulfide"[tw] OR "dienogest"[Supplementary Concept] OR "dienogest"[tw] OR "Dimethisterone"[mesh] OR "Dimethisterone"[tw] OR "estradiol 17 beta-cypionate"[Supplementary Concept] OR "estradiol 17 beta-cypionate"[tw] OR "estradiol 3-benzoate"[Supplementary Concept] OR "estradiol 3-benzoate"[tw] OR "estropipate"[Supplementary Concept] OR "estropipate"[tw] OR "Ethinyl Estradiol"[mesh] OR "Ethinyl Estradiol"[tw] OR "ethinyl estradiol, levonorgestrel drug combination"[Supplementary Concept] OR "ethinyl estradiol, levonorgestrel drug combination"[tw] OR "Ethinyl Estradiol-Norgestrel Combination"[mesh] OR "Ethinyl Estradiol-Norgestrel Combination"[tw] OR "Ethinodiol Diacetate"[mesh] OR "Ethinodiol Diacetate"[tw] OR "etonogestrel"[Supplementary Concept] OR "etonogestrel"[tw] OR "Femovan"[supplementary concept] OR "Femovan"[tw] OR "fluprosteno"[Supplementary Concept] OR "fluprosteno"[tw] OR "Gestodene"[Supplementary Concept] OR "Gestodene"[tw] OR "Gestrinone"[mesh] OR "Gestrinone"[tw] OR "gossypol acetic acid"[supplementary concept] OR "gossypol acetic acid"[tw] OR "Infecundin"[Supplementary Concept] OR "Infecundin"[tw] OR "levonorgestrel butanoate"[Supplementary Concept] OR "levonorgestrel butanoate"[tw] OR "Levonorgestrel"[mesh] OR "Levonorgestrel"[tw] OR "Lynestrenol"[mesh] OR "Lynestrenol"[tw] OR "Medroxyprogesterone Acetate"[mesh] OR "Medroxyprogesterone Acetate"[tw] OR "Medroxyprogesterone"[mesh] OR "Medroxyprogesterone"[tw] OR "Megestrol"[mesh] OR "Megestrol"[tw] OR "Mestranol"[mesh] OR "Mestranol"[tw] OR "Mifepristone"[mesh] OR "Mifepristone"[tw] OR "neem oil"[supplementary concept] OR "neem oil"[tw] OR "Nonidet P-40"[supplementary concept] OR "Nonidet P-40"[tw] OR "Non-Ovlon"[Supplementary Concept] OR "Non-Ovlon"[tw] OR "Nonoxynol"[mesh] OR "Nonoxynol"[tw] OR "Noracycline"[Supplementary Concept] OR "Noracycline"[tw] OR "norelgestromin"[Supplementary Concept] OR "norelgestromin"[tw] OR "Norethindrone Acetate"[mesh] OR "Norethindrone Acetate"[tw] OR "norethindrone acetate, ethinyl estradiol, ferrous fumarate drug combination"[Supplementary Concept] OR "norethindrone acetate, ethinyl estradiol, ferrous fumarate drug combination"[tw] OR "norethindrone enanthate"[Supplementary Concept] OR "norethindrone enanthate"[tw] OR "Norethindrone"[mesh] OR "Norethindrone"[tw] OR "Norethynodrel"[mesh] OR "Norethynodrel"[tw] OR "norgestimate"[Supplementary Concept] OR "norgestimate"[tw] OR "norgestimate, ethinyl estradiol drug combination"[Supplementary Concept] OR "norgestimate, ethinyl estradiol drug combination"[tw] OR "Norgestrel"[mesh] OR "Norgestrel"[tw] OR "Norgestrienone"[mesh] OR "Norgestrienone"[tw] OR "Norinyl"[Supplementary Concept] OR "Norinyl"[tw] OR "NuvaRing"[Supplementary Concept] OR "NuvaRing"[tw] OR "Octoxynol"[mesh] OR "Octoxynol"[tw] OR "Ortho Evra"[Supplementary Concept] OR "Ortho Evra"[tw] OR "Oviol"[Supplementary Concept] OR "Oviol"[tw] OR "SC-11800 EE"[Supplementary Concept] OR "SC-11800 EE"[tw] OR "ST 1435"[Supplementary Concept] OR "ST 1435"[tw] OR "sulprostone"[Supplementary Concept] OR "sulprostone"[tw] OR "trestolone"[Supplementary Concept] OR "trestolone"[tw] OR "Triptorelin Pamoate"[mesh] OR "Triptorelin Pamoate"[tw] OR "Trisequens"[Supplementary Concept] OR "Trisequens"[tw] OR "ulipristal acetate"[Supplementary Concept] OR "ulipristal acetate"[tw] OR "Estrogens"[Mesh] OR "Estrogens"[Pharmacological Action] OR "estrogen"[tw] OR "estrogens"[tw] OR estrogen*[tw] OR "oestrogen"[tw] OR "oestrogens"[tw] OR oestrogen*[tw] OR "2,3-bis(3'-hydroxybenzyl)butane-1,4-diol"[Supplementary Concept] OR "2,3-bis(3'-hydroxybenzyl)butylolactone"[Supplementary Concept] OR "4-octylphenol"[Supplementary Concept] OR "8-prenylnaringenin"[Supplementary Concept] OR "biochanin A"[Supplementary Concept] OR "bisphenol A"[Supplementary Concept] OR "Chlorotrianisene"[mesh] OR "Coumestrol"[mesh] OR "daidzein"[Supplementary Concept] OR "Dienestrol"[mesh] OR "Diethylstilbestrol"[mesh] OR "diethylstilbestrol dipropionate"[Supplementary Concept] OR "Epimestrol"[mesh] OR "Equal"[mesh] OR "Estradiol"[mesh] OR "estradiol enanthate"[Supplementary Concept] OR "Estrogenic Steroids, Alkylated"[mesh] OR "Estrogens, Conjugated (USP)"[mesh] OR "estrogens, conjugated synthetic A"[Supplementary Concept] OR "estrogens, conjugated synthetic B"[Supplementary Concept] OR "Estrogens, Esterified (USP)"[mesh] OR "Estrone"[mesh] OR "Ethinyl Estradiol"[mesh] OR "formononetin"[Supplementary Concept] OR "Genistein"[mesh] OR "glycitein"[Supplementary Concept] OR "glycitin"[Supplementary Concept] OR "Hexestrol"[mesh] OR

"Infecundin"[Supplementary Concept] OR "Mestranol"[mesh] OR "nylestriol"[Supplementary Concept] OR "o,p'-DDT"[Supplementary Concept] OR "O-desmethylangolensin"[Supplementary Concept] OR "polyestradiol phosphate"[Supplementary Concept] OR "Quinestrol"[mesh] OR "secoisolaricresinol"[Supplementary Concept] OR "Zearalenone"[mesh] OR "Zeranol"[mesh] OR "2,3-bis(3'-hydroxybenzyl)butane-1,4-diol"[tw] OR "2,3-bis(3'-hydroxybenzyl)butyrolactone"[tw] OR "4-octylphenol"[tw] OR "8-prenylnaringenin"[tw] OR "biochanin A"[tw] OR "bisphenol A"[tw] OR "Chlorotrianisene"[tw] OR "Coumestrol"[tw] OR "daidzein"[tw] OR "Dienestrol"[tw] OR "Diethylstilbestrol"[tw] OR "diethylstilbestrol dipropionate"[tw] OR "Epimestrol"[tw] OR "Equol"[tw] OR "Estradiol"[tw] OR "estradiol enanthate"[tw] OR "Estrogenic Steroids, Alkylated"[tw] OR "Estrogens, Conjugated (USP)"[tw] OR "estrogens, conjugated synthetic A"[tw] OR "estrogens, conjugated synthetic B"[tw] OR "Estrogens, Esterified (USP)"[tw] OR "Estrone"[tw] OR "Ethinyl Estradiol"[tw] OR "formononetin"[tw] OR "Genistein"[tw] OR "glycitein"[tw] OR "glycitin"[tw] OR "Hexestrol"[tw] OR "Infecundin"[tw] OR "Mestranol"[tw] OR "nylestriol"[tw] OR "o,p'-DDT"[tw] OR "O-desmethylangolensin"[tw] OR "polyestradiol phosphate"[tw] OR "Quinestrol"[tw] OR "secoisolaricresinol"[tw] OR "Zearalenone"[tw] OR "Zeranol"[tw]) AND ("Depression"[Mesh] OR "Depressive Disorder"[Mesh] OR "depression"[tw] OR "depressive"[tw] OR "Mood Disorders"[Mesh] OR "Affect"[Mesh] OR "mood"[tw] OR "Cyclothymic Disorder"[tw] OR "Dysthymic Disorder"[tw] OR "Premenstrual Dysphoric Disorder"[tw] OR "Seasonal Affective Disorder"[tw] OR "Cyclothymic Disorders"[tw] OR "Dysthymic Disorders"[tw] OR "Premenstrual Dysphoric Disorders"[tw] OR "Seasonal Affective Disorders"[tw]) NOT (("Male"[mesh] OR "Men"[mesh] OR "male"[ti] OR "males"[ti] OR "men"[ti]) NOT ("Female"[mesh] OR "Women"[mesh] OR "Female"[ti] OR "Women"[ti])) NOT (("Infant"[mesh] OR "Child"[mesh] OR "Adolescent"[mesh]) NOT ("Adult"[mesh])) NOT ("Animals"[mesh] NOT "Humans"[mesh]) AND english[la] AND ((systematic review[ti] OR meta-analysis[pt] OR meta-analysis[ti] OR systematic literature review[ti] OR this systematic review[tw] OR pooling project[tw] OR (systematic review[tiab] AND review[pt]) OR meta synthesis[ti] OR meta-analy*[ti] OR integrative review[tw] OR integrative research review[tw] OR rapid review[tw] OR umbrella review[tw] OR consensus development conference[pt] OR practice guideline[pt] OR drug class reviews[ti] OR cochrane database syst rev[ta] OR acp journal club[ta] OR health technol assess[ta] OR evid rep technol assess summ[ta] OR jbi database system rev implement rep[ta]) OR (clinical guideline[tw] AND management[tw]) OR ((evidence based[ti] OR evidence-based medicine[mh] OR best practice*[ti] OR evidence synthesis[tiab]) AND (review[pt] OR diseases category[mh] OR behavior and behavior mechanisms[mh] OR therapeutics[mh] OR evaluation study[pt] OR validation study[pt] OR guideline[pt] OR pmcbook)) OR ((systematic[tw] OR systematically[tw] OR critical[tiab] OR (study selection[tw]) OR (predetermined[tw] OR inclusion[tw] AND criteri*[tw]) OR exclusion criteri*[tw] OR main outcome measures[tw] OR standard of care[tw] OR standards of care[tw]) AND (survey[tiab] OR surveys[tiab] OR overview*[tw] OR review[tiab] OR reviews[tiab] OR search*[tw] OR handsearch[tw] OR analysis[ti] OR critique[tiab] OR appraisal[tw] OR (reduction[tw] AND (risk[mh] OR risk[tw]) AND (death OR recurrence))) AND (literature[tiab] OR articles[tiab] OR publications[tiab] OR publication[tiab] OR bibliography[tiab] OR bibliographies[tiab] OR published[tiab] OR pooled data[tw] OR unpublished[tw] OR citation[tw] OR citations[tw] OR database[tiab] OR internet[tiab] OR textbooks[tiab] OR references[tw] OR scales[tw] OR papers[tw] OR datasets[tw] OR trials[tiab] OR meta-analy*[tw] OR (clinical[tiab] AND studies[tiab]) OR treatment outcome[mh] OR treatment outcome[tw] OR pmcbook)) NOT (letter[pt] OR newspaper article[pt]))

Supplement 24. Quality assessment: GRADE for the effects of hormonal contraceptives on depressive symptoms.

Here we describe the Grading of Recommendations Assessment, Development and Evaluation (GRADE) process resulting in the overall judgement of the certainty of evidence of a specific network estimate to be high, moderate, low, or very low. This resulted in an overall judgement per comparison. See *Supplement 24 Table 1*. The overall judgement per comparison is based on the judgement of each of GRADE domain: within-study bias, reporting bias, indirectness, imprecision, heterogeneity, and incoherence. The criteria for downgrading the certainty of evidence for a specific comparison within a certain GRADE domain are described below.

Supplement 24 Table 1. Overall judgement of the certainty of evidence per network estimate.

Formulation	Compared to	n*	Within-study bias	Reporting bias	Indirectness	Imprecision	Heterogeneity	Incoherence	Confidence rating
<i>Mixed evidence</i>									
Desogestrel (75ug)	Ethinylestradiol/desogestrel (20ug/150ug)	1							Very low
Desogestrel (75ug)	Ethinylestradiol/etonogestrel (15ug/120ug)	1							Very low
Estradiol/dienogest (3;2;2;1mg/0;2;3;0mg)	Ethinylestradiol/levonorgestrel (30ug/150ug)	1							Very low
Estradiol/nomegestrol (1.5mg/2.5mg)	Ethinylestradiol/drospirenone (30ug/3mg)	1							Very low
Estradiol/nomegestrol (1.5mg/2.5mg)	Placebo	1							Very low
Ethinylestradiol/desogestrel (20ug/150ug)	Ethinylestradiol/etonogestrel (15ug/120ug)	1							Very low
Ethinylestradiol/desogestrel (20ug/150ug)	Ethinylestradiol/devonorgestrel (20ug/100ug)	1							Very low
Ethinylestradiol/drospirenone (30ug/3mg)	Ethinylestradiol/etonogestrel (15ug/120ug)	1							Very low
Ethinylestradiol/drospirenone (30ug/3mg)	Ethinylestradiol/levonorgestrel (30ug/150ug)	2							Very low
Ethinylestradiol/levonorgestrel (20ug/100ug)	Placebo	1							Very low
Ethinylestradiol/levonorgestrel (30ug/150ug)	Levonorgestrel (30ug)	2							Very low
Ethinylestradiol/levonorgestrel (30ug/150ug)	Placebo	4							Very low
Levonorgestrel (30ug)	Placebo	2							Very low
<i>Indirect evidence</i>									
Desogestrel (75ug)	Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)	-							Very low
Desogestrel (75ug)	Estradiol/nomegestrol (1.5mg/2.5mg)	-							Very low
Desogestrel (75ug)	Ethinylestradiol/drospirenone (30ug/3ug)	-							Very low
Desogestrel (75ug)	Ethinylestradiol/levonorgestrel (20ug/100ug)	-							Very low
Desogestrel (75ug)	Ethinylestradiol/levonorgestrel (30ug/150ug)	-							Very low

Formulation	Compared to	n*	Within-study bias	Reporting bias	Indirectness	Imprecision	Heterogeneity	Incoherence	Confidence rating
Desogestrel (75ug)	Levonorgestrel (30ug)	-	●	●	●	●	●	●	Very low
Desogestrel (75ug)	Placebo	-	●	●	●	●	●	●	Very low
Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)	Estradiol/nomegestrol (1.5mg/2.5mg)	-	●	●	●	●	●	●	Very low
Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)	Ethinylestradiol/desogestrel (20ug/150ug)	-	●	●	●	●	●	●	Very low
Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)	Ethinylestradiol/drospirenone (30ug/3mg)	-	●	●	●	●	●	●	Very low
Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)	Ethinylestradiol/etonogestrel (15ug/120ug)	-	●	●	●	●	●	●	Very low
Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)	Ethinylestradiol/levonorgestrel (20ug/100ug)	-	●	●	●	●	●	●	Very low
Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)	Levonorgestrel (30ug)	-	●	●	●	●	●	●	Very low
Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)	Placebo	-	●	●	●	●	●	●	Very low
Estradiol/nomegestrol (1.5mg/2.5mg)	Ethinylestradiol/desogestrel (20ug/150ug)	-	●	●	●	●	●	●	Very low
Estradiol/nomegestrol (1.5mg/2.5mg)	Ethinylestradiol/etonogestrel (15ug/120ug)	-	●	●	●	●	●	●	Very low
Estradiol/nomegestrol (1.5mg/2.5mg)	Ethinylestradiol/levonorgestrel (20ug/100ug)	-	●	●	●	●	●	●	Very low
Estradiol/nomegestrol (1.5mg/2.5mg)	Ethinylestradiol/levonorgestrel (30ug/150ug)	-	●	●	●	●	●	●	Very low
Estradiol/nomegestrol (1.5mg/2.5mg)	Levonorgestrel (30ug)	-	●	●	●	●	●	●	Very low
Ethinylestradiol/desogestrel (20ug/150ug)	Ethinylestradiol/drospirenone (30ug/3mg)	-	●	●	●	●	●	●	Very low
Ethinylestradiol/desogestrel (20ug/150ug)	Ethinylestradiol/levonorgestrel (30ug/150ug)	-	●	●	●	●	●	●	Very low
Ethinylestradiol/desogestrel (20ug/150ug)	Levonorgestrel (30ug)	-	●	●	●	●	●	●	Very low
Ethinylestradiol/desogestrel (20ug/150ug)	Placebo	-	●	●	●	●	●	●	Very low
Ethinylestradiol/drospirenone (30ug/3mg)	Ethinylestradiol/levonorgestrel (20ug/100ug)	-	●	●	●	●	●	●	Very low
Ethinylestradiol/drospirenone (30ug/3mg)	Levonorgestrel (30ug)	-	●	●	●	●	●	●	Very low
Ethinylestradiol/drospirenone (30ug/3mg)	Placebo	-	●	●	●	●	●	●	Very low
Ethinylestradiol/etonogestrel (15ug/120ug)	Ethinylestradiol/levonorgestrel (20ug/100ug)	-	●	●	●	●	●	●	Very low
Ethinylestradiol/etonogestrel (15ug/120ug)	Ethinylestradiol/levonorgestrel (30ug/150ug)	-	●	●	●	●	●	●	Very low

Formulation	Compared to	n*	Within-study bias	Reporting bias	Indirectness	Imprecision	Heterogeneity	Incoherence	Confidence rating
Ethinylestradiol/etonogestrel (15ug/120ug)	Levonorgestrel (30ug)	-							Very low
Ethinylestradiol/etonogestrel (15ug/120ug)	Placebo	-							Very low
Ethinylestradiol/levonorgestrel (20ug/100ug)	Ethinylestradiol/levonorgestrel (30ug/150ug)	-							Very low
Ethinylestradiol/levonorgestrel (20ug/100ug)	Levonorgestrel (30ug)	-							Very low
			no concerns	some concerns	high concerns				

*Refers to the number of studies investigating the comparison.

The table describes the judgement of certainty of evidence for each of the GRADE domains that results in an overall judgement of certainty of evidence per comparison. Ethinylestradiol/dienogest (3;2;2;1mg/0;2;3;0mg) is a multiphasic oral contraceptive pill that has four different dosages of hormones throughout a 4-week cycle. The dosages before the / refer to the different dosages of ethinylestradiol, and the ones after the / to those of dienogest.

1. Within study bias

We rated the overall risk of bias within each study according to the rating of the domain rated with the highest risk of bias. Of the trials included in the meta-analysis, we judged six trials to have a high risk of bias,^{330, 404, 416-418, 421} five trials to have some concerns,^{346, 405, 414, 415, 420} and only one study to have low risk of bias.⁴¹⁹ Most often the bias arose from missing outcome data.^{330, 346, 403, 404, 416-418, 420, 421} In these cases, plus two other trials, there were also some^{330, 346, 403-405, 414-416, 418, 420} or high^{417, 421, 423} concerns about the selection of the reported results. Two studies that were not eligible for inclusion in the network meta-analysis as the interventions were not connected with the network^{403, 423} were also judged to have a high risk of bias due to missing data on the outcome⁴⁰³ or due to selection of the reported results.⁴²³ See also *Supplement 24 Table 2*.

Supplement 24 Table 2. Study specific risk of bias judgements.

Study	Randomization	Deviation from intended intervention	Missing outcome data	Measurements of outcome	Selection of the reported result	Overall
<i>Included in the network meta-analysis</i>						
Battaglia	●	●	●	●	●	●
Davis	●	●	●	●	●	●
Elaut	●	●	●	●	●	●
Engman	●	●	●	●	●	●
Graham	●	●	●	●	●	●
Kelly	●	●	●	●	●	●
Lundin	●	●	●	●	●	●
O'Connell	●	●	●	●	●	●
Sangthawan	●	●	●	●	●	●
Winkler	●	●	●	●	●	●
Witjes	●	●	●	●	●	●
Zethraeus	●	●	●	●	●	●
<i>Included in the systematic review only</i>						
Greco	●	●	●	●	●	●
Legro	●	●	●	●	●	●

● low concerns, ● some concerns, ● high concerns

The table shows the judged risk of bias in each domain for every study included in the review. Risk of bias were determined using the Cochrane Risk of Bias tool.⁴¹²

We downgraded the confidence rating for a network estimate by one level when the contributions were on average coming from “some concerns” risk of bias comparisons or with two levels when the contributions were mostly coming from “high” risk of bias comparisons.⁴¹⁰ See *Supplement 24 Figure 1*.

Formulation*Mixed evidence*

Desogestrel (75ug)
 Desogestrel (75ug)
 Estradiol/dienogest (3;2;2;1mg/0;2;3;0mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Levonorgestrel (30ug)

Indirect evidence

Desogestrel (75ug)
 Desogestrel (75ug)
 Desogestrel (75ug)
 Desogestrel (75ug)
 Desogestrel (75ug)
 Desogestrel (75ug)
 Desogestrel (75ug)
 Desogestrel (75ug)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Ethinylestradiol/levonorgestrel (20ug/100ug)

● no concerns ● some concerns ● high concerns

Compared to

Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Placebo
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Placebo
 Levonorgestrel (30ug)
 Placebo
 Placebo

Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Levonorgestrel (30ug)
 Placebo
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Levonorgestrel (30ug)
 Placebo
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Levonorgestrel (30ug)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Levonorgestrel (30ug)
 Placebo
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Levonorgestrel (30ug)
 Placebo
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Levonorgestrel (30ug)
 Placebo
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Levonorgestrel (30ug)

Supplement 24 Figure 1. Network estimate specific risk of bias judgements.

Risk or bias judgement for average quality of network estimate. Each bar represents a relative treatment effect estimated from the network shown in *Figure 11.2*. White vertical lines separate the percentage contribution of different studies. Each bar shows the percentage contribution from studies judged to be at no concerns, some concerns, and high concerns. See the text above below the corresponding header “within-study bias” for details about rules used to judge comparison as having no, some or high concerns. Ethinylestradiol/dienogest (3;2;2;1mg/0;2;3;0mg) is a multiphasic oral contraceptive pill that has four different dosages of hormones throughout a 4-week cycle. The dosages before the / refer to the different dosages of ethinylestradiol, and the ones after the / to those of dienogest.

2. Reporting bias

The symmetrical comparison adjusted funnel plot did not suggest presence of overall publication bias in comparison with placebo. See *Supplement 24 Figure 2*. We also managed to retrieve additional supplementary and unpublished data that showed similar effect sizes as the published data. Hence, none of the studies was downgraded for potential publication bias.⁴¹⁰

Supplement 24 Figure 2. Funnel plot of randomized clinical trials of hormonal contraceptives versus placebo.

The color of the nodes corresponds with the type of hormonal contraceptives compared in the specific study (arm) (oral contraceptive pill; progesterone-only pill; vaginal ring; or placebo).

3. Indirectness

The distribution of age was balanced across comparisons (see also *Table 11.1*). However, full assessment of transitivity was challenging as information on many potential effect moderators, such as previously negative experiences with hormonal

contraceptive use and previous MDD, was not provided. Moreover, a major group of users, first-time users, was not or only very limited at most, included in the trials. Hence, the sample was not fully representative of hormonal contraceptive users. Also, the majority of interventions were only poorly connected to the network and therefore the evaluation of transitivity for these comparisons is unclear.⁴¹⁰ Hence, we downgraded comparisons to “some concerns” if they did not include first-time users or a group of women with previously negative experiences with hormonal contraceptive use, and additionally downgraded to “major concerns” if the comparison was only investigated once for this domain. Subsequently, we downgraded the confidence rating for a network estimate by one level when the contributions were on average coming from “some concerns” risk of bias comparisons or with two levels when the contributions were predominantly coming from “high” risk of bias comparisons.⁴¹⁰ See *Supplement 24 Figure 3*.

4. Imprecision

We considered 0.20 as the minimally clinically meaningful important difference. This was based on the rule of thumb for small effect sizes.²¹⁸ A rating of ‘major concerns’ was assigned to comparisons if the 95% credible interval of the effect extended beyond the area of equivalence on the opposite side of the no effect line as the point estimate (credible interval ranged from below -0.20 to beyond 0.20), so that the estimated treatment effect is compatible with clinically important effects in both directions. A rating of ‘some concerns’ was assigned if the credible interval extended into but not beyond the area of equivalence on the opposite side of the no effect line. ‘No concerns’ were assigned if the credible interval was entirely on one side of the no effect line (below -0.20, or beyond 0.20), or if it was entirely within the area of equivalence (beyond -0.20 and below 0.20).⁴¹⁰

5. Heterogeneity

Imprecision and incoherence are related to heterogeneity. Hence, we only rated heterogeneity for comparisons that were not already downgraded on the domains of imprecision and incoherence. In these two cases, the amount of heterogeneity was assessed with the I^2 . The I^2 of the network meta-analysis was acceptable (43.7%), suggesting moderate between-trial heterogeneity. Hence, we rated these network estimate as “some concerns” for heterogeneity.

6. Incoherence

For incoherence, we looked at the results of the node splitting and we downgraded the comparisons with important inconsistency ($p < 0.10$).⁴¹⁰ For the majority of closed loops (five out of eight) the indirect and direct effects were in the same direction. For a minority of closed loops however (i.e., ethinylestradiol/drospirenone [30ug/3mg] versus estradiol/nomegestrel [1.5mg/2.5mg], placebo versus estradiol/nomegestrel [1.5mg/2.5mg], and ethinylestradiol/levonorgestrel [30ur/150ug] versus ethinylestradiol/drospirenone [30ug/3mg]) these effects were in the opposite direction, though these differences were not significant ($p = 0.17$, $p = 0.17$, and $p = 0.19$, respectively). As none of the p-values were below 0.10, we did not downgrade any of these comparisons. See *Supplement 24 Figure 4*. In case there was only direct or indirect evidence for a comparison, we rated the certainty of evidence for the incoherence domain as “some concerns”.⁴¹⁰ See *Supplement 24 Table 1*.

Formulation*Mixed evidence*

Desogestrel (75ug)
 Desogestrel (75ug)
 Estradiol/dienogest (3;2;2;1mg/0;2;3;0mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Levonorgestrel (30ug)

Indirect evidence

Desogestrel (75ug)
 Desogestrel (75ug)
 Desogestrel (75ug)
 Desogestrel (75ug)
 Desogestrel (75ug)
 Desogestrel (75ug)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Ethinylestradiol/levonorgestrel (20ug/100ug)

Compared to

Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Placebo
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Placebo
 Levonorgestrel (30ug)
 Placebo
 Placebo

Estradiol/dienogest (3;2;2;1mg/ 0;2;3;0mg)
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Levonorgestrel (30ug)
 Placebo
 Estradiol/nomegestrol (1.5mg/2.5mg)
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Levonorgestrel (30ug)
 Placebo
 Ethinylestradiol/desogestrel (20ug/150ug)
 Ethinylestradiol/etonogestrel (15ug/120ug)
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Levonorgestrel (30ug)
 Ethinylestradiol/drospirenone (30ug/3mg)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Levonorgestrel (30ug)
 Placebo
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Levonorgestrel (30ug)
 Placebo
 Ethinylestradiol/levonorgestrel (20ug/100ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Levonorgestrel (30ug)
 Ethinylestradiol/levonorgestrel (30ug/150ug)
 Levonorgestrel (30ug)

● no concerns ● some concerns ● high concerns

Supplement 24 Figure 3. Network estimate specific indirectness.

Indirectness judgement for average quality of network estimate. Each bar represents a relative treatment effect estimated from the network shown in *Figure 11.2*. White vertical lines separate the percentage contribution of different studies. Each bar shows the percentage contribution from studies judged to be at no concerns, some concerns, and high concerns. See the text above below the corresponding header “indirectness” for details about rules used to judge comparison as having no, some or high concerns. Ethinylestradiol/dienogest (3;2;2;1mg/0;2;3;0mg) is a multiphasic oral contraceptive pill that has four different dosages of hormones throughout a 4-week cycle. The dosages before the / refer to the different dosages of ethinylestradiol, and the ones after the / to those of dienogest.

Supplement 24 Figure 4. Estimates of direct and indirect effects.

Summary of a node-split analysis consisting of nine separate node-splitting models that show direct, indirect and the network estimate.

Chapter 12

Supplement 25. Quality assessment: GRADE for the efficacy of oral contraceptive pills on depressive symptoms in women with premenstrual syndrome of premenstrual dysphoric disorder.

Here we describe the GRADE process resulting in the overall judgement of the certainty of evidence of a specific network estimate to be high, moderate, low, or very low. This resulted in an overall judgement per comparison. See *Supplement 25 Table 1*. The overall judgement per comparison is based on the judgement of each of GRADE domain: within-study bias, reporting bias, indirectness, imprecision, heterogeneity, and incoherence. The criteria for downgrading the certainty of evidence for a specific comparison within a certain GRADE domain are described below.

Supplement 25 Table 1. Overall judgement of the certainty of evidence per network estimate.

Formulation	Compared to	n*	Within-study bias	Reporting bias	Indirectness	Imprecision	Heterogeneity	Incoherence	Confidence rating
<i>Mixed evidence</i>									
Estradiol norgestrel (1.5mg/2.5mg; 24/4)	Placebo	1	●	●	●	●	●	●	Very low
Ethinylestradiol drospirone (20mg/3mg; 21/7)	Ethinylestradiol drospirone (20mg/3mg; continuous)	1	●	●	●	●	●	●	Very low
Ethinylestradiol drospirone (20mg/3mg; 21/7)	Placebo	1	●	●	●	●	●	●	Very low
Ethinylestradiol drospirone (20mg/3mg; 24/4)	Ethinylestradiol desogestrel (20mg/150mg; 24/4)	1	●	●	●	●	●	●	Very low
Ethinylestradiol drospirone (20mg/3mg; 24/4)	Placebo	2	●	●	●	●	●	●	Very low
Ethinylestradiol drospirone (20mg/3mg; continuous)	Placebo	1	●	●	●	●	●	●	Very low
Ethinylestradiol levonorgestrel (20mg/90mg; continuous)	Placebo	2	●	●	●	●	●	●	Very low

Formulation	Compared to	n*	Within-study bias	Reporting bias	Indirectness	Imprecision	Heterogeneity	Incoherence	Confidence rating
Ethinylestradiol drosperinone (30mg/3mg; 21/7)	Placebo	1	●	●	●	●	●	●	Very low
Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0; 21/7)	Placebo	1	●	●	●	●	●	●	Very low
<i>Indirect evidence</i>									
Estradiol nomegestrol (1.5mg/2.5mg; 24/4)	Ethinylestradiol drosperinone (20mg/3mg; 21/7)	-	●	●	●	●	●	●	Very low
Estradiol nomegestrol (1.5mg/2.5mg; 24/4)	Ethinylestradiol drosperinone (20mg/3mg; 24/4)	-	●	●	●	●	●	●	Very low
Estradiol nomegestrol (1.5mg/2.5mg; 24/4)	Ethinylestradiol drosperinone (20mg/3mg; continuous)	-	●	●	●	●	●	●	Very low
Estradiol nomegestrol (1.5mg/2.5mg; 24/4)	Ethinylestradiol desogestrel (20mg/150mg; 24/4)	-	●	●	●	●	●	●	Very low
Estradiol nomegestrol (1.5mg/2.5mg; 24/4)	Ethinylestradiol levonorgestrel (20mg/90mg; continuous)	-	●	●	●	●	●	●	Very low
Estradiol nomegestrol (1.5mg/2.5mg; 24/4)	Ethinylestradiol drosperinone (30mg/3mg; 21/7)	-	●	●	●	●	●	●	Very low
Estradiol nomegestrol (1.5mg/2.5mg; 24/4)	Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0; 21/7)	-	●	●	●	●	●	●	Very low
Ethinylestradiol drosperinone (20mg/3mg; 21/7)	Ethinylestradiol drosperinone (20mg/3mg; 24/4)	-	●	●	●	●	●	●	Very low
Ethinylestradiol drosperinone (20mg/3mg; 21/7)	Ethinylestradiol desogestrel (20mg/150mg; 24/4)	-	●	●	●	●	●	●	Very low
Ethinylestradiol drosperinone (20mg/3mg; 21/7)	Ethinylestradiol levonorgestrel (20mg/90mg; continuous)	-	●	●	●	●	●	●	Very low
Ethinylestradiol drosperinone (20mg/3mg; 21/7)	Ethinylestradiol drosperinone (30mg/3mg; 21/7)	-	●	●	●	●	●	●	Very low
Ethinylestradiol drosperinone (20mg/3mg; 21/7)	Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0; 21/7)	-	●	●	●	●	●	●	Very low
Ethinylestradiol drosperinone (20mg/3mg; 24/4)	Ethinylestradiol drosperinone (20mg/3mg; continuous)	-	●	●	●	●	●	●	Very low

Formulation	Compared to	n*	Within-study bias	Reporting bias	Indirectness	Imprecision	Heterogeneity	Incoherence	Confidence rating
Ethinylestradiol drospirone (20mg/3mg; 24/4)	Ethinylestradiol levonorgestrel (20mg/90mg; continuous)	-							Very low
Ethinylestradiol drospirone (20mg/3mg; 24/4)	Ethinylestradiol drospirone (30mg/3mg; 21/7)	-							Very low
Ethinylestradiol drospirone (20mg/3mg; 24/4)	Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0; 21/7)	-							Very low
Ethinylestradiol drospirone (20mg/3mg; continuous)	Ethinylestradiol desogestrel (20mg/150mg; 24/4)	-							Very low
Ethinylestradiol drospirone (20mg/3mg; continuous)	Ethinylestradiol levonorgestrel (20mg/90mg; continuous)	-							Very low
Ethinylestradiol drospirone (20mg/3mg; continuous)	Ethinylestradiol drospirone (30mg/3mg; 21/7)	-							Very low
Ethinylestradiol drospirone (20mg/3mg; continuous)	Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0; 21/7)	-							Very low
Ethinylestradiol desogestrel (20mg/150mg; 24/4)	Ethinylestradiol levonorgestrel (20mg/90mg; continuous)	-							Very low
Ethinylestradiol desogestrel (20mg/150mg; 24/4)	Ethinylestradiol drospirone (30mg/3mg; 21/7)	-							Very low
Ethinylestradiol desogestrel (20mg/150mg; 24/4)	Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0; 21/7)	-							Very low
Ethinylestradiol desogestrel (20mg/150mg; 24/4)	Placebo	-							Very low
Ethinylestradiol levonorgestrel (20mg/90mg; continuous)	Ethinylestradiol drospirone (30mg/3mg; 21/7)	-							Very low
Ethinylestradiol levonorgestrel (20mg/90mg; continuous)	Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0; 21/7)	-							Very low
Ethinylestradiol drospirone (30mg/3mg; 21/7)	Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0; 21/7)	-							Very low

no concerns some concerns high concerns

*Refers to the number of studies investigating the comparison. The table describes the judgement of certainty of evidence for each of the GRADE domains that results in an overall judgement of certainty of evidence per comparison. Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0; 21/7) is a multiphasic oral contraceptive pill that has four different dosages of progesterone throughout a 4-week cycle. The dosages after the / refer to the different dosages of progesterone.

1. Within study bias

We rated the overall risk of bias within each study according to the rating of the domain rated with the highest risk of bias. Of the trials included in the meta-analysis, we judged four trials to have a high risk of bias,^{330, 437, 438, 443} and five trials to have some concerns.^{435, 436, 438-440, 444} Most often the bias arose from the fact that trials had missing outcome data for more than 5% of the participants and due to the selection of the reported results. See also *Supplement 25 Table 2*.

We downgraded the confidence rating for a network estimate by one level when the contributions were on average coming from “some concerns” risk of bias comparisons or with two levels when the contributions were mostly coming from “high” risk of bias comparisons.⁴¹⁰ If the contribution was coming from equal amounts of “some concerns” and “high” risk of biases, we downgraded the confidence for that specific network estimate with two levels. See *Supplement 25 Figure 1*.

Supplement 25 Table 2. Study specific risk of bias judgements.

Study	Randomization	Deviation from intended intervention	Missing outcome data	Measurements of outcome	Selection of the reported result	Overall
Eisenlohr-Moul
Freeman'12
Freeman'01
Graham
Halbreich
Lundin
Pearlstein
Wichianpitaya
Yonkers

no concerns some concerns high concerns

The table shows the judged risk of bias in each domain for every study included in the review. Risk of bias were determined using the Cochrane Risk of Bias tool.⁴¹²

Supplement 25 Figure 2. Funnel plot of randomized clinical trials on the effects of oral contraceptive pills versus placebo on depressive symptoms.

3. Indirectness

The distribution of age was balanced across comparisons. See *Table 12.1*. However, the majority of interventions were only poorly connected to the network and therefore the evaluation of transitivity for these comparisons is unclear.⁴¹⁰ Hence, we downgraded comparisons to “some concerns” for this domain if the comparison was only investigated once. Subsequently, we downgraded the confidence rating for a network estimate by one level when the contributions were on average coming from “some concerns” risk of bias comparisons.⁴¹⁰ If the contribution was coming from equal amounts of “some concerns” and “high” risk of biases, we downgraded the confidence for that specific network estimate with two levels. See *Supplement 25 Figure 3*.

Formulation

Mixed evidence

Ethinylestradiol drospirenone (20mg/3mg, 21/7)
Ethinylestradiol drospirenone (20mg/3mg, 21/7)
Ethinylestradiol drospirenone (20mg/3mg, 24/4)
Ethinylestradiol drospirenone (20mg/3mg, 24/4)
Ethinylestradiol drospirenone (20mg/3mg, continuous)
Ethinylestradiol levonorgestrel (20mg/90mcg, continuous)
Ethinylestradiol drospirenone (30mg/3mg, 21/7)
Ethinylestradiol norethindrone (35mcg/0.5; 1.0; 0.5; 21/7)

Compared to

Placebo
Ethinylestradiol drospirone (20mg/3mg; continuous)
Placebo
Ethinylestradiol desogestrel (20mg/150mg; 24/4)
Placebo
Placebo
c) Placebo
Placebo
Placebo

Indirect evidence

Ethinylestradiol/norgestrel (1.5mg/2.5mg; 24/4)
Ethinylestradiol/norgestrel (1.5mg/2.5mg; 24/4)
Ethinylestradiol/norgestrel (1.5mg/2.5mg; 24/4)
Ethinylestradiol/norgestrel (1.5mg/2.5mg; 24/4)
Ethinylestradiol/norgestrel (1.5mg/2.5mg; 24/4)
Ethinylestradiol/norgestrel (1.5mg/2.5mg; 24/4)
Ethinylestradiol/drospirenone (20mg/3mg; 21/7)
Ethinylestradiol/drospirenone (20mg/3mg; 21/7)
Ethinylestradiol/drospirenone (20mg/3mg; 21/7)
Ethinylestradiol/drospirenone (20mg/3mg; 21/7)
Ethinylestradiol/drospirenone (20mg/3mg; 21/7)
Ethinylestradiol/drospirenone (20mg/3mg; 24/4)
Ethinylestradiol/drospirenone (20mg/3mg; 24/4)
Ethinylestradiol/drospirenone (20mg/3mg; 24/4)
Ethinylestradiol/drospirenone (20mg/3mg; continuous)
Ethinylestradiol/drospirenone (20mg/3mg; continuous)
Ethinylestradiol/drospirenone (20mg/3mg; continuous)
Ethinylestradiol/drospirenone (20mg/3mg; continuous)
Ethinylestradiol/desogestrel (20mg/150mg; 24/4)
Ethinylestradiol/desogestrel (20mg/150mg; 24/4)
Ethinylestradiol/desogestrel (20mg/150mg; 24/4)
Ethinylestradiol/desogestrel (20mg/150mg; 24/4)
Ethinylestradiol/levonorgestrel (20mg/90mg; continuous)
Ethinylestradiol/levonorgestrel (20mg/90mg; continuous)
Ethinylestradiol/drospirenone (30mg/3mg; 21/7)

Ethinylestradiol drospirenone (20mg/3mg; 21/7)
Ethinylestradiol drospirenone (20mg/3mg; 24/4)
Ethinylestradiol drospirenone (20mg/3mg; continuous)
Ethinylestradiol desogestrel (20mg/150mg; 24/4)
Ethinylestradiol levonorgestrel (20mg/90mg; continuous)
Ethinylestradiol drospirenone (30mg/3mg; 21/7)
Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0.5; 21/7)
Ethinylestradiol drospirenone (20mg/3mg; 24/4)
Ethinylestradiol desogestrel (20mg/150mg; 24/4)
Ethinylestradiol levonorgestrel (20mg/90mg; continuous)
Ethinylestradiol drospirenone (30mg/3mg; 21/7)
Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0.5; 21/7)
Ethinylestradiol drospirenone (20mg/3mg; continuous)
Ethinylestradiol levonorgestrel (20mg/90mg; continuous)
Ethinylestradiol drospirenone (30mg/3mg; 21/7)
Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0.5; 21/7)
Ethinylestradiol desogestrel (20mg/150mg; 24/4)
Ethinylestradiol levonorgestrel (20mg/90mg; continuous)
Ethinylestradiol drospirenone (30mg/3mg; 21/7)
Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0.5; 21/7)
Ethinylestradiol levonorgestrel (20mg/90mg; continuous)
Ethinylestradiol drospirenone (30mg/3mg; 21/7)
Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0.5; 21/7)
Placebo
s)Ethinylestradiol drospirenone (30mg/3mg; 21/7)
s)Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0.5; 21/7)
s)Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0.5; 21/7)

● no concerns ● some concerns ● high concerns

Supplement 25 Figure 3. Network estimate specific indirectness.

Indirectness judgement for average quality of network estimate. Each bar represents a relative treatment effect estimated from the network shown in *Figure 12.3*. White vertical lines separate the percentage contribution of different studies. Each bar shows the percentage contribution from studies judged to be at no concerns, some concerns, and high concerns. See the text above below the corresponding header “indirectness” for details about rules used to judge comparison as having no, some or high concerns. Ethinylestradiol norethindrone (35mg/0.5;1;0.5;0; 21/7) is a multiphasic oral contraceptive pill that has four different dosages of progesterone throughout a 4-week cycle. The dosages after the / refer to the different dosages of progesterone.

4. Imprecision

We considered 0.20 as the minimally clinically meaningful important difference. This was based on the rule of thumb for small effect sizes.²¹⁸ A rating of 'major concerns' was assigned to comparisons if the 95% credible interval of the effect extended beyond the area of equivalence on the opposite side of the no effect

line as the point estimate (credible interval ranged from below -0.20 to beyond 0.20), so that the estimated treatment effect is compatible with clinically important effects in both directions. A rating of 'some concerns' was assigned if the credible interval extended into but not beyond the area of equivalence on the opposite side of the no effect line. 'No concerns' were assigned if the credible interval was entirely on one side of the no effect line (below -0.20, or beyond 0.20), or if it was entirely within the area of equivalence (beyond -0.20 and below 0.20).⁴¹⁰

4. Heterogeneity

Imprecision and incoherence are related to heterogeneity. Hence, we only rated heterogeneity for comparisons that were not already downgraded on the domains of imprecision and incoherence. In these cases, the heterogeneity could not be reliably tested due to the low number of comparisons. Hence, we rated these network estimate as "some concerns" for heterogeneity.

5. Incoherence

As there we no closed loops in the network, no node split procedure could be established to assess the incoherence of direct and indirect effects. In case there was only indirect or direct evidence for a comparison, we rated the certainty of evidence for the incoherence domain as "some concerns".⁴¹⁰ See *Supplement 25 Table 1*.

Summary in Dutch / Nederlandse samenvatting

In dit proefschrift is onderzocht of geslachtshormonen een rol kunnen spelen bij depressie, en zo ja, hoe en bij wie. Dit is gedaan door een aantal studies rondom twee kernvragen uit te voeren:

1. Is er een verband tussen geslachtshormoongehalte en depressie?
2. Kan pilgebruik depressieve klachten veroorzaken?

Toen bleek dat deze onderwerpen veel losmaakten in de media, werd nog een derde thema besproken:

3. Wat kunnen wetenschappers doen om wetenschapsjournalistiek te verbeteren?

In deze samenvatting wordt eerst uitgelegd waarom de rol van geslachtshormonen bij depressie is onderzocht. Daarna worden de resultaten van de verschillende studies besproken. Ook wordt uitgelegd wat deze resultaten betekenen voor patiënten met depressieve klachten en voor toekomstig onderzoek op dit gebied. In de conclusie wordt teruggekomen op de vraag waar dit proefschrift mee begon: "Can hormones get you down?".

Geslachtshormonen en depressie?

Depressie komt helaas veel voor. Één op de vier vrouwen en één op de acht mannen krijgt deze ziekte. Een depressie kan worden vastgesteld als een patiënt minimaal vijf verschillende depressieve klachten ervaart gedurende twee weken. De aard van deze klachten loopt sterk uiteen. De belangrijkste klacht is somberheid, maar ook andere klachten - zoals vermoeidheid en slaapproblemen - komen vaak voor. In ieder geval hebben depressieve klachten (ook zonder de diagnose depressie) een grote invloed op het welzijn.

Er is niet één oorzaak aan te wijzen voor het ontstaan van een depressie. Meestal speelt een combinatie van persoonlijke eigenschappen, stressvolle gebeurtenissen en biologische factoren een rol in het ontstaan van de ziekte. Een mogelijk belangrijke biologische factor is de invloed van geslachtshormonen. Deze hormonen beïnvloeden namelijk processen in de hersenen die belangrijk zijn voor het onderhouden van een goede stemming. Sommige vrouwen merken dit doordat

ze bij veranderingen in hormoongehaltenes meer depressieve klachten ervaren. Een ander voorbeeld is dat vrouwen die “de pil” gebruiken, soms depressieve klachten melden als bijwerking.

Het lijkt er dus op dat er verschillende manieren zijn waarop geslachtshormonen een rol zouden kunnen spelen bij het ontstaan en het in stand houden van een depressie. Onbekend is hoe deze geslachtshormonen dat dan precies doen. Om hier meer duidelijkheid over te krijgen, is in dit proefschrift onderzocht hoe geslachtshormonen samenhangen met depressie. Dit biedt mogelijk handvatten om in de toekomst patiënten bij wie de depressie (mede) is ontstaan door de invloed van geslachtshormonen, beter te kunnen behandelen.

Vraag 1: Is er een verband tussen geslachtshormoongehaltenes en depressie?

Er zijn vele verschillende geslachtshormonen in ons lichaam actief, waarvan de bekendste twee testosteron en oestrogeen zijn. Zowel mannen als vrouwen hebben beide hormonen in hun bloed, maar de spiegels verschillen tussen de geslachten. Mannen hebben bijvoorbeeld gemiddeld 14 keer zoveel testosteron in hun bloed als vrouwen. Het zorgt bijvoorbeeld voor spiergroei, lichaamsbehaarings, de ontwikkeling van de penis en de clitoris, en zin in seks. Oestrogeen zorgt daarentegen voor borstgroei en het behoud van sterke botten, en is betrokken bij de menstruatie. Omdat oestrogeen en testosteron, naast hun invloed op lichamelijke kenmerken, ook invloed uitoefenen op processen in de hersenen die belangrijk zijn voor het onderhouden van een goede stemming, spelen deze geslachtshormonen mogelijk ook een rol bij depressie. Daarom werd hun samenhang met depressie en depressieve klachten in dit eerste deel van het proefschrift onderzocht.

Hoe meet je testosteron?

Het gehalte aan geslachtshormonen wordt in grote onderzoeken meestal gemeten met een bloedtest, omdat bekend is dat dit een betrouwbare meetmethode is. Een bloedafname kan echter pijnlijk zijn, waardoor niet iedereen hieraan wil meewerken. Geslachtshormonen meten met een speekseltest zou daarom een gemakkelijk alternatief zijn, maar de betrouwbaarheid zo’n meting is nog onvoldoende bekend. Daarom is onderzocht of de in het speeksel gemeten

testosterongehalten van 2.048 vrouwen en mannen overeenkwamen met de testosterongehalten gemeten in het bloed van dezelfde personen (**hoofdstuk 2**). Helaas bleken de resultaten van deze twee meetmethodes niet goed overeen te komen. Daarom is in het onderzoek voor dit proefschrift gebruik gemaakt van testosterongehalten gemeten in bloed.

Hebben vrouwen met een hoog en mannen met een laag testosterongehalte vaker een depressie?

Vrouwen met te hoge testosterongehalten en mannen met te lage testosterongehalten zijn vaker depressief. Er wordt daarom gedacht dat ook vrouwen met een relatief hoog en mannen met een relatief laag testosterongehalte een grotere kans hebben op het krijgen van een depressie dan vrouwen en mannen met gemiddelde testosterongehalten (zie *Figuur 1*). Of dit waar was werd onderzocht bij 2.842 vrouwen en mannen bij wie eenmalig testosteron was gemeten en van wie bekend was of zij depressief waren of werden in de jaren erna (**hoofdstuk 3 en 4**). Er was echter geen verband tussen de hoogte van het testosterongehalte en het hebben of krijgen van een depressie.

Figuur 1. Theorie dat vrouwen met een hoog en mannen met een laag testosterongehalte een grotere kans zouden hebben op het krijgen van een depressie.

Het figuur is de Nederlandse versie van *Figuur 1.5*. Links staat de verdeling van testosterongehalten bij vrouwen en rechts die van mannen. Het gebied tussen beide verdelingen verwijst naar testosterongehalten die hoog zijn voor vrouwen en laag voor mannen.

Hebben oudere vrouwen en mannen met een depressie én een hoog of laag testosterongehalte een grotere kans op het krijgen van het metabool syndroom?

Het metabool syndroom is een verzameling van klachten die voorkomen bij een gestoorde vet- en suikerstofwisseling. Patiënten hebben hierbij minimaal drie van de volgende kenmerken: een verhoogd cholesterolgehalte, hoge bloeddruk, overgewicht en een hoog bloedsuikergehalte. Het is bekend dat zowel hoge testosterongehaltes bij vrouwen en lage testosterongehaltes bij mannen, als het hebben van een depressie de kans op het krijgen van het metabool syndroom verhoogd. Mogelijk is deze kans extra verhoogd bij mensen met de combinatie van deze kenmerken, het hebben van een depressie én een afwijkend testosterongehalte. Om dit te onderzoeken is bij 478 oudere vrouwen en mannen gekeken of patiënten met beide kenmerken een grotere kans hadden op het metabool syndroom dan patiënten met één kenmerk (**hoofdstuk 5**). Dit bleek echter niet het geval te zijn. De kans op het metabool syndroom bij afwijkende testosterongehaltes was gelijk voor vrouwen en mannen met en zonder een depressie.

Doen vrouwen en mannen met een hoog testosterongehalte vaker een zelfmoordpoging?

Elke dag doen er in Nederland ongeveer 25 mensen een zelfmoordpoging waaraan een vijfde van deze mensen komt te overlijden. Zelfmoordgedachtes komen vaak voor bij patiënten met een depressie en het hebben van deze gedachtes is een belangrijke risicofactor voor het doen van een zelfmoordpoging. Toch blijft het lastig om te voorspellen wie een zelfmoordpoging zal doen. Daarom moeten andere risicofactoren voor het doen van een zelfmoordpoging beter in kaart worden gebracht.

Hoge testosterongehaltes maakt mensen mogelijk agressiever of impulsiever en deze kenmerken kunnen gemengde gevoelens over zelfmoord laten omslaan in het doen van een zelfmoordpoging. Daarom is meermaals onderzocht onderzocht of vrouwen en mannen met hoge testosterongehaltes vaker een zelfmoordpoging doen dan mensen met lagere testosterongehaltes. Sommige onderzoeken lieten dit inderdaad zien. Echter, deze onderzoeken hebben testosterongehaltes gemeten na de zelfmoordpoging. Dit is een probleem omdat het testosterongehalte dan ook verhoogd zou kunnen zijn door de enorme stress die wordt ervaren rondom een zelfmoordpoging. Daardoor is niet te bepalen of een hoog testosterongehalte de

kans op een zelfmoordpoging verhoogt, of omgekeerd, dat een zelfmoordpoging het testosterongehalte verhoogt. Daarom is bij 2.861 vrouwen en mannen gekeken of een eenmalig gemeten testosterongehalte een zelfmoordpoging in de negen jaar erna kon voorspellen (**hoofdstuk 6**). Dat bleek niet zo te zijn. De 93 deelnemers die een zelfmoordpoging deden, hadden vergelijkbare testosterongehaltes als de deelnemers die geen zelfmoordpoging deden.

Hebben vrouwen met meer hormoonstommelingen meer last van depressieve klachten en prikkelbaarheid?

De overgang is de periode in het leven van de vrouw rond de laatste menstruatie. Tijdens de overgang gaat het gehalte van het geslachtshormoon oestrogeen meer op en neer dan ervoor. Deze schommelingen en andere hormonale veranderingen veroorzaken opvliegers, maar mogelijk ook depressieve klachten en prikkelbaarheid. Daarom is bij 50 vrouwen in de overgang gekeken of de variabiliteit (de mate van schommelingen, gemeten van week tot week) in oestrogeen depressieve klachten (**hoofdstuk 7**) en prikkelbaarheid (**hoofdstuk 8**) kon voorspellen. Vrouwen met meer variabiliteit in oestrogeengehaltes bleken meer depressieve klachten maar minder prikkelbaarheid te ervaren dan vrouwen met minder variabiliteit in oestrogeen (zie *Figuur 2*).

Figuur 2. Samenhang tussen variabiliteit in oestrogeen en depressieve klachten en prikkelbaarheid bij vrouwen tijdens de overgang.

Het linker figuur is gebaseerd op *Figure 7.1 (links)* en het rechter figuur is gebaseerd *Figure 8.1 (linksboven)*. Beide figuren tonen het verband tussen de mate van variabiliteit van oestrogeen en de gemiddelde ernst van depressieve klachten (links) en prikkelbaarheid (rechts). Een hogere score op de verticale as betekent dat er meer klachten waren. Elke stip vertegenwoordigt één vrouw.

Betekenis resultaten uit het eerste deel van dit proefschrift voor patiënten met een depressie of depressieve klachten en voor toekomstig onderzoek op dit gebied

Samengevat, is er in dit eerste deel van het proefschrift geen verband gevonden tussen eenmalig gemeten geslachtshormonen en depressie. Wel is er een relatie gevonden tussen de variabiliteit in oestrogeen en stemmingsproblematiek bij vrouwen in de overgang. Deze bevindingen ondersteunen het belang van meerdere geslachtshormoonbepalingen in toekomstig onderzoek, omdat dat een veel beter beeld geeft van de dynamiek van het hormonale systeem. Dat is belangrijk omdat deze resultaten erop wijzen dat vrouwen in de overgang juist door schommelende oestrogeengehaltes mogelijk meer depressieve klachten hebben.

Om te bepalen of schommelende oestrogeengehaltes echt de oorzaak zijn van meer depressieve klachten zal een ander soort onderzoek moeten worden gedaan. Hierbij zouden vrouwen in de overgang door loting moeten worden verdeeld over een groep die hormoontherapie krijgt en een groep die een placebo krijgt (nepmedicijn). Er zijn al wel van dit soort onderzoeken uitgevoerd waaruit bleek dat hormoontherapie (oestrogeen via een pleister of een combinatie van oestrogeen en progesteron via een pil) depressieve klachten kon verminderen. Echter, aan deze studies nam maar een klein aantal vrouwen deel. Daarom zal het onderzoek herhaald moeten worden bij een grotere groep vrouwen.

Vraag 2: Kan pilgebruik depressieve klachten veroorzaken?

De pil bevat synthetische vormen van de geslachtshormonen oestrogeen en progesteron. Pilgebruik heeft veel voordelen. Het voorkomt ongewenste zwangerschappen, wat veel betekend heeft voor de emancipatie van de vrouw. Ook vermindert het jeugdpuistjes en menstruatiepijn. Toch kunnen er, zoals bij elk medicijn, ook bijwerkingen optreden. Een voorbeeld van een mogelijke bijwerking is het voorkomen van depressieve klachten. Hoe vaak deze klachten optreden en bij wie, en of pilgebruik deze klachten veroorzaakt, werd in dit proefschrift onderzocht.

Komen depressieve klachten vaker voor bij pilgebruiksters?

Eerst is onderzocht of depressieve klachten inderdaad vaker voorkomen bij pilgebruiksters dan bij niet-pilgebruiksters (**hoofdstuk 9**). Hiervoor zijn de gegevens onderzocht van 1.010 vrouwen die negen jaar waren gevolgd. Het bleek

dat de 16-jarigen die de pil gebruikten ongeveer 20% meer depressieve klachten hadden dan de 16-jarigen die de pil niet gebruikten. Dit verschil in klachten verdween toen deze vrouwen ouder werden. Op 19-, 22- en 25-jarige leeftijd hadden wel en niet-pilgebruiksters evenveel depressieve klachten (zie *Figuur 3*). De resultaten van dit onderzoek bespraken we een podcast¹ en beschreven in een review² voor Dokter Media.

Figuur 3. Depressieve klachten in pil en niet-pil-gebruiksters.

Het figuur is een simpelere weergave van *Figure 9.1*. De vierkanten geven de gemiddelde ernst van depressieve klachten met één standaardfout weer. De gegevens zijn gecorrigeerd voor de leeftijd, etniciteit en sociaaleconomische status van de vrouwen. Een hogere score op de verticale as betekent dat er meer depressieve klachten waren.

Hebben pilgebruiksters een grotere kans op een depressie later in het leven?

De vorige studie toonde dat tijdens pilgebruik de 16-jarigen meer depressieve klachten ervaren. Naar aanleiding van deze resultaten werd in een volgende studie onderzocht of ook in de jaren na het pilgebruik de kans op een depressie is verhoogd (**hoofdstuk 10**). Hiervoor zijn de gegevens van 725 vrouwen onderzocht. Van deze vrouwen was bekend of zij voor het 20^{ste} levensjaar de pil

¹ Maakt 'de pil' depressief? FactGurus Podcast op 29 oktober 2019. Te beluisteren via: <https://www.bnr.nl/podcast/factgurus/10393426/maakt-de-pil-depressief>

² "Meisjes van 16 die aan de pil zijn hebben vaker depressieve gevoelens". Dokter Media Review op 7 oktober 2019. Te lezen via: <https://doktermedia.nl/reviews/medicatie/meisjes-van-16-die-aan-de-pil-zijn-hebben-vaker-depressieve-gevoelens/>

hadden gebruikt en of zij in de zes jaren daarna een depressief waren. Het bleek dat pilgebruik tijdens de pubertijd de kans op het krijgen van een depressie voor het 25^{ste} levensjaar vergrootte.

Veroorzaakt pilgebruik depressieve klachten?

Met de twee vorige onderzoeken kon niet vastgesteld worden of pilgebruik de depressieve klachten veroorzaakte. Om een oorzakelijk verband vast te stellen, is namelijk een ander soort onderzoek nodig: een gerandomiseerde klinische studie. In zo'n onderzoek worden de vrouwen door loting verdeeld over een groep die de pil krijgt en een groep die een placebo (nepmedicijn) krijgt. Door de resultaten van dit soort vergelijkbare onderzoeken samen te vatten, kan met meer zekerheid worden bepaald of de pil depressieve klachten kan veroorzaken. Voor dit proefschrift is zo'n onderzoek uitgevoerd (**hoofdstuk 11**). Dit onderzoek bevatte de gegevens van 12 verschillende onderzoeken waaraan 5.833 vrouwen met een gemiddelde leeftijd van 27 jaar deelnamen. Door de gegevens van deze onderzoeken te vergelijken konden de effecten op depressieve klachten van acht verschillende soorten pillen worden vastgesteld en worden vergeleken met die van placebo. Geen van de onderzochte pillen bleek de depressieve klachten van de vrouwen te verergeren (zie *Figuur 4*).

Figuur 4. Invloed van de pil op depressieve klachten.

Het figuur is een simpelere weergave van *Figure 11.3*. Als de horizontale lijn door een vierkant (de onzekerheidsmarge) groter/kleiner is dan nul, vermindert/verergerd die pil depressieve klachten.

Kan pilgebruik bij sommige vrouwen depressieve klachten verminderen?

Hoewel sommige vrouwen mogelijk meer depressieve klachten ervaren bij pilgebruik, zijn er waarschijnlijk ook vrouwen die juist minder klachten hebben door pilgebruik. Bijvoorbeeld, vrouwen met het premenstrueel syndroom (PMS). Bij PMS ervaren vrouwen lichamelijke en/of emotionele klachten (zoals depressieve klachten of prikkelbaarheid) in de week voor de menstruatie. Ongeveer 12% van de vrouwen in de vruchtbare leeftijd heeft hier last van. De helft van deze vrouwen heeft vooral last van ernstige emotionele klachten, zoals depressieve klachten en prikkelbaarheid. Dit ziektebeeld noemen we een premenstrueel dysfore stoornis (PMDD). Waarschijnlijk worden de klachten bij deze vrouwen veroorzaakt door een hogere gevoeligheid voor natuurlijke hormoonstommelingen. Pilgebruik vermindert deze natuurlijke schommelingen, en daarmee mogelijk ook de klachten. Via dezelfde aanpak als in hoofdstuk 11 is daarom onderzocht of pilgebruik de depressieve klachten van deze vrouwen inderdaad kan verminderen (**hoofdstuk 12**). In dit onderzoek werden de gegevens van 1.205 vrouwen met een gemiddelde leeftijd van 32 jaar meegenomen. Pilgebruik bleek de depressieve klachten niet of bijna niet te verminderen (zie *Figuur 5*). Wel hadden de vrouwen minder andere emotionele en lichamelijke klachten.

Figuur 5. Invloed van pilgebruik op depressieve klachten in vrouwen met het premenstrueel syndroom of een premenstrueel dysfore stoornis.

Het figuur is een simpelere weergave van *Figuur 12.3*. Het figuur geeft per onderzoek weer wat de geschatte kans is op een bepaalde uitkomst van de pil (links op depressieve klachten, rechts op de ernst van alle klachten samen). Onderaan het figuur is de schatting van het globale effect op basis van de daarboven getoonde studies weergegeven.

Betekenis resultaten uit het tweede deel van dit proefschrift voor pilgebruiksters en patiënten met PMS of PMDD en voor toekomstig onderzoek op dit gebied

Samengevat, is er in het tweede deel van dit proefschrift geen wetenschappelijk bewijs gevonden dat pilgebruik depressieve klachten veroorzaakt. Twee onderzoeken lieten zien dat jonge pilgebruiksters meer depressieve klachten ervaren dan niet-pilgebruiksters, maar de oorzaak hiervan kon niet worden bepaald. In het onderzoek waarin dat wel kon, is er geen verband gevonden tussen pilgebruik en depressieve klachten. De gegevens voor dit laatste onderzoek kwamen echter van vrouwen die meestal al eens eerder de pil hadden gebruikt. Hierdoor zijn de resultaten van dit onderzoek en die van de eerdere twee onderzoeken moeilijk met elkaar te vergelijken. Om de vraag te beantwoorden of pilgebruik depressieve klachten kan veroorzaken bij jonge vrouwen of bij vrouwen die de pil nog niet eerder hebben gebruikt, is daarom meer onderzoek nodig. In ieder geval is het belangrijk dat vrouwen die na het starten van pilgebruik meer depressieve klachten ervaren, dit met hun arts bespreken. Samen kunnen zij bepalen of eventueel een ander type voorbehoedsmiddel gewenst is.

Ook bij vrouwen met PMS of PMDD bleek pilgebruik geen invloed te hebben op het aantal depressieve klachten. Met andere woorden, pilgebruik verminderde de depressieve klachten niet. Voor deze specifieke klachten zullen daarom bij deze vrouwen andere behandelopties overwogen moeten worden. Wel bleek dat pilgebruik andere klachten die vrouwen met PMS of PMDD vaak ervaren te verminderen. Vrouwen met PMS of PMDD kunnen dus wel baat hebben bij pilgebruik

Vraag 3: Wat kunnen wetenschappers doen om de wetenschapsjournalistiek te verbeteren?

Wetenschappers delen hun onderzoeksresultaten met collega's door deze te publiceren in wetenschappelijke tijdschriften. Soms zijn deze resultaten ook interessant voor een breder publiek. Omdat dit bredere publiek geen wetenschappelijke achtergrond heeft, moeten de resultaten worden "vertaald" in gewone begrijpelijke taal. Dit vertalen is een samenwerkingsproces tussen wetenschappers, wetenschapsvoorlichters en journalisten. Helaas gaat er bij dit vertalen, bewust of onbewust, nog wel eens iets mis. Zo is over de resultaten van hoofdstuk 9 door Radar geschreven dat "Meisjes aan de anticonceptiepil vaker

depressief" zijn. Dat klopt niet helemaal. Er werd namelijk gevonden dat 16-jarige pilgebruiksters meer depressieve klachten hadden, maar of ze ook vaker een depressie hadden, is niet onderzocht.

Wat wetenschappers kunnen doen om zo'n verkeerde uitleg van resultaten te voorkomen werd onderzocht in de laatste studie van dit proefschrift (**hoofdstuk 13**). Een advies dat daar uitkwam is dat wetenschappers persberichten beter moeten controleren op overdrijving of feitelijke onjuistheid. De persberichten vormen namelijk een belangrijke bron van informatie voor de journalisten. Bovendien liet een eerder onderzoek al zien dat de kans groot is dat fouten in een persbericht worden overgenomen in het mediabericht.

Conclusie

In dit proefschrift is het verband tussen geslachtshormonen en depressie onderzocht. Er bleek geen verband tussen eenmalig gemeten testosterongehaltes en depressie of zelfmoordpogingen. Wel bleek dat hormoonstommelingen bij vrouwen in de overgang mogelijk kunnen leiden tot meer depressieve klachten. Ook bleek dat jonge vrouwen gevoeliger zijn voor het krijgen van depressieve klachten bij pilgebruik, maar dat volwassen vrouwen over het algemeen geen depressieve klachten ervaren. Of de gevonden verbanden ook oorzakelijk van aard zijn, zal toekomstig onderzoek moeten uitwijzen. Deze vervolgonderzoeken gaan definitievere antwoorden geven op de vraag "Can hormones get you down?".

List of publications

Included in this thesis

de Wit AE, de Vries YA, de Boer MK, Scheper C, Fokkema AA, Janssen CAH, Giltay EJ, Schoevers RA. Efficacy of oral contraceptive pills for depressive symptoms and overall symptomatology in premenstrual syndrome: pairwise and network meta-analysis of randomized trials. *Am J Obstet Gynecol*. Jul 2021;doi:10.1016/j.ajog.2021.06.090

Anderl C*, **de Wit AE***, Giltay EJ, Oldehinkel AJ, Chen F. Association between adolescent oral contraceptive use and future major depressive disorder: a prospective cohort study. *Am J Psychiatry*. Jul 2021; doi:10.1111/jcpp.13476

*Shared first author

de Wit AE, de Vries YA, de Boer MK, Scheper C, Fokkema AA, Schoevers RA, Giltay EJ. Hormonal contraceptive use and depressive symptoms: systematic review and network meta-analysis of randomized trials. *BJPsych Open*. Jun 2021;7(4):e110. doi:10.1192/bjo.2021.64

de Wit AE, Giltay EJ, de Boer MK, Bosker FJ, Nolen WA, Penninx BWJH, Schoevers RA. Plasma androgens and the presence and course of depression in a large cohort of men. *Psychoneuroendocrinology*. May 2021;130:105278. doi:10.1016/j.psyneuen.2021.105278

de Wit AE, Giltay EJ, de Boer MK, Bosker FJ, Cohn AY, Nolen WA, Kaiser UB, Joffe H, Penninx BWJH, Schoevers RA. Plasma androgens and the presence and course of depression in a large cohort of women. *Transl Psychiatry*. Feb 2021;11(1):124. doi:10.1038/s41398-021-01249-2

de Wit AE, Giltay EJ, de Boer MK, Nathan M, Wiley A, Crawford S, Joffe H. Predictors of irritability symptoms in mildly depressed perimenopausal women. *Psychoneuroendocrinology*. Jan 2021;126:105128. doi:10.1016/j.psyneuen.2021.105128

de Wit AE, De Boer MK, Bosker FJ, van der Does WAJ, Gooren LJG, Nolen WA, Penninx BWJH, Schoevers RA, Giltay EJ. Associations of plasma androgens with suicidality among men and women: A 9-year longitudinal cohort study. *J Affect Disord*. May 2020;269:78-84. doi:10.1016/j.jad.2020.03.032

Joffe H, **de Wit AE**, Coborn J, Crawford S, Freeman M, Wiley A, Athappilly G, Kim S, Sullivan KA, Cohen LS, Hall JE. Impact of Estradiol Variability and Progesterone on Mood in Perimenopausal Women With Depressive Symptoms. *J Clin Endocrinol Metab*. Mar 2020;105(3) doi:10.1210/clinem/dgz181.

de Wit AE, Booij SH, Giltay EJ, Joffe H, Schoevers RA, Oldehinkel AJ. Association of Use of Oral Contraceptives With Depressive Symptoms Among Adolescents and Young Women. *JAMA Psychiatry*. Oct 2019; doi:10.1001/jamapsychiatry.2019.2838

de Wit AE, Giltay EJ, de Boer MK, Bosker FJ, van der Mast RC, Comijs HC, Oude Voshaar RC, Schoevers RA. Associations between testosterone and metabolic syndrome in depressed and non-depressed older men and women. *Int J Geriatr Psychiatry*. 03 2019;34(3):463-471. doi:10.1002/gps.5040

de Wit AE, Bosker FJ, Giltay EJ, de Kloet CS, Roelofs K, van Pelt J, Penninx BWJH, Schoevers RA. Testosterone in human studies: Modest associations between plasma and salivary measurements. *Andrologia*. 2017; doi:10.1111/and.12779

Additional publications

Jessurun CAC, Hulsbergen AFC, **de Wit AE**, Tewarie IA, Snijders TJ, Verhoeff J, Phillips J, Reardon DA, Mekary RA, Broekman MLD. The combined use of steroids and immune checkpoint inhibitors in brain metastasis patients: a systematic review and meta-analysis. *Neuro Oncol*. Feb 2021; doi:10.1093/neuonc/noab046

Submitted for publication

Morssinkhof MWL, Lamers F, Hoogendoorn AW, **de Wit AE**, Riese H, Giltay EJ, van den Heuvel OA, Penninx BWJH, Broekman BFP. Oral contraceptives, insomnia and depressive symptoms in women with and without depression.

Acknowledgements / Dankwoord

Mijn proefschrift is klaar. Ik kijk terug op een tijd waarin ik, naast veel te hebben gelachen, ook veel heb geleerd en de kans heb gekregen het maximale uit mezelf te halen. Dit alles kon niet zonder de hulp van velen.

Allereerst wil ik een aantal mensen bedanken die op een directe manier hebben bijgedragen aan dit proefschrift:

Mijn promotor en copromotoren: **Robert, Erik en Marrit**. **Robert**, jij riep meerdere malen dat geen zee voor mij te hoog is, maar dat geldt evengoed voor jou. Proberen kan altijd, toch? Voor dit proberen kreeg ik vrijheid en vertrouwen, en daar ben ik je ontzettend dankbaar voor. **Erik**, jouw enthousiasme voor wetenschap is onuitputtelijk. Dit, in combinatie met je bescheidenheid en vakinhoudelijke kennis, maakt je een fantastische begeleider. **Marrit**, jij was er altijd om, naast de inhoud van mijn proefschrift, de minstens zo belangrijke zaken buiten de wetenschap te bespreken. Jij bent ontzettend sterk in het meedenken, relativeren en het geven van concrete adviezen. Samen waren jullie een gouden team, waardoor ik me altijd gesteund heb gevoeld.

De persoon die elke promovendus in zijn team zou moeten hebben: **Fokko**. Dit proefschrift startte op de zevende verdieping van het UCP. Jij toverde een asbak uit een labdoosje en al rokende (met mijn toestemming) vertelde jij hoe je een 5-HT receptor had geïsoleerd. Wat ik daar vooral van begreep, was dat jij, als biochemicus en enthousiasteling, onmisbaar zou worden voor dit project. Jij moedigde mij aan om mijn eigen weg te volgen en dat was het beste advies dat je me kon geven.

My mentor from Boston: **Hadine**. I am ever grateful for the opportunity you gave me to obtain my “Been to America” degree. I admire your enormous knowledge on both the research and clinical aspects of our field. Your attention to detail regarding the writing has been an eye-opener for me and pushed our papers to a next level.

My inspiring co-authors: **Christine, Jamie, Sanne, Tineke** and **Ymkje Anna**. **Christine**, you initiated our collaboration which resulted in a paper that has been a huge acquisition of knowledge and skills for me. **Jamie**, you have been - and still are - very supportive when I needed help, which made you the best colleague to share my office with during my time in Boston. **Sanne**, jij bent een kei in het je vastbijten in een probleem, maar bovenal waardeer ik je oog voor het welzijn van de mensen om je heen. **Tineke**, jouw opmerkingen slaan altijd de spijker op z'n kop. **Ymkje Anna**, voor jou is geen probleem een echt probleem, want je

weet altijd wel een creatieve oplossing te bedenken. Kortom, **Christine, Jamie, Sanne, Tineke** en **Ymkje Anna**: zonder jullie hulp waren de hoofdstukken uit dit proefschrift nooit zo goed geworden. Mijn dank is groot!

Collega's van mijn onderzoeksthuisbasis: **de afdeling ICPE/UCP van het UMCG**. Jullie toewijding, ondersteuning en kennis vormden een onuitputtelijke inspiratiebron voor dit proefschrift. Met naam wil ik nog de hulp van **Paulien, Gerda** en **Lynn** noemen. **Paulien**, jij bent een onmisbare schakel gebleken in de voortgang van dit project. Vele zaken waren binnen vijf minuten geregeld als ik je om hulp vroeg bij het plannen van afspraken en het rondkrijgen van administratieve zaken. **Gerda**, jouw vastberadenheid en optimisme hebben de NESDA bio-opslag in Groningen weer op de rit gekregen en me enorm geholpen tijdens de androgeenbepaling in het lab. Als laatste, **Lynn**, jij wierp je op als mijn stagebegeleider zodat ik met de beste begeleiding mijn eerste stappen in de wetenschap kon zetten.

Collega's van mijn tijdelijke onderzoeksplekken: **de kantoortuin van de afdeling Psychiatrie van het LUMC, the WHARP group from Harvard Medical School**, en **het lab Bijzondere Chemie van het UMCG**. Allereerst mijn **kantoortuingenoten**, bedankt voor jullie gezelligheid en de koffies op de vrijdagen. My colleagues from **WHARP**, thank you so much for making me feel at home in Boston. I feel lucky to have had the chance to be part of the crew. Especially **Aleta**, thank you for introducing me into Boston's bike scene. Eveneens **Hanneke**, jij bent met je advies vanaf de zijlijn een enorme steun voor me geweest tijdens mijn tijd in Boston. Als laatste, **Claude, Henk, Twan, Hillie** en **Martijn**, bedankt voor het mij wegwijs maken in het lab en het mij enthousiasmeren voor de LC-MS/MS machine!

Mijn opleider Epidemiologie: **prof. dr. Truuske de Bock**. Bedankt voor de begeleiding tijdens mijn epidemiologie opleiding. De opgedane kennis is van enorme waarde geweest voor dit proefschrift.

Deelnemers aan de studies **NESDA, NESDO, TRAILS, PERI** en alle gerandomiseerde klinische studies die zijn gebruikt voor de netwerk meta-analyses, zonder jullie inzet was er geen proefschrift geweest.

Tot slot, leden van de leescommissie: **prof. dr. Daan Touw, prof. dr. Liesbeth van Rossum**, en **prof. dr. Nic van der Wee**. Bedankt voor de tijd en moeite die jullie hebben genomen om mijn proefschrift te lezen en te beoordelen.

Daarnaast wil ik ook graag de vele mensen bedanken die indirect belangrijk zijn geweest in de totstandkoming van dit proefschrift.

De **key population van Stichting New Beginnings/Chances for life**. Jullie ontvingen mij met open armen toen ik mij wilde verdiepen in de problematiek van transgenders in Suriname. Het contact met jullie gaf mij waardevolle inzichten in de impact van geslachtshormonen op het lichaam. Maar belangrijker, jullie leerden mij hoe vreselijk groot de kansenongelijkheid in deze wereld is. Het onrecht waartegen jullie moeten vechten is immens. Ik hoop met heel mijn hart op een betere acceptatie en emancipatie van transgenders. Jullie verdienen het meer dan eenieder ander.

Mijn mede-PhDs uit het ICPE: **Arnout, Fionneke, Marieke H, Marieke S, Robin en Yoram**. We gingen meerdere keren samen op vakantie en ondanks de lockdown spreken we elkaar dagelijks. Dit maakt dat “vriendschap” een betere omschrijving is dan “collega’s” voor de band die ik voel met jullie. Ik ben dan ook blij dat we binnenkort de laatste term kunnen laten vallen en alleen de “vriendschap” overblijft. Speciaal wil ik nog mijn dank benoemen aan mijn paranimf **Robin**. Jouw beperkte zelfmedelijden en liefde voor sport en feestjes maken je de ideale metgezel voor het ondernemen van activiteiten. Ik prijs mezelf dan ook gelukkig dat we, vijf jaar na start van onze promotietrajecten, weer in dezelfde stad wonen.

Mijn mede-MD/PhD-ers: **Fabian, Irene, Joylene, Mathilda en Stan**. Jullie waren altijd beschikbaar voor koffie, support en gezelligheid, waarvoor mijn grote dank!

Mijn collega’s bij Dokter Media: **Tijs, Lester en Donica**. Ik vind het heerlijk dat ik bij jullie alle ruimte krijg om ideeën uit te proberen en om mijn academische onderzoeksresultaten te vertalen naar een concreet advies voor de praktijk. Op een mooie toekomst voor Dokter Media en nog vele gezellige borrels.

Mijn vrienden: **Harmen, Ineke, Jolien, Mariëlle, Nick, Yldau en Veerle**. **Ineke**, de MOLM-13 cellijn blijkt onze gouden basis voor een vriendschap, want sinds onze JSM-labstage ben jij mijn steun en toeverlaat geweest voor dit proefschrift. Jouw huis werd over de jaren mijn thuis in Groningen. Samen met **Nick**, ben jij mijn Groningse rots in de branding waar ik altijd op kan rekenen. Bedankt dat je op de dag van de verdediging naast me wil staan als paranimf! **Jolien**, onze gedachtestromen zitten vaak op dezelfde golflengte en dat is het ingrediënt voor geweldige uitspattingen. Ik kan niet wachten tot de volgende. **Mariëlle**, ik begrijp nog steeds niet waarom wij ooit samen in een roeiboot zijn begonnen, maar gelukkig zitten we tegenwoordig op het droge. Je eerlijkheid en betrouwbaarheid maken je een geweldige vriendin. **Yldau**, jij liep altijd een

paar maanden op me voor en was daarmee de ideale vraagbaak. Maar bovenal, je nuchterheid en gezelligheid maken me trots om je mijn vriendin te mogen noemen. **Veerle**, huisgenoot uit Boston en inmiddels buurvrouw in Den Haag, gastvrijheid kent bij jou en **Harmen** geen grenzen. Bij jullie voel ik me altijd thuis. Aan jullie allen, dank voor de feestjes, vakanties en logeerpartijtjes, hier heb ik ontzettend van genoten. Ik voel me gelukkig jullie tot mijn vrienden te mogen rekenen.

Mijn **huisgenoten** uit Groningen, Hardenberg, Zwolle, Paramaribo, Boston en Leiden. Mijn studententijd kreeg de gewenste start in Huize 19-pand-a aan de Groningse gracht, maar ook later had ik altijd het geluk om in geweldige huizen terecht te komen waar vooral veel aandacht was voor het leven buiten het ziekenhuis. De groep is te groot om jullie allen bij naam te noemen, maar ik ben jullie eeuwige dank verschuldigd. Speciale dank wil ik uiten naar **Wijnanda** en **Max**, mijn schoonouders en senior-huisgenoten, wiens deur altijd voor me openstond als ik weer een coschap kwam lopen in de buurt.

Mijn zussen: **Mirthe** en **Sterre**. Als super trotse zus zie ik hoe jullie elk je eigen weg vinden. Het leeftijdsverschil wordt met de jaren verwaarloosbaar en dat vind ik ontzettend leuk. **Mirthe**, achter “attent” in de Dikke Van Dale zou jouw naam moeten staan. Er ontgaat jou werkelijk niets. **Sterre**, jouw grenzeloze creativiteit en vurigheid zorgen voor veel vertier en hartstochtelijke discussies. Ik kijk uit naar de toekomst waarin we nog vele dingen samen mogen doen.

Mijn vriend: **Arnold**. Voor jou was dit promotietraject een onderbetaalde abstracte aaneenschakeling van medisch geneuzel en politieke spelletjes. Maar sinds wij elkaar zeven jaar geleden vroeg in de ochtend in ‘t Vaatje troffen, ben jij er altijd voor mij geweest. Het had wat tijd en omzwervingen via Hardenberg, Boston, Leiden, Paramaribo en zelfs je ouders in Zwolle nodig, voordat ik mijn weg naar Amsterdam vond. Het feit dat onze relatie al deze afstanden, mijn promotieonderzoek en jouw projecten zoals de overname van HEMA en Keesing Media Group tijdens de lockdown heeft overleefd, sterkt mij in de gedachte: onze liefde is onvoorwaardelijk.

De aanstichters van dit project: **mijn ouders**. Jullie doctoraatsbullen hingen thuis in de hal waardoor ik al jong wist te melden dat ik later “universiteit” wilde worden. Dat laatste is niet gelukt, maar de bul is onderweg. Jullie zette mij op deze wereld met de overtuiging dat hard werken wordt beloond en dat je uit moet gaan van je eigen kracht. Dit zijn twee bijzonder belangrijke overtuigingen gebleken

voor het doen van dit promotieproject waarvoor mijn enorme dank. Nog meer wil ik jullie bedanken voor jullie onvoorwaardelijke steun en geloof in mijn kunnen. Betere ouders kan ik me niet wensen.

Anouk

Mei 2021

Curriculum Vitae

Anouk de Wit (Zeist, 1993) started studying Medicine at the University of Groningen in 2011. During the bachelor phase of her studies she followed the Junior Scientific Master program which sparked her enthusiasm for science. After completing this honors program, she successfully applied for a personal MD PhD grant. This enabled her to pursue her studies of Medicine and to follow a PhD trajectory including the Epidemiologist B program at the department of Psychiatry of the University Medical Center Groningen under the supervision of professor Robert Schoevers.

Between 2016 and 2019, de Wit alternated clinical rotations at different hospitals for her master studies in Medicine with months of research for her PhD project. She moved out of the Netherlands twice. First, she visited Paramaribo for two months to study the access to healthcare for transgender people in Suriname, a project for which she was awarded the Groninger Universiteitsfonds Talent Award. Thereafter, she moved to Boston (United States) for half a year to collaborate with Hadine Joffe (Department of Psychiatry, Harvard Medical School). Together, they worked on the link between sex steroid dynamics and mood deterioration during the menopause transition. This work was nominated for the Greenblatt Award. Upon her return, de Wit's work on the association between oral contraceptive pill use and depressive symptoms (JAMA Psychiatry, 2019) received significant media attention. Because some of the media misinterpreted the findings, she became aware of how important it is for correctly disseminating research findings to the public that researchers, public information officers and journalists understand each other. This motivated her to find ways to create more awareness among researchers of their role and responsibility in science journalism. To this end, she started working as an editor at Dokter Media - a platform that aims to improve the transmission of medical news - and she wrote an article on this theme, which was awarded with the Open Research Award.

With a final clinical internship at the department of Neurosurgery at the Leiden University Medical Centre, de Wit completed her master studies in Medicine in 2020. Her PhD project resulted in presentations at several (international) conferences, multiple papers, and eventually in this thesis. With the completion of her thesis, de Wit will also finalize the Epidemiologist B program. She currently works as a doctor at the department of Neurosurgery at the Haaglanden Medical Center and she is still an editor at Dokter Media.