Cyprus University of Technology

Faculty of Health Sciences

Doctoral Dissertation

THE USE OF META-RESEARCH TO EVALUATE THE ROBUSTNESS OF EVIDENCE IN PERINATAL EPIDEMIOLOGY

Konstantinos Giannakou

Limassol, July 2018

CYPRUS UNIVERSITY OF TECHNOLOGY

FACULTY OF HEALTH SCIENCE

CYPRUS INTERNATIONAL INSTITUTE OF ENVIRONMENTAL AND PUBLIC HEALTH

Doctoral Dissertation

THE USE OF META-RESEARCH TO EVALUATE THE ROBUSTNESS OF EVIDENCE IN PERINATAL EPIDEMIOLOGY

Konstantinos Giannakou

July 2018

Doctoral Dissertation

THE USE OF META-RESEARCH TO EVALUATE THE ROBUSTNESS OF EVIDENCE IN PERINATAL EPIDEMIOLOGY

Presented by

Konstantinos Giannakou

Supervisor: Stefania I. Papatheodorou, Assistant Professor

Signature _____

Member of the committee: Evridiki Papastavrou, Associate Professor

Signature _____

Member of the committee: Zacharias Zachariou, Professor

Signature _____

Cyprus University of Technology

Limassol, July 2018

Copyrights

Acknowledgements

"As long as the centuries continue to unfold, the number of books will grow continually, and one can predict that a time will come when it will be almost as difficult to learn anything from books as from the direct study of the whole universe. It will be almost as convenient to search for some bit of truth concealed in nature as it will be to find it hidden away in an immense multitude of bound volumes." — Denis Diderot, "Encyclopédie" [1755]

This PhD thesis could not have been completed without the assistance, guidance and kind support from many people. My deepest gratitude goes first to my supervisor Dr Stefania Papatheodorou, who showed constant motivation and understanding, provided me with unstinting support and expertly guided me throughout this work. Thank you for encouraging my research and for allowing me to grow as a research scientist.

I am greatly indebted to the members of my research committee, Dr Panayiotis Yiallouros, Dr Nicos Middleton, and Dr Costas Christophi for investing time and providing interesting and valuable feedback. I feel proud and honoured that you have accepted to be on my committee. Moreover, I want to thank the members of the examining committee, Dr Zacharias Zachariou and Dr Evridiki Papastavrou for reading the manuscript of this thesis and for your recommendations to improve it.

My appreciation also extends to my fellow PhD and master's students at Cyprus International Institute, for their valuable help and inspiration. I want to especially thank my colleague, Dr Despo Pampaka. Your presence was very important in a process that is often felt as tremendously solitaire. Finally, I want to thank my friends and my family for their constant encouragement and support these past 4 years.

ABSTRACT

Introduction

The volume of literature pertinent to healthcare is growing at an increasing rate with nearly one million articles on research involving human subjects are published each year. With the ever-increasing of published studies, scientists turn into systematic reviews and meta-analyses to summarize the evidence, using multiple related studies for a single research question. There are tens of thousands of systematic reviews already published, but their production is still increasing at a phenomenal rate. Even though systematic reviews have become a very popular type of research study that increased the scientific knowledge and inform clinical and policy decision making, their credibility is under threat as most appear to be either not useful or of uncertain utility. The problem is that the majority are unnecessary, inaccurate or misleading due to biases in the methodology and selective reporting of results, or they address questions that have no clinical value. The increase in the number of systematic reviews, along with escalating demand from policy makers for rapid reviews of research, has emerged an evolving scientific discipline, and a newer form of synthesis, umbrella reviews. An umbrella review is a new method that provide a comprehensive assessment of the body of information that is available on a given topic using the evidence from multiple systematic reviews and meta-analyses. This assessment is fundamental not only for understanding the reliability of an evidence-base but also serves as the foundation for clinical and public health recommendations.

Aims

Towards further expand the mapping and the critical evaluation of research evidence across published literature of clinical identities with a large impact on the perinatal

epidemiology field, we aimed first to systematically appraise the evidence across published systematic reviews and meta-analyses on the risk factors and/or interventions for preeclampsia and gestational diabetes, and second identify whether any interventions or fields of risk factors include epidemiological credible evidence.

Methods

In three separate umbrella reviews, all major electronic databases were searched using appropriate terms towards identifying eligible systematic reviews and meta-analyses examining associations between risk or protective factors for preeclampsia and gestational diabetes, respectively, and pharmacologic and non-pharmacologic interventions for preeclampsia prevention. For each meta-analysis we estimated the summary effect size by random-effects and fixed-effects models, the 95% confidence interval and the 95% prediction interval. We also assessed the between-study heterogeneity expressed by I^2, evidence of small-study effects (large studies had significantly more conservative results than smaller studies) and evidence of excess significance bias (too many studies with statistically significant results). We further applied standardized methodological criteria to evaluate the epidemiological credibility of the statistically significant associations.

Results

Fifty-eight eligible meta-analyses of observational studies were identified providing data on 130 putative risk factors associated with preeclampsia. Sixty-five (50%) associations had nominally statistically significant findings at $P<0.05$, while sixteen (12%) were significant at $P<10^{-6}$. Sixty-five (50%) associations had large or very large heterogeneity. Evidence for small-study effects and excess significance bias was found in ten (8%) and twenty-six (20%) associations, respectively. Oocyte donation vs

spontaneous conception was the only non-genetic risk factor that presented convincing evidence for an association with preeclampsia. Across the statistically significant genetic risk factors (P<0.05), only PAI-1 4G/5G polymorphism was supported with strong evidence for a contribution to the pathogenesis of preeclampsia.

Twenty-nine eligible meta-analyses of randomized controlled trials were identified, providing data on 57 pharmacologic and non-pharmacologic interventions for preeclampsia prevention. Twenty-four (42%) associations had nominally statistically significant findings at P<0.05, while only 10 (18%) were significant at $P<10^{-3}$ under the random-effects model. Sixteen (28%) associations had large or very large heterogeneity. Evidence of excess significance bias was found in 15 (26%) associations. After applying our classification criteria, the following three interventions were classified as Class I level of evidence including low dose aspirin ≤16 weeks of gestation for preterm preeclampsia, diet and nutrition counselling and dietary interventions.

Twenty-one eligible meta-analyses of observational studies were identified, providing data on 43 putative risk factors associated with gestational diabetes mellitus (GDM). Thirty-eight (88%) associations had nominally statistically significant findings at P<0.05, while only 14 (32%) were significant at $P<10^{-6}$ under the random-effects model. Eighteen (42%) associations had large or very large heterogeneity. Evidence for small-study effects and excess significance bias was found in three (7%) and four (9%) associations, respectively. Only five risk factors presented convincing evidence for an association with GDM: vitamin D deficiency, low vs. normal BMI (cohort studies), BMI ~30-35 kg/m^2 vs. normal BMI, BMI >35 kg/m^2 vs. normal BMI, and hypothyroidism.

Conclusions

The results from this PhD thesis suggest that the evidence in the field of risk factors or interventions for preeclampsia and GDM, suffers from the presence of large between-study heterogeneity and statistical biases, that threat their validity and hind the identification of robust risk factors or interventions. Although a large proportion of meta-analyses reported nominally statistically significant associations, only a minority of these associations provided convincing evidence without indications of bias. Oocyte donation vs spontaneous conception and PAI-1 4G/5G polymorphism (recessive model) show the strongest consistent evidence for a contribution to the pathogenesis of preeclampsia. Vitamin D deficiency, low vs. normal BMI, moderately and severely obese vs. normal weight, and hypothyroidism show the strongest consistent evidence for GDM development. These risk factors represent a starting point for further etiopathological research, improvement of the prediction of preeclampsia and GDM, and identification of the women at high risk. From the available interventions for preeclampsia prevention, early administration of low dose aspirin in women with preterm preeclampsia, diet and nutrition counselling and dietary interventions had the strongest epidemiologic evidence suggesting their effectiveness. We believe this evaluation of research evidence that includes a robust hierarchical classification of the published evidence and its interpretation can be used to inform decision-making to support clinicians, public health professionals, regulatory officials, and policymakers.

ΠΕΡΙΛΗΨΗ

Εισαγωγή

Ο όγκος της βιβλιογραφίας που σχετίζεται με την υγεία αυξάνεται με εκπληκτικό ρυθμό, με περίπου ένα εκατομμύριο επιδημιολογικά άρθρα για τον άνθρωπο να δημοσιεύονται κάθε χρόνο. Λόγω των αυξανόμενων δημοσιευμένων μελετών, οι επιστήμονες στέφονται στις συστηματικές ανασκοπήσεις και μετα-αναλύσεις για να συνοψίσουν τα δεδομένα, χρησιμοποιώντας πολλαπλές σχετικές μελέτες για μια συγκεκριμένη ερευνητική ερώτηση. Μέχρι σήμερα υπάρχουν δεκάδες χιλιάδες συστηματικές ανασκοπήσεις. Παρ' όλα αυτά η παραγωγή τους εξακολουθεί να αυξάνεται με εκπληκτικό ρυθμό. Αν και θεωρούνται ένα πολύ δημοφιλές είδος ερευνητικής μελέτης που αύξησε την επιστημονική γνώση και συνέβαλε στη λήψη κλινικών και πολιτικών αποφάσεων, η αξιοπιστία τους διακυβεύεται καθώς η πλειονότητα αυτών εμφανίζεται να είναι είτε μη χρήσιμη, είτε ασαφής. Το πρόβλημα είναι ότι η πλειοψηφία των συστηματικών ανασκοπήσεων είναι πλεονάζουσα, ανακριβής ή παραπλανητική εξαιτίας των μεροληψιών στη μεθοδολογία και της επιλεκτικής αναφοράς των αποτελεσμάτων ή επειδή εξετάζουν πεδία που δεν έχουν κλινική σημασία. Η αύξηση του αριθμού των συστηματικών ανασκοπήσεων, καθώς και η μεγάλη ζήτηση ερευνητικών ανασκοπήσεων από τους υπεύθυνους χάραξης πολιτικής, έχουν αναδείξει ένα εξελισσόμενο επιστημονικό κλάδο, και μια νεότερη μορφή σύνθεσης της βιβλιογραφίας, «umbrella reviews». Αυτή η νέα μέθοδος παρέχει μια ολοκληρωμένη αξιολόγηση του συνόλου των πληροφοριών που είναι διαθέσιμα για ένα συγκεκριμένο θέμα, χρησιμοποιώντας δεδομένα από πολλαπλές συστηματικές ανασκοπήσεις και μετα-αναλύσεις. Η αξιολόγηση αυτή είναι θεμελιώδους σημασίας

όχι μόνο για την κατανόηση της αξιοπιστίας μιας βάσης δεδομένων, αλλά και ως βάση για συστάσεις που αφορούν τη δημόσια υγεία.

Στόχοι

Προκειμένου να διευρυνθεί περαιτέρω η χαρτογράφηση και η κριτική αξιολόγηση δημοσιευμένων ερευνητικών στοιχείων σε κλινικά πεδία με μεγάλη επίδραση στην περιγεννητική επιδημιολογία, η παρούσα εργασία στοχεύει, πρώτον στη συστηματική αξιολόγηση των στοιχείων από συστηματικές ανασκοπήσεις και μετα-αναλύσεις που εξετάζουν συσχετίσεις μεταξύ παραγόντων κινδύνου και παρεμβάσεων για την προεκλαμψία και τον διαβήτη κύησης, και δεύτερον, να προσδιορίσει την επιδημιολογική εγκυρότητα των προτεινόμενων παρεμβάσεων ή των πεδίων παραγόντων κινδύνου.

Μέθοδοι

Σε τρεις ξεχωριστές ανασκοπήσεις (umbrella reviews), όλες οι σημαντικές ηλεκτρονικές βάσεις δεδομένων έχουν ερευνηθεί με τη χρήση κατάλληλων όρων έτσι ώστε να εντοπιστούν οι συστηματικές ανασκοπήσεις και μετα-αναλύσεις που αναφέρουν συσχετίσεις μεταξύ παραγόντων κινδύνου για την προεκλαμψία και τον διαβήτη κύησης, και φαρμακολογικών και μη φαρμακολογικών παρεμβάσεων για την πρόληψη της προεκλαμψίας, αντίστοιχα. Για κάθε μετα-ανάλυση εκτιμήσαμε το μέγεθος της επίδρασης της περίληψης του αποτελέσματος, τα 95% διαστήματα εμπιστοσύνης και τα 95%. διάστημα πρόβλεψης. Εκτιμήσαμε επίσης την ετερογένεια μεταξύ των μελετών που εκφράζεται από το I^2, ενδείξεις επιδράσεων λόγω μικρής μελέτης (μεγάλες μελέτες είχαν στατιστικά σημαντικά πιο συντηρητικά αποτελέσματα σε σχέση με μικρότερες μελέτες) και υπέρμετρης μεροληψίας (συστηματικά

σφάλματα). Τυποποιημένα μεθοδολογικά κριτήρια εφαρμόστηκαν έτσι ώστε να αξιολογηθεί η επιδημιολογική εγκυρότητα των στατιστικά σημαντικών στοιχείων.

Αποτελέσματα

Πενήντα οκτώ μετα-αναλύσεις μελετών παρατήρησης εντοπίστηκαν, παρέχοντας δεδομένα για 130 υποτιθέμενους παράγοντες κινδύνου σε σχέση με την προεκλαμψία. Εξήντα πέντε (50%) συσχετίσεις είχαν στατιστικά σημαντικά ευρήματα ($P<0.05$), ενώ μόνο δεκαέξι (12%) ήταν στατιστικά σημαντικές σε επίπεδο σημαντικότητας $P<10^{-6}$. Εξήντα πέντε (50%) συσχετίσεις είχαν μεγάλη ή πολύ μεγάλη ετερογένεια. Ενδείξεις επιδράσεων λόγω μικρής μελέτης και υπέρμετρης μεροληψίας εντοπίστηκαν σε δέκα (8%) και είκοσι έξι (20%) συσχετίσεις, αντίστοιχα. Η δωρεά ωοκυττάρων έναντι της φυσιολογικής σύλληψης ήταν ο μόνος μη-γενετικός παράγοντας κινδύνου που παρουσίασε πειστικές αποδείξεις σε σχέση με την προεκλαμψία. Ανάμεσα στους στατιστικά σημαντικούς γενετικούς παράγοντες κινδύνου ($P<0.05$), μόνο ο πολυμορφισμός PAI-1 4G/5G (recessive model) παρουσίασε ισχυρές ενδείξεις για συμβολή στην παθογένεση της προεκλαμψίας.

Είκοσι εννέα μετα-αναλύσεις τυχαιοποιημένων κλινικών δοκιμών εντοπίστηκαν, παρέχοντας δεδομένα για 57 φαρμακολογικές και μη φαρμακολογικές παρεμβάσεις για την πρόληψη της προεκλαμψίας. Είκοσι τέσσερις (42%) παρεμβάσεις είχαν στατιστικά σημαντικά ευρήματα ($P<0.05$), ενώ μόνο 10 (18%) ήταν στατιστικά σημαντικά στο επίπεδο σημαντικότητας $P<10^{-3}$. Δεκαέξι (28%) συσχετίσεις είχαν μεγάλη ή πολύ μεγάλη ετερογένεια. Ενδείξεις υπέρμετρης μεροληψίας εντοπίστηκαν σε 15 (26%) παρεμβάσεις. Μετά την εφαρμογή των κριτηρίων ταξινόμησης, τρεις παρεμβάσεις ταξινομήθηκαν στην "Κατηγορία Ι" σύμφωνα με την επιδημιολογική

τους εγκυρότητα: χαμηλή δόση ασπιρίνης <16 εβδομάδες κύησης για πρόωρη προεκλαμψία, διαιτητική συμβουλευτική και διαιτητικές παρεμβάσεις.

Είκοσι μία μετα-αναλύσεις μελετών παρατήρησης εντοπίστηκαν, παρέχοντας δεδομένα για 43 υποτιθέμενους παράγοντες κινδύνου που σχετίζονται με τον διαβήτη κύησης. Τριάντα οκτώ (88%) παράγοντες κινδύνου είχαν στατιστικά σημαντικά ευρήματα ($P<0.05$), ενώ μόνο 14 (32%) ήταν στατιστικά σημαντικοί σε επίπεδο σημαντικότητας $P<10^{-6}$. Δεκαοκτώ (42%) παράγοντες κινδύνου είχαν μεγάλη ή πολύ μεγάλη ετερογένεια. Ενδείξεις επιδράσεων λόγω μικρής μελέτης και υπέρμετρης μεροληψίας εντοπίστηκαν σε τρείς (7%) και τέσσερις (9%) συσχετίσεις, αντίστοιχα. Μόνο πέντε παράγοντες κινδύνου παρουσίασαν πειστικές αποδείξεις για συσχέτιση με τον διαβήτη κύησης: ανεπάρκεια βιταμίνης D, χαμηλό σε σχέση με φυσιολογικό ΔΜΣ (μελέτες κοόρτης), ΔΜΣ ~ 30-35 kg/m^2 έναντι κανονικού ΔΜΣ, ΔΜΣ > 35 kg/m^2 έναντι φυσιολογικού, και υποθυρεοειδισμός.

Συμπεράσματα

Τα αποτελέσματα αυτής της διδακτορική διατριβής υποδηλώνουν ότι τα ερευνητικά στοιχεία στον τομέα των παραγόντων κινδύνου ή παρεμβάσεων για την προεκλαμψία και τον διαβήτη κύησης πάσχουν από την ύπαρξη μεγάλης ετερογένειας μεταξύ των μελετών, όπως και στατιστικών σφαλμάτων που απειλούν την εγκυρότητά τους και εμποδίζουν τον εντοπισμό ισχυρών παραγόντων κινδύνου ή παρεμβάσεων. Αν και σε ένα μεγάλο ποσοστό των μετα-αναλύσεων εντοπίστηκαν στατιστικά σημαντικές συσχετίσεις, μόνο η μειοψηφία αυτών ήταν πειστικές χωρίς ενδείξεις προκατάληψης. Η δωρεά ωοκυττάρων έναντι της φυσιολογικής σύλληψης και ο πολυμορφισμός PAI-1 4G/5G (recessive model) παρουσιάζουν τα ισχυρότερα πειστικά στοιχεία στην παθογένεση της προεκλαμψίας. Η ανεπάρκεια της βιταμίνης D, ο χαμηλός σε

σύγκριση με τον φυσιολογικό ΔΜΣ, μετρίως και σοβαρή παχυσαρκία έναντι του φυσιολογικού βάρους και ο υποθυρεοειδισμός παρουσιάζουν τα ισχυρότερα πειστικά στοιχεία για την ανάπτυξη του διαβήτη κύησης. Αυτοί οι παράγοντες κινδύνου αποτελούν ένα σημείο εκκίνησης για περαιτέρω αιτιοπαθολογική έρευνα, για τη βελτίωση της πρόβλεψης της προεκλαμψίας και του διαβήτη κύησης, καθώς και για την αναγνώριση των γυναικών που διατρέχουν υψηλό κίνδυνο. Από τις διαθέσιμες παρεμβάσεις για πρόληψη της προεκλαμψίας, η έγκαιρη χορήγηση χαμηλής δόσης ασπιρίνης σε γυναίκες με πρόωρη προεκλαμψία, η διαιτητική συμβουλευτική και διαιτητικές παρεμβάσεις είχαν τα ισχυρότερα επιδημιολογικά στοιχεία που υποδηλώνουν την αποτελεσματικότητά τους. Πιστεύουμε ότι αυτή η αξιολόγηση των ερευνητικών στοιχείων που περιλαμβάνει μια ισχυρή ιεραρχική ταξινόμηση των δημοσιευμένων τεκμηρίων και της ερμηνείας τους, μπορεί να συμβάλει στη λήψη αποφάσεων για την υποστήριξη των κλινικών ιατρών, των επαγγελματιών στη δημόσια υγεία, των ρυθμιστικών αρχών, και των υπευθύνων χάραξης πολιτικής.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 4

Chapter 5

Chapter 6

LIST OF FIGURES

Chapter 2

Chapter 4

Chapter 5

Chapter 6

LIST OF ABBREVIATIONS

ACOG	American College of Obstetricians and Gynecologists
AACE	American Association of Clinical Endocrinologists
AMSTAR	Assessment of Multiple Systematic Reviews
BMI	Body Mass Index
CI	Confidence Intervals
CONSORT	Consolidated Standards of Reporting Trials
CVD	Cardiovascular Disease
FDA	Food and Drug Administration
GDM	Gestational Diabetes Mellitus
GWAS	Genome Wide Association Studies
HCG	Human Chorionic Gonadotropin
HELLP	Hemolysis, Elevated Liver enzyme, Low platelets
HIV	Human Immunodeficiency Virus
HTA	Health Technology Assessment
IOM	Institute of Medicine
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
LAGB	Laparoscopic Adjustable Gastric Band
LGA	Large for Gestational Age
MOOSE	Meta-analysis Of Observational Studies in Epidemiology
MINORS	Methodological Index for Non-Randomized Studies
NICE	National Institute for Health and Care Excellence
NO	Nitric Oxide
NOS	Newcastle Ottawa Scale

OD	Oocyte Donation
OQAQ	Overview of Quality Assessment Questionnaire
PE	Preeclampsia
PCOS	Polycystic Ovary Syndrome
PI	Prediction Intervals
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
PROSPERO	Prospective Register of Systematic Reviews
QUOROM	Quality of Reporting of Meta-analyses
RCT	Randomized Controlled Trial
SE	Standard Error
STREGA	Strengthening the Reporting of Genetic Association Studies
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TSH	Thyroid Stimulating Hormone
USA	United States
USPSTF	United States Preventive Services Task Force
WHO	World Health Organization

Chapter 1 – Introduction

1.1 Brief overview

Several million new research papers are published annually, where from those over than 20,000 are controlled trials of healthcare interventions (1,2). With the ever-increasing of published studies, scientists cannot be expected to examine in detail every single new paper relevant to their interests, where clinicians and patients who are making medical decisions need to know which treatment works best among all treatments for the condition of interest (3–5). Consequently, they increasingly turn to systematic reviews and meta-analyses, which are comprehensive synthesis tools that provide valid, cumulative evidence on relevant topics (4–6). Since their initial publication in 1979, meta-analyses and systematic reviews have become a very popular type of research study that increased the uptake and application of knowledge of clinical and policy decision-makers and ultimately informed policy for public health (7,8). Over the past several decades, the publication rate of systematic reviews and meta-analyses has rapidly accelerated, and their production is still increasing dramatically (9). In 2014, it was estimated that more than 8000 systematic reviews were indexed annually on MEDLINE, a threefold increase over the last decade (10).

Although systematic reviews and meta-analyses are widely considered the highest level of evidence, most of them appear to be either not useful or of uncertain utility (9,11). The problem is that many meta-analyses are not novel as overlapping redundant

meta-analyses on the same topic are very common, whereas when these meta-analyses report discordant results and conclusions, this can cause confusion amongst readers and probably mislead clinicians and policymakers (12–14). Antidepressants offer a case study of the confusing effects of having redundant meta-analyses with different conclusions, as between 2007 and 2014, 185 meta-analyses of antidepressants for depression were published. These meta-analyses are often produced either by industry employees or by authors with industry ties and results are aligned with sponsor interests (9). China has rapidly become the most prolific producer of English-language, PubMed-indexed meta-analyses, with the most massive publication is on genetics. However, genetic association meta-analyses from China typically neglected genome-wide data and combine fragmented information from mostly abandoned era of candidate genes studies, that almost all the times lead to misleading results (15,16). Furthermore, many contracting companies working on evidence synthesis, are currently contracted by pharmaceutical and medical device industries, to produce meta-analyses, many of which probably remain unpublished, creating a skewed picture of the evidence (9,17). Another underlying concern about the methodology and bias of systematic reviews is the quality of the published medical research on which they are based as there are concerns that most of current published research findings are false or grossly overestimated, either because of incorrect or inappropriate statistical analysis of results, or because they include different types of bias in favour of positive statistically significant results (18–20) Evidence from a meta-analysis is also highly depends on the overall methodological rigor that a systematic review or meta-analysis was conducted, which is a function of proper reporting and using accurate methods to limit bias and ensure the internal validity of the findings (10,21). Considering that many systematic reviews and meta-analyses have serious methodological flaws that

limit the validity of their findings, of the remaining, most have weak or insufficient evidence to inform decision making. Few systematic reviews and meta-analyses are both non-misleading and truly informative (9).

While systematic reviews and meta-analyses are successful at summarizing the evidence on a particular research question using multiple related studies, a limitation is that often, a single meta-analysis can address one treatment comparison or one risk factor for a specific outcome, which may offer a limited view of the evidence (22). In addition, the evaluation of biases (e.g. publication bias, reporting bias, selection bias, etc.), in each single meta-analysis is difficult as the data are usually limited (23). To address these shortcomings, an evolving scientific discipline, meta-research, also known as "research on research", has emerged (24). The key characteristic of meta-research is the emphasis on the broader picture since its primary objective is to combine the evidence from multiple systematic reviews and meta-analyses on multiple topics and offer insights about how common and how consistent certain biases are across a large field or multiple fields (24). This type of research can be useful in providing an overview of evidence within a particular area, helps to recognize the relative merits of all available interventions, and consequently be more useful for health technology assessments, evidence-based guidelines and medical decision-making (24). An umbrella review has been reported as one of the four types of "next-generation" systematic reviews that may raise the bar and help shape a new generation of more reliable evidence synthesis (11). The principle reason for the conduct of an umbrella review is to provide an overall examination of the body of information that is available on a given topic using the evidence from multiple systematic reviews and meta-analyses that may be based on outcomes, risk factors or interventions (11,25).

1.2 Aims

To further expand the mapping and the critical appraisal of research evidence across the published literature of clinical identities with a large impact on the perinatal epidemiology field, the umbrella review approach was used. The main aim of the thesis is to systematically overview, analyze and summarize evidence across the published literature on the perinatal epidemiology field, namely preeclampsia and gestational diabetes, and map whether any interventions or fields of risk factors include convincing evidence to support their results. This evaluation of the quality of research evidence and its translation will help to inform medical decision-making and policy-makers.

In summary, this PhD thesis focused on the application of umbrella review approach to:

i. Systematically appraise the evidence on the risk or protective factors for preeclampsia and identified those that supported by high epidemiological credibility.

ii. Systematically examine the evidence on the pharmacologic and non-pharmacologic interventions for preeclampsia prevention and identify those with robust evidence.

iii. Systematically assess the evidence on the risk factors that have been associated with gestational diabetes and detect which factors present the most convincing epidemiological evidence.

1.3 Thesis Overview

Chapter 1 describes the motivation of this work and the main aims of this study.

Chapter 2 provides an extensive literature review, which includes a discussion about the unnecessary, conflicted, and misleading systematic reviews and meta-analyses, the

publication and other reporting biases, the serious methodological flaws that many systematic reviews and meta-analyses have, and the current methodological challenges in perinatal epidemiology field.

Chapter 3 presents meta-research methods, focusing on umbrella review methodology and its contribution to this research study. Basic concepts of key methodology principles used throughout this study such as assessment of summary effect and heterogeneity, evaluation of excess statistical significance, and assessment of epidemiologic credibility are also described.

Chapter 4 presents the umbrella review of systematic reviews and meta-analyses of observational studies on genetic and non-genetic risk factors for preeclampsia. A more detailed description of the methodology used to summarize evidence from the literature on the protective or risk factors for preeclampsia, evaluation of the presence of statistical biases and identification of the associations with robust epidemiologic evidence is provided. The work presented in Chapter 4 has been published in *Ultrasound in Obstetrics & Gynecology* as a research manuscript titled "Genetic and non-genetic risk factors for pre-eclampsia: an umbrella review of systematic reviews and meta-analyses of observational studies".

Chapter 5 presents the umbrella review of meta-analyses and systematic reviews of randomized trials of interventions for preventing preeclampsia. In view of the importance of guidelines for prevention, this study provides a comprehensive summary of the range of pharmacologic and non-pharmacologic interventions, present the magnitude, direction, significance of the reported associations, assess the potential biases, and identify those that present the most convincing epidemiological evidence. The work presented in Chapter 5 is under review for publication in *Clinical*

Epidemiology as a research manuscript titled “Randomized clinical trials for preventing preeclampsia: an umbrella review of the literature”.

Chapter 6 presents the umbrella review of meta-analyses of observational studies on risk factors for gestational diabetes. In this study we applied the methodology of umbrella review, to summarize and evaluate the evidence from all the environmental protective or risk factors that have been associated with gestational diabetes, evaluate whether there are indications of biases in this literature and how these manifest and, finally, identify which of the previously reported associations are supported by convincing evidence. There work presented in Chapter 6 under review for publication in *BJOG: An International Journal of Obstetrics & Gynaecology* as a research manuscript titled “Risk factors for gestational diabetes: An umbrella review of meta-analyses of observational studies”. Finally, Chapter 7 summarizes the conclusive points of this work, its major limitations and strengths, and discusses the implications for future research.

Chapter 2 – Literature review

2.1 Increase in Published Systematic Reviews and Meta-analyses

Currently, there are nearly approximately 17 million articles in PubMed tagged with 'human(s)', with >700,000 articles identified as 'clinical trials', and >1,8 million as 'reviews'. Nearly one million articles on humans are added each year (26). With the ever-increasing number of publications, interest has risen in the development of several methods in order to inform users about the most current evidence that is available from scientific literature towards supporting decision making (27). The most well-known and more frequently used method to summarize available evidence for a particular topic is the performance of a systematic literature review. Unlike other type of reviews such as narrative reviews, a systematic review is expected to involve a more rigorous scientific process characterized by transparency and repeatability (28,29). By examining the accumulated body of evidence rather than the results of single studies, systematic reviews can provide more reliable results for a range of health care enquiries and can also identify gaps in knowledge and inform future research agendas (30,31). A systematic review uses a thorough search strategy and certain eligibility criteria to identify relevant studies that provide evidence to address a particular research question. The purpose of conducting a systematic review in such a stepwise, thorough fashion is to limit the introduction of bias—any process that systematically and non-randomly causes a deviation of results and inferences from the truth—thus making the conclusions of the review more reliable and accurate (32). These reviews,

in their ideal form, include an explicit description of how they were conducted and incorporate methods to minimise bias and maximise precision (33,34). Such methods include the detail description of the methodology (search strategy), a systematic search across several databases to identify studies using predefined eligibility criteria, an assessment of the validity of the findings, and the systematic synthesis and presentation of the characteristics and findings of the included studies (35). When the data allow a quantitative synthesis of results, a systematic review may include a meta-analysis. Meta-analysis refers to the statistical approach that allows the statistical integration of results to produce a pooled-effect estimate from several independent studies addressing the same research question (36). Meta-analysis allows for increased power and precision to detect true differences (and, by definition, a reduced chance of false-negative results, or Type II error) and therefore is less influenced by the findings of any one study. In addition, meta-analyses can help researchers answer additional questions and develop new hypotheses to explain differences between the included studies (35,36).

Publication of systematic reviews and meta-analyses has increased rapidly during the last decade. An inspection of PubMed-indexed in the period January 1, 1986 to December 4, 2015 shows 266,782 items tagged as “systematic reviews” and 58,611 items tagged as “meta-analyses” (Figure 2.1). In 1991, only 1,024 and 334 articles were published as systematic reviews and meta-analyses, respectively. The annual publications of systematic reviews and meta-analyses in 2014 were 28,959 and 9,135, respectively. This corresponds to an increase in the publication rate of 2,728% for systematic reviews and 2,635% for meta-analyses versus an increase of only 153% for all PubMed indexed items (9). This increased publication rate of systematic reviews

and meta-analyses continues to be impressive as between 2010 and 2014 corresponding to 67% and 132% increases, respectively, compared to only 27% increase for all PubMed-indexed items (9). This discouraging situation is also existing for meta-analyses of clinical research, especially of randomized controlled trials, as it is likely that more systematic reviews of clinical trials than new randomized trials are published each year (9).

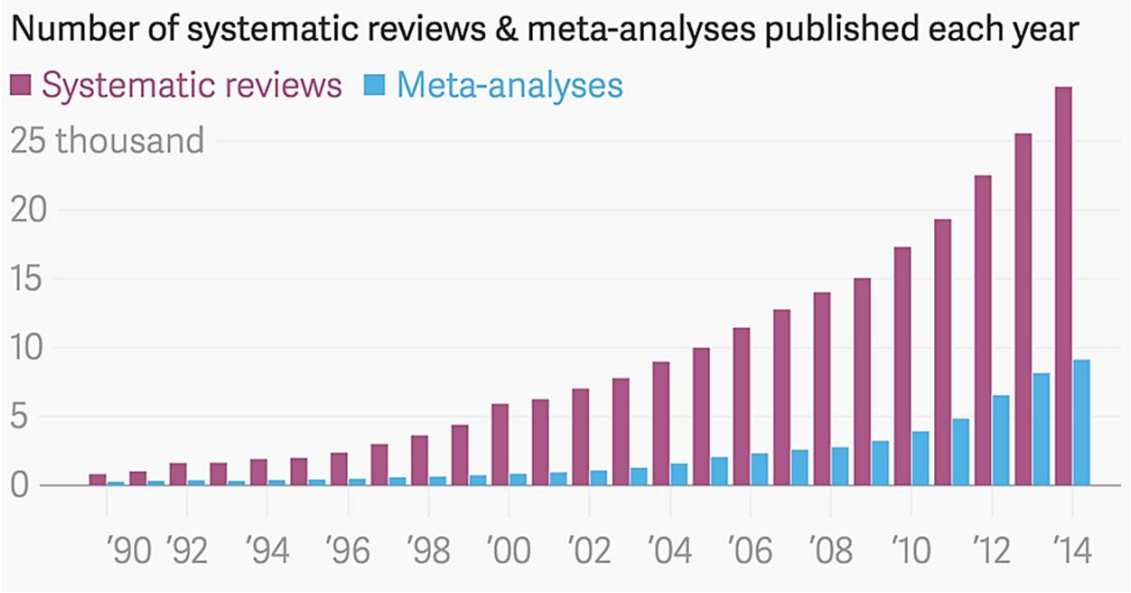

Figure 2.1. Number of PubMed-Indexed articles published each year between 1986 and 2014 that carry the tag "Systematic Review" or "Meta-analysis" for type of publication. Data from: Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Q 2016; 94: 485–514.

Estimates based from a search in MEDLINE in November 2004 conducted by Moher et al. suggested 300 systematic reviews indexed in that month, which corresponded to an annual publication rate of 2500 systematic reviews. The majority (71%) focused on clinical questions (as opposed to a diagnosis, prognosis, or epidemiological question), and 20% were Cochrane systematic reviews (37). The reporting quality varied, with

only 66% reporting the years of their search, 69% assessing study risk of bias/quality, 50% using the term "systematic review" or "meta-analysis" in the title or abstract, 23% formally assessing evidence for publication bias, and 60% reporting the funding source of the systematic review (37). This trend was revised in a recent cross-sectional study of systematic reviews by Page et al. published in 2016, who identified 682 systematic reviews indexed in a single month, suggesting that more than 8000 systematic reviews are being indexed in MEDLINE per year, corresponding to a 3-fold increase over the last decade (10). The majority of systematic reviews addressed a therapeutic question and Cochrane systematic reviews accounted for 15% of the sample. Quality of reporting was highly variable: at least a third of reviews did not report use of a protocol, the search logic for at least one database, methods for data extraction and risk of bias assessment, or the funding source of the review. In addition, at least a third used statistical methods that are discouraged by leading systematic review organizations such as the Cochrane Collaboration (10).

The main deficiency 25 years ago was that there were very few meta-analyses of randomized trials of humans. In 1992, the Cochrane Collaboration was launched with goal to systematically integrate evidence on all medical and health care-related interventions as at that time meta-analyses of randomized trials were rare (9,38,39). As of December 4, 2015, the Cochrane Database of Systematic Reviews included 9,170 entries, which is very close to their original expectation of 10,000 reviews that would be needed to cover the medical and health-care evidence completely (40). Notably, the number of systematic reviews and meta-analyses on the effects of medical interventions from the Cochrane Collaboration is only a small minority of all the published literature (2).

Why are so many systematic reviews and meta-analyses being produced? It is unknown if this mass production of systematic reviews and meta-analyses occurred because of the availability of software that can be used by minimally trained individuals, the limited knowledge on meta-analysis methods in the previous decades or it represents a reflection of efforts to catch up with reviewing the existing published literature (9,12). In addition, researchers face pressures to publish (or perish) in order to advance their careers, whereas journal editors recognize that publishing systematic reviews and meta-analyses can help increase their impact factors since such articles tend to be cited more than other types of studies (41). Perhaps other reasons of this enormous production exist, such as the large impact and importance that meta-analyses have in medical research as the top of the pyramid in most hierarchies of evidence, industry employees can use the results of meta-analyses as a marketing device for their products or because they can be performed with little or no money and can be published in prestigious journals which are often heavily cited (41–43).

An examination of the geographic derivation to detect the countries that are mostly responsible for this massive production of meta-analyses, directs China as the most prolific producer of English-language PubMed-indexed meta-analyses (9). In 2014, over a third (34%) of articles classified as “meta-analyses” in PubMed, have author affiliations from China and only 9% from the United States (USA), which has a distant second place. The change in the geographic origin of meta-analysis occurred in a very short period of time, since in 2005 meta-analyses form China were rare compared to the US (n = 539 from the US vs n = 33 from China). By the 2012 China surpassed the US in production, and currently is publishing 4 times more meta-analyses than the US (Figure 2.2) (9). The rise of meta-analyses from China pertains to all types of meta-

analyses, including those of randomized trials, epidemiological studies, diagnostic-test studies, and any other kind of design (16). However, the most massive rise of Chinese meta-analyses is on the field of genetics, where in 2014, China published 1210 (63%) such genetic meta-analysis articles out of a global total of 1,910, while the US published only 136 (7%) (16).

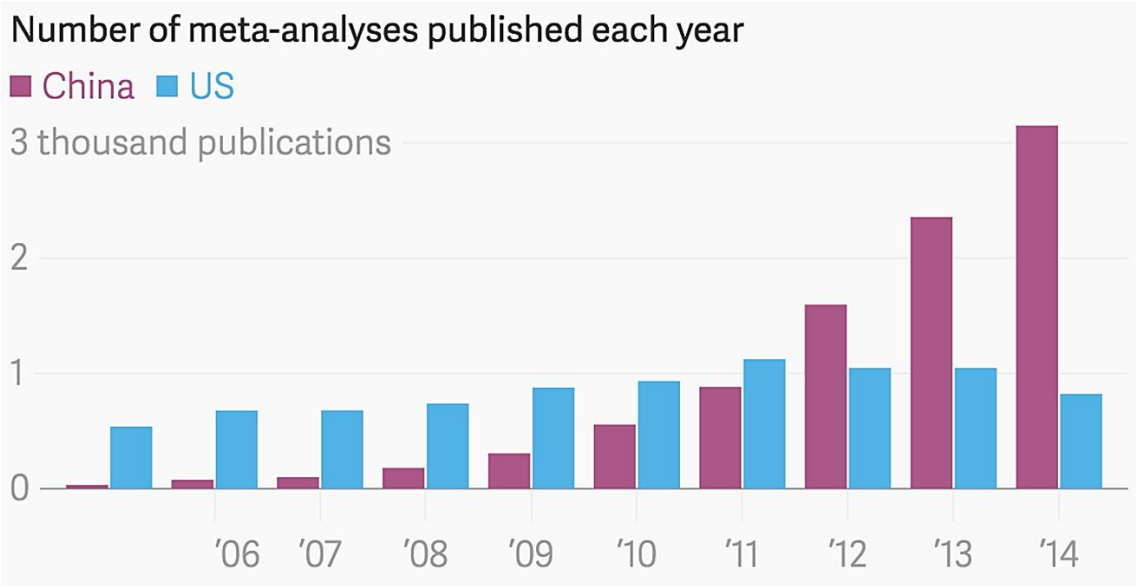

Figure 2.2. Number of PubMed-Indexed articles published each year between 2005 and 2014 that carry the tag "Meta-analysis" for type of publication and have author affiliations from china or from the United States (USA). Data from: Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Q 2016; 94: 485–514

2.2 Unnecessary, Conflicted and Misleading Systematic Reviews and Meta-analyses

2.2.1 Mass Production of Redundant Systematic Reviews and Meta-analyses

As previously discussed, the number of meta-analyses published in recent years has dramatically increased. However, many meta-analyses are not novel as overlapping meta-analyses on the same topic are very common. In the past, multiple independent meta-analyses on the same topic in various research fields have been identified (12). These meta-analyses are representing either serial updates of the same subject by the same team of authors or an independently reproduction of a meta-analysis on the same research topic (12,13). Replication is useful in any scientific field and similarly, independent replication of meta-analyses by different teams could be useful to clarify whether they reach the same results and conclusions (12,16). When new evidence emerges, some meta-analyses might need updating especially if this evidence is likely to modify the conclusions (44). Also, new meta-analyses might be required to examine different outcomes that were not included in the original meta-analysis (16).

A recent study examined how common it is to have multiple overlapping meta-analyses of randomized trials published on the same topic by selecting a random sample (5%) of meta-analyses of randomized trails that were published in 2010. Of 73 eligible meta-analyses published in 2010, 49 (67%) had at least one overlapping meta-analysis published on the same topic by the end of 2012. The median of overlapping meta-analyses was 2, but the maximum was up to 13 meta-analyses (12). Authors from that study also reported that 65% of the subsequent meta-analyses published in 2010 did not include any additional outcomes and 23% of them included one or more authors of the original meta-analyses (12). A cause for concern is that even when published

meta-analyses on the same topic examine different outcomes, the practice of presenting these outcomes in different articles is deficient and confusing (12).

The topic of statins for atrial fibrillation after cardiac surgery provides an example where the extent of unnecessary meta-analyses of randomized trials is most clear. Over the period between 2008 and 2012, 11 meta-analyses of statins for prevention of atrial fibrillation after cardiac surgery were published with a relatively steady appearance of new meta-analyses every few months. Eight of the 11 included only randomized trials, while three also included observational studies. With the exception of the first one, which it was inconclusive and had non-statistically significant results, the remaining showed a highly statistically summary effect and clinically important benefit of statins on the occurrence of postoperative atrial fibrillation, and the treatment effect was consistently large with summary risk ratios ranging between 0.54 and 0.57 and summary odds ratios ranging between 0.40 and 0.78. Of note, some of those had even practically identical results (12). An extension of the search for any additional meta-analyses published until December 2015 identified another 10 potentially eligible meta-analyses on the same topic. This raises doubts whether is reasonable to have newer meta-analyses on the same topic when their incremental value was uncertain and can reflect wasted efforts and inefficiency in the process of summarizing evidence. It is also a matter of interest that the following meta-analyses did not cite systematically the prior meta-analyses on the same topic (9,12).

2.2.2 Mass Production of Conflicted Meta-analyses

Usually multiple systematic reviews or meta-analyses on the same topic would find the same results, however, overlapping meta-analyses may report discordant results and conclusions, especially when the number of the following meta-analyses increases (13). Significant differences in selection criteria, types of studies selected, outcome definition, statistical methods, occasional errors, or even diverse subjective interpretation between overlapping meta-analyses led to discordant estimates. The interpretation of even the same results can differ across systematic reviews and meta-analyses on the same topic, especially when the authors have strong motivations to reach specific conclusion. This phenomenon has been reported previously for a variety of research fields of both meta-analyses of randomized and non-randomized studies (13,45–51). These discordant results can cause confusion amongst readers, waste in research resources, as well as leading to unnecessary duplications, incomplete reporting and public disenchantment with clinical science (13,14,52).

Antidepressants offer a case study of the confusing effects of having redundant meta-analyses with different conclusions and clear example of an area where meta-analyses are used as a powerful marketing tool (9). The market of antidepressants is worth many tens of billions of dollars per year as in the United States only, approximately 10% of people currently take antidepressants, and the use of these drugs has increased fourfold over the last 15 years (53). Given that evidence-based medicine has become so popular, an increasing number of physicians and even patients want to read a systematic review and meta-analysis to be convinced that a treatment is worth adopting (9). An empirical evaluation searched in PubMed for meta-analyses assessing antidepressants for depression published from January 2007 through March 2014 identify 185 meta-

analyses of antidepressants for depression published over these 7 years. Of the 185 meta-analyses, 147 (79%) had a direct involvement from industry (sponsorship, authors who were industry employees and/or authors with industry conflicts of interest) and 54 (29%) had authors who were employees of the assessed drug's manufacturer (53). This represents a massive presence of the industry in generating a prolific production of meta-analyses in this field. Meta-analyses by industry authors often lack a systematic review and focus on pooling individual data from industry trials on a specific manufactured drug. Out of the 185 meta-analyses, only 58 (31%) reported any negative statement about the treatment (e.g. any caveat about their efficacy or safety) in the concluding statement of their abstract. Among those 58 meta-analyses, only one had an author who was an employee at a pharmaceutical company at the time, even though 54 of the 185 total meta-analyses (about 30%) had at least one industry author. That means, when a meta-analysis that had an author who was an employee of the manufacturer of the assessed drug were 22 times less likely to report negative statements about the drug in the abstract that summarizes the conclusions of the work about the antidepressants assessed compared to the other meta-analyses (1/54 [2%] vs 57/131 [44%], $p < 0.001$) (53).

2.2.3 Misleading Genetic Association Meta-analyses from China

As previously discussed, China is the most prolific producer of English-language PubMed-indexed meta-analyses, where the increase was most prominently seen in genetic association meta-analyses (9). An empirical evaluation study compared in-depth 50 genetic association meta-analyses from China versus 50 from USA, published in 2012. Although at face value genetic association meta-analyses from China looked excellent as their reporting was done appropriately, with careful tabulations, and were

published in respectable English-language journals, however, the majority were likely to have reached misleading conclusions. Meta-analyses from China typically neglected genome-wide data, and often included candidate gene studies published in Chinese-language journals, while many USA meta-analyses used genome-wide approaches and raw data. Genetic association meta-analyses from China almost always used only literature-based data (92%) and focused on one or two genes (94%) and variants (78%) identified with candidate gene approaches (88%) (16). The combine fragmented information from mostly abandoned era of candidate genes that led to many thousands of articles with misleading results by American and European teams in the 1990s and early 2000s. This is because, candidate gene studies with single or a few genes and variants addressed one at a time, by single teams, with small sample sizes and with fragmented reporting of the literature subject to publication bias. Almost always, meta-analyses that include such studies give nominally statistically significant results ($p<0.05$), but, this means very little based on what is known in the current era of genomics as almost 99% of the claimed associations were not validated were tested in very large consortia where the entire genome was assessed (15,16). The vast majority of diseases are the result of the interaction between many genes and many environmental factors, hence by selectively choose information about one or a handful of genes has no practical use. Likewise, empirical investigations in some other fields, including single genetic association studies of candidate genes, clinical trials, and randomized trials on acupuncture have suggested that Chinese studies present a prominent excess of significant results that requires cautious interpretation (54–56).

2.2.4 Production of Meta-analyses by Contractors

Contractors is an additional group that is apparently involved in enormous production of meta-analyses. Over the past decade, many contracting companies operating in the domain of evidence synthesis, such as the Mapi Group, Abacus International, Evidera, and Precision for Value. These companies are contracted mostly by pharmaceutical and medical device industries to run meta-analyses for a fee (9). These industries are highly interested in such evidence synthesis tools not only for the reasons that were previously discussed, but also as a means to obtain further insights about the relative merits of their products and of those manufactured by competitors. The meta-analyses are done professionally and at high efficiency, often using advanced techniques, for example, network meta-analysis (57). Using network meta-analysis, it is possible to assess the comparative effectiveness of multiple interventions using both direct and indirect evidence (58–60). These new methods are attractive for clinical researchers because they are particularly useful for clinical guideline development and policy since they seem to respond to their main concern: quantifying relative treatment effects and eventually determining the best available treatments options for efficacy and/or safety. Network meta-analyses can also inform cost-effectiveness analyses and therefore healthcare resource allocation decisions. National agencies for health technology assessment and drug regulators increasingly use such methods (17,61,62).

In contrast to Chinese genetic association meta-analyses, much of the time there is little or no incentive to publish the results of contractor-produced meta-analyses. Non-publication may occur for several reasons, including but not limited to the time and effort to prepare the manuscript and then go through painful reviews and revisions, unfavorable results for the manufacturer, pharmaceutical and medical device

corporations may not wish to share with the public (and consequently also with competitors) private information and/or information that they consider important to give them insights and strategic advantages, low priority for publication for meta-analysis topics that might have already been covered in other published papers, or simply no strong incentives for the manufacturer or contracting company to publish the results (9,17). This produces a skewed picture of the evidence, which is exactly what systematic reviews and meta-analyses are supposed to refrain from.

A recent study was aimed at estimating the number of network meta-analyses performed by consulting companies contracted by industry and explore whether the results of these meta-analyses were published and, if not, why they remain unpublished (17). Two searches were performed to identify the contracting companies. First, MEDLINE was searched from inception until 6 May 2015, for network meta-analyses of randomized trials to found whether they had authors affiliated with any contracting company. Second, the list of the exhibitors at the 20th Annual International Meeting (May 2015) of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) was searched for contracting companies. Afterwards, surveys with questions related to the number of performed network meta-analyses, number published, and reasons for non-publication were sent to these companies. In 162 of 822 (20%) network meta-analyses found, authors were affiliated to 66 contracting companies, while another 36 contracting companies were identified by the exhibitors list. Three companies had no contact information and six merged with others therefore 93 companies were contacted. Thirty seven out of ninety-three (40%) companies responded, and 19 indicated that they had performed a total of 476 network meta-analyses, but only 102 (21%) papers were published. Sixteen out of 19 companies

replied to the second survey, but only 13 provided reasons for non-publication. Based on the replies by the 13 companies, 45 out of 174 (26%) conducted network meta-analyses had already been published and of the 129 still unpublished meta-analyses there was intent to publish about half of the meta-analyses in the peer-reviewed literature while some others have been used for health technology assessment (HTA) submissions with unclear plans for further publication in the scientific literature. This study also revealed that unwillingness of the industry sponsor to allow publication was the most common specified reason for lack of a plan for publication. It is unknown whether the decision for non-publication was made before or after seeing the results and thus whether non-publication reflects the presence of unfavorable results, unwillingness to share with the public, or low priority for publishing meta-analyses on topics already covered in other published papers (17).

2.3 Publication and Other Selecting Reporting Biases

Until today, there are many millions of papers of clinical research and around 1 million papers from clinical trials have been published to date (Ioannidis JP, 2016b). It was estimated that over US$100 billion investment in biomedical research worldwide generated 1 million research publications each year (63). However, there are still concerns that many completed research studies have not been formally published whereas, true and readily valid major discoveries are far fewer since many of the new proposed associations are false or grossly overestimated as they may do not reflect genuine associations but include different types of bias in favor of positive statistically significant results (18–20). The terms publication bias and selective reporting bias refers to the differential choice to publish studies or report particular results depending on the nature or the directionality of findings (35). It has been reported that

approximately 50% of completed studies may remain unpublished (64–66), whereas at the same time, empirical research consistently suggests that published work is more likely to be statistically significant or "positive" than unpublished research (67). The pursuit of statistical significant results may be generated with several different forms, including study publication bias, selective outcome reporting bias, selective analysis reporting bias and fabrication bias (68–71). A most enticing group of such biases are those that can be clustered under the term "significance-chasing biases" (71) (Figure 2.3).

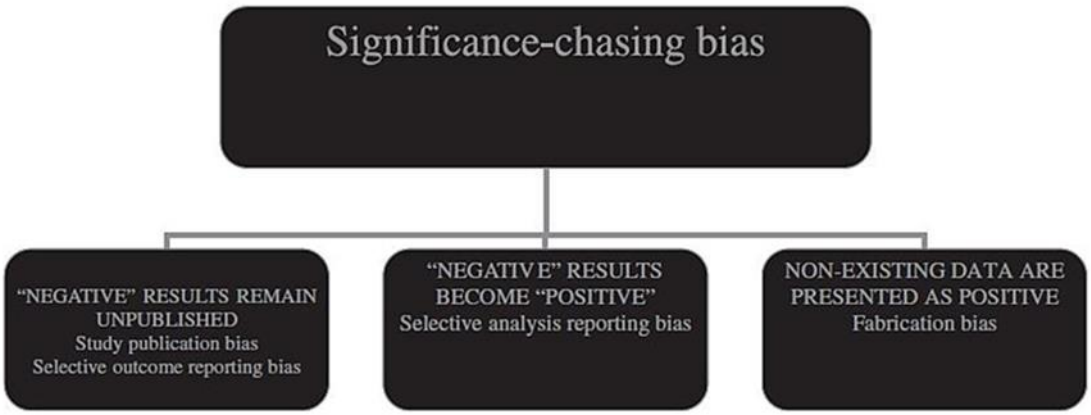

Figure 2.3. Significance-chasing biases. Reprint from: Ioannidis JPA. Meta-research: the art of getting it wrong. Res Synth Methods 2010; 1:169–84

Study publication bias arises when authors are more likely to submit and/or editors are more likely to publish studies when they reach ‘‘positive’’ results. In general, studies with statistically significant or positive results are more likely to be published than those with nonsignificant or negative results (64,67,70). The prevalence of this bias may vary across different scientific fields, proportional to the ease of making a study disappear and the difficulty of making a ‘‘negative’’ study become ‘‘positive’’ with

changes in the analysis plans and/or outcome definition (72). A previous research has demonstrated that only 51% of the antidepressant trials registered with the Food and Drug Administration (FDA) had been "positive", where by contrast, as many as 94% of trials published in the peer-reviewed literature evaluating antidepressant agents were "positive" (73). Publication bias has been recognized as a problem in medical research for many years. When the research that is readily available differs in its results from the results of all the research that has been done in an area, readers and reviewers of that research are in danger of drawing the wrong conclusion about what that body of research shows (74). The first article with the term "publication bias" that could be identified by searching PubMed was published in 1979 and since then, the number of references that are potentially relevant to publication bias has considerably increased. This increase in the number of relevant studies on publication bias may reflect the increased awareness of publication and related biases (75).

Selective outcome reporting bias can occur in three ways; when multiple outcomes are evaluated in a study and the outcomes found to be significant are more likely to be published; selective reporting of a specific outcome, for example, when an outcome is measured and analyzed at several time points but not all results are reported; and incomplete reporting of an outcome (71,76). Selective analysis reporting bias occurs where certain data are analyzed using different analytical options such as subgroup analyses or intention-to–treat analyses versus per–protocol analyses, and publication favors the more impressive, statistically significant results (70,77,78). Non-existing data may be presented as "positive", but fabrication bias is unlikely to be as common to other types of bias in favor of statistically significant results (69).

There is additional evidence indicate that research without statistically significant results takes longer to achieve publication than research with significant results, which further biasing evidence over time (65,79,80). This "time-lag bias" is another form of bias that can also affect perceived efficacy of interventions. For example, one study assessing efficacy trials of human immunodeficiency virus (HIV) treatments concluded that the time from study enrollment to publication was significantly longer for negative trials than that for positive trials (80). A recent meta-analysis of published and unpublished randomized controlled trials (RCTs) of serotonin reuptake inhibitors in subjects less than 18 years old with major depressive disorder examine if there is evidence of a time-lag bias in the publication of pediatric antidepressant trials. Despite the small number of trials, authors found a significant evidence of time-lag bias in the publication of findings and concluded that time-lag bias is not unique to child psychiatry and reflects a larger problem in scientific publishing (81).

In addition, a number of other potential information suppression mechanisms exist, including: language bias (selective inclusion of studies published in English); availability bias (selective inclusion of studies that are easily accessible to the researcher); cost bias (selective inclusion of studies that are available free or at low cost); and familiarity bias (selective inclusion of studies only from one's own discipline (74). All of these biases lead to the same consequence, namely that the literature located by a systematic reviewer will be unrepresentative of the population of completed studies, hereafter all present the same threat to a review's validity. For this reason, it has been suggested that a single, broadly encompassing term, dissemination bias, to be used to refer to the problem (64).

These “significance-chasing biases” eventually can cause a relative excess of published statistically significant results that distort the totality of the available evidence on a research question and leads in misleading estimates of treatment effects and associations between study variables (64,82). Selective reporting biases affecting specific outcomes and specific analyses within studies is probably the greatest and most intangible concern that distorts the literature across many fields (83–87).

Consequences of these biases depend on types of research (basic biomedical, observational, or clinical studies) and levels of result acceptability, but the detrimental consequences are the avoidable suffering of patients and waste of limited resources (75). For instance, in basic medical research, due to biased results from falsely positive studies, subsequent clinical trials may waste limited resources and fail to confirm the previous published results (88,89). This observation was revealed in a recent analysis of 4445 animal studies in 160 meta-analyses of neurological diseases, where 112 meta-analyses (70%) found nominally ($p<0.05$) statistically significant results. Authors concluded that perhaps the majority of the data were either suppressed or recast in a way that truly negative studies would be published as positive results since there were just too many positive results published to be true. This observation also suggests strong biases, with selective analysis and outcome reporting biases being plausible explanations (90). It is estimated that over 50% of preclinical research can’t be replicated, placing the approximate annual cost of irreproducibility in the US alone at US$28 billion, whereas unsurprisingly, drug discovery has reduced, and its costs have risen, as preclinical interventions in animal models are rarely recapitulated in clinical trials (91). Results of observational studies are often highly contradictory over an extensive variety of risk factors, which might be due to publication bias. For example,

publication bias may cause highly contradictory results observed in early published studies of genetic associations (92). Publication bias in clinical trials has a direct impact on patients' and populations' health (75). When the relative efficacy of a treatment is overestimated because of publication bias, health resources can be wasted by obtaining more expensive interventions, instead of cheaper alternatives, without corresponding improvement in outcome (75). There are also many reported cases in which patients have received ineffective or harmful treatments (64).

But how systematic reviews and meta-analyses can be affected from these biases? Information from multiple primary studies can be synthesized either prospectively or retrospectively. Ideally, meta-analyses should be conducted in consortia where investigators collaborate preventively with embedded replication across teams and joint analyses (9). In the past, large consortia have been successfully conducted in prospective meta-analyses of genome data (93). However, teamwork, collaboration, and replication are rare in most fields due to lack of incentives and therefore most systematic reviews and meta-analyses conducted today are retrospective (9). Hence, evidence from a retrospective meta-analysis highly depends on the quality of the included studies, and if poor-quality data, overly biased data, or data that do not make sense are combining together, then systematic reviews and meta-analyses will have misleading inferences and estimates which can cause major negative effects on the credibility and value of research evidence and turn out to be unreliable for decision making (71,75,82,94).

2.4 Flawed Meta-analyses and Correct but Non-Informative Meta-analyses

As discussed earlier, evidence from a meta-analysis highly depends on the quality of the studies included and the overall methodological rigor which the meta-analysis was conducted. Therefore, a rigorous evaluation of the validity of primary studies is fundamental to the validity of the assumptions derived by the meta-analysis (95).

A major advance in evidence-based medicine has been the development of initiatives to improve methodological quality and reporting of systematic reviews and meta-analyses for various forms of evidence (e.g. randomized or non-randomized), that include among other principles of research question formulation, the use of a comprehensive search strategy, assessment of methodological quality of the primary studies and evaluation of heterogeneity (35,96,97). The first checklist specific to meta-analyses was the Quality of Reporting of Meta-analyses (QUOROM), which was published nearly two decades ago, designed to address the suboptimal reporting of meta-analyses (98). QUOROM is similar to the Consolidated Standards of Reporting Trials (CONSORT) for reporting of RCTs, which was published in 1996 (99–101). In 2009, QUOROM was revised to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) to encompass both systematic reviews and meta-analyses and address several conceptual and practical advances in the science of secondary research (102). Because of the increasing number of published meta-analyses using observational studies, the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) Group created a 6-section that contains specifications for reporting of meta-analyses of observational studies, including background, search strategy, methods, results, discussion, and conclusion (103).

Adequate reporting does not necessarily ensure that the contents of the document are valid and precise. The overall quality of a systematic review or meta-analysis is a function of proper reporting but, much more importantly, of using accurate methods to limit bias and ensure the internal validity of the findings (10,21). As such, similar to checklists for reporting, checklists for methodology have been created such as the Overview of Quality Assessment Questionnaire (OQAQ), the Potsdam guidelines, and the Sacks instrument (97,104–106). In 2007, the Assessment of Multiple Systematic Reviews (AMSTAR) was developed by combining elements of the OQAQ and the Sacks instrument as well as other items based on their methodological advances (107), that it has been found to be valid, feasible, reliable, and to have good inter-rater agreement (108,109). Since their development, PRISMA and AMSTAR have become widely accepted by many journals as the tools to ensure proper reporting and methodology of systematic reviews and meta-analyses (21).

Until today, many studies have assessed the reporting and methodological quality of published systematic reviews and meta-analyses of both observational and randomized evidence in a variety research fields, in order to assess the prevalence of methodological flaws in the design, conduct, analysis and reporting. Despite the available guidelines, these studies have revealed serious methodological flaws of most of the included systematic reviews and meta-analyses as essential methodological components of the systematic review process, such as conducting a thorough literature search and assessing risk of bias of primary studies were frequently missing in their reports, even when published in journals with high impact factors. This may impair the validity of these publications and thus limit their value to guide policy decisions and clinical practice or their use for educational and research purposes

(21,37,61,95,110–116). In addition, many studies reported that Cochrane meta-analyses have higher overall quality scores compared with those published in peer reviewed journals (37,111,113,115), whereas a trend of an overestimation of treatment effect in meta-analyses of lower quality scores was also observed (110,117,118).

Moreover, even when published systematic reviews and meta-analyses are well performed with no evidence of methodological flaws, may still not be informative. It is very common, especially in meta-analyses of randomized trials; authors to conclude that the available evidence is weak or inefficient to answer the key research question. Hence, the correct but non-informative meta-analyses fail to inform decision making on patient care or health policy (119–121). For instance, a recent study that evaluated 3,826 systematic reviews produced by the Cochrane Collaboration that involve physiotherapeutic treatments reported that only 0.5% of the reviews concluded that the intervention presented a positive effect and that further studies were not recommended, whereas a significant proportion (46.9%) found that the evidence was insufficient for clinical practice and recommended further research (122). These results are comparable with those from another study that analyzed a random sample of Cochrane systematic reviews of a variety of interventions (e.g. drug therapy, surgery etc.), that found only 0.98% of the 1016 reviews did the authors find insufficient evidence to support or refute the indication, while around half of the reviews examined (47.83%) did not offer enough evidence for clinical practice, and the authors asked for further research (123).

2.5 Animal Studies in Human Research

Over the past centuries animal research have been successfully used in many areas of science, such as in basic research, and played an important role in the development of modern medical treatments (124,125). Research based on animals has brought new and deeper understanding about basic mechanisms of the human body and have provided valuable contributions to the development of great medical advances that impact diseases such as polio and Parkinson's disease. Advances in surgeries and treatments including kidney and heart transplantation were also perfected with the use of animals (126–128). Experiments using animals not only helped to the development of new vaccines for the treatment of infectious diseases like diphtheria, tetanus, tuberculosis, poliomyelitis, and measles, but it also led to the development of greatly needed medicines, such as antibacterial and antibiotic drugs (125,129). Furthermore, animal studies can provide unique insights into the pathophysiology and causes of disease, and often reveal novel targets for directed treatments. Pre-clinical studies in selected animal species are also needed to formulate hypotheses that justify clinical trials. Without such studies it would be unethical to test unproven chemicals in humans and it may not be necessary to test new treatments on humans if preliminary testing on animals shows that they are not clinically useful (130,131). In addition, extensive animal testing is required from regulatory authorities concerned with public protection to screen new treatments for toxicity and to establish safety (131).

Although, the history of today's therapeutic armamentarium has always involved animal testing, we cannot overlook the fact that the use of animals in research has always aroused controversy on ethical and technical grounds. Up until today there is an ongoing debate over the propriety and value of using animals in medical and

scientific research (130). Decades of animal experimentation for specific diseases such as cancer and diabetes have produced little or nothing of value to humans as encouraging results in animal's studies often does not translated to successful human randomized trials (89,132,133). For instance, the traditional mouse models for cancer has now been widely discredited as human cancer cell lines are more accurate for identifying effective cancer drugs compared to animals, and in fact the traditional mouse allograft model is not predictive at all (134–136). Similarly, the entire field of mouse immunology research is tainted by the recent discovery that, unlike humans, mice have a second thymus gland (137). In addition, despite the existence of numerous successful animal models for traumatic brain injury, diabetes and stoke treatments, each one has failed to confirm benefits for humans (130,138,139).

Several analyses have set out to understand why the extrapolation of results from animals to human sometimes fail. One obvious reason is the difference not so much in organ composition and functions, but the greater complexity of man compared to all the animal species. Even though the lab animals have many similar features to humans and usually animal models are excellent representations of most human characteristics and attributes, still, vast anatomical, physiological, and genetic differences between humans and animals, might be a reason of the poor translation of the results from animals to humans (130,140,141). In addition, the human organism often differs dramatically from the animal species involved in pre-clinical studies with respect to absorption, distribution and excretion of substances, and forms very different metabolites of the same drug (129,130). Another explanation is that animal models may not adequately mimic human pathophysiology. Lab animals are often young, rarely have comorbidities, and are not exposed to the range of competing interventions

that humans often receive. The timing, route, and formulation of the intervention may also introduce problems (131,142).

Moreover, there is growing opinion among scientists that an important part of discrepancy between animal and human studies is because of the poor quality and methodological biases in animal experimentation as well as the lack of adequate reporting of animal data (129–131,143). Bias related to randomization, double blinding, surrogate end-points, calculation of sample size, statistical analysis, and nonpublication of negative results still greatly limits the extrapolation of animal findings to human (130,133,144,145). For instance, an analysis of 76 animal studies published in top journals between 1980 and 2000 show that only around a third of highly cited animal research translated at the level of human randomized trials and only 49% as having good methodological quality (133). In one another analysis of 290 animal experiments presented at emergency medicine meetings, animal studies that did not use randomization or blinding were much more likely to report a treatment effect than studies that were randomized or blinded (144). In a recent analysis of 4445 animal studies in 160 meta-analyses of neurological diseases, authors concluded that perhaps most of the data were either suppressed or recast in a way that truly negative studies would be published as positive results, suggesting strong biases, with selective analysis and outcome reporting biases being plausible explanations (90). Similarly, systematic reviews of animal studies have also revealed evidence of selective analysis and outcome reporting bias as well as publication bias leading to overstatement of the validity of entire bodies of research (89,146–149).

In response to the serious deficiencies found in the conduct and reporting of animal studies the Animal Research: Reporting In Vivo Experiments (ARRIVE) guidelines were produced in 2010 (150). Recent attempts to improve translation within the animal research also include the "co-clinical trial" in which preclinical trials explicitly parallel ongoing human phase I and II trials (151). Nevertheless, there is certainly plenty room for substantial improvement in animal research to improve their credibility and reproducibility. The importance and reliance on animal models may change in the future along with the development of more advanced nonanimal research technologies (e.g. computational models, bioinformatics, stem cell methods, and genetic methods), yet, for the time being, experiments involving animals remain an essential aspect of human research design.

2.6 Challenges in Perinatal Epidemiology

One of the most vulnerable periods of human life is the period of intrauterine growth and development. Events during pregnancy are important influences on the outcome of pregnancy and the health and wellbeing of the newborn (152). What happens in pregnancy and the very early stages of childhood will have a profound impact on child and adolescent development (153). There is also increasing evidence for the role of early adverse experiences during pregnancy on childhood and adult health as well as interest in the possibility of intergenerational effects of events (i.e. effects of events during pregnancy on the outcomes of pregnancy and health in subsequent generations) (152). The "fetal origins hypothesis" describes that maternal health and nutrition in the prenatal period send signals to the fetus about the relative harshness of world in which he or she will be born (154). For instance, supporting this hypothesis, several studies

have found associations between low birth weight and long-term health outcomes such as diabetes and heart disease (155–158).

Perinatology is a medical specialty field that was established to provide integrated care to mother and fetus and to bridge the gap between the obstetrician's concern for the pregnant woman and the pediatrician's concern for the infant (152). Building on the existence of perinatology as a medical specialty, perinatal epidemiology has developed as a subspecialty of epidemiology (152). Perinatal epidemiology research is concerned with identifying the effects of events during pregnancy on pregnancy outcome, including maternal, fetal, and neonatal health outcomes (159). It also encompasses the study of the effects of factors inherent to the pregnant woman such as age and ethnicity, voluntary harmful exposures during pregnancy (e.g. smoking and alcohol use), environmental exposures, diet, genetic constitution, the effects of illness, and the use of medications (152). While the focus of epidemiology has traditionally been on "disease" not its converse health, although they are clearly interrelated, perinatal epidemiology research differs in at least three ways (160). First, the broader view of health rather than disease is especially appropriate in perinatal research. Pregnancy is in most cases a healthy life transition, where changes in social and role function are expected and many of the symptoms of pregnancy, such as first trimester nausea or third trimester backache are considered "normal". Hence, the model is *not* one of curing the disease, and outcomes evaluations should consider the normal process of childbearing and its impact on normative functioning (160). Second, with pregnancy as opposed to the most acute and chronic disease there is a predictable progression and time course, which is generally 40 weeks' gestation (± 2 weeks, from the last menstrual period), with a key definable outcome to the health state-delivery of the infant. Third,

during the perinatal period there are two patients, the mother and the baby and measures to assess outcomes need to include both patients (160).

One important methodological challenge in the design and conduct of perinatal research is that randomizing patients is not always feasible. The importance of evidence from RCTs is now widely recognized, as they are considered the most appropriate way to evaluate the impact of an intervention in clinical practice and often referred to as the "gold standard" of research methods (161–164). Randomization is the theoretically ideal way to draw strong inferences about the effect of an exposure on maternal, fetal, neonatal, and infant outcomes, because randomization ensures that the intervention and control group(s) are comparable in terms of factors other than the one being studied (152). However, many factors that affect the outcome of pregnancy cannot be assigned at random, consequently when these factors are of interest, a randomized trial cannot be conducted. Also, some factors that cannot be assigned at random, such as age, ethnicity, and genetic constitution, are non-modifiable, and as result they cannot be studied in randomized trials as they are not subject to manipulation by the researcher. In addition, some factors including cigarette smoking, alcohol use, and cocaine and heroin use, cannot be assigned at random for practical or ethical reasons (152). Even when an exposure, like medication use, can be assigned at random, attaining sufficient enrollment for an adequately powered RCT in a reasonable amount of time can also be challenging. For instance, it is estimated that asthma occurs in approximately 5% of pregnancies. A randomized trial comparing two medications for the treatment of asthma, that sought to enroll 200 women (100 in each group) would require a base population of 4000 pregnant women if all of the women with asthma were eligible for the study and consented to enroll. However, eligibility

criteria and unwillingness to participate would reduce the number of pregnant women available for a trial, hence; a study involving 200 pregnant women with asthma would require a large base population (e.g. 16,000-20,000) of pregnant women to be successful (152).

Because of the inability of randomization, most epidemiological studies conducted during pregnancy are non-experimental or non-randomized. The concern of non-experimental studies is bias, which might arise from flaws in the study design, conduct of the study or in the presentation of the results (152). Addressing confounding is another key methodological issue in non-randomized studies. Mixing the effect of exposure on occurrence of outcome with a third factor, called confounder, happens when a confounder is an independent risk factor for the outcome and has an independently statistical association with the exposure of interest (165,166). A confounder should also not be at intermediate pathway between exposure and outcome (166). Depending on the interrelation between confounder with exposure and outcome, uncontrolled confounding leads to over or under estimation of measure of association and consequently to erroneous conclusions (167). The issue of baseline population comparability, often referred to as risk-adjusted or adjustment for case-mix, is a primary methodological issue in the design and conduct of perinatal outcome studies (160). When comparisons are made across treatments, programs, providers, or institution the case-mix of those groups must be considered (168). For instance, comparing maternal or neonatal outcomes between women who deliver at levels I vs. levels III regional perinatal hospitals should consider the perinatal risk of women being treated at each hospital, since the perinatal outcomes of the level III hospital would be expected to be worse as these hospitals typically have more high-risk patients (160).

The need refinement of traditional perinatal outcomes, such as low birth weight is another issue that stand out as methodological challenge in the perinatal research (160). Many traditional measures in perinatal research could be considered intermediate measures. One much used example is low birth weight (<2500 g). Although, the number of studies that highlighted birthweight as a predictor of neonatal and infant mortality increased dramatically, birthweight, *per se*, is not a disease, but birthweight <2500 g is highly predictive of many diseases of the newborn. Despite the fact this research has contributed to our understanding of the predictors of neonatal health and has had considerable effects on public health programs, we must still recognize it as an intermediate outcome (160). One more issue in trying to define outcomes is to differentiate process from outcome. An example in perinatal epidemiology is the frequent use of Caesarean section as an outcome for maternal health (169–172). However, Caesarean section as a dichotomous variable merely describes that, it is a procedure and not an outcome that reflects the actual health status of either the mother or the infant (160).

Another methodological issue is perinatal epidemiology is how long time-period should be considered. Conceptually, research attempts to move beyond the defined medical event, to examine the wider and sometimes longer-term impact of medical care on the individual or population (160). However, because of the potentially lengthy lifetime of a mother and newborn after birth, long-term examinations can be unbearable, and a shortened period of interest might be used, such as the first few months or the first few years of life. Yet, there may be potential bias when using shorter time periods, as significant events beyond the specified period would not be accounted for (160). Finally, the methodological issue of the multiplicity of outcomes of interest

is perinatal epidemiology should also be considered. Factors that affect pregnancy outcome are complexly interrelated and this makes the field challenging because it requires an understanding of the outcome's pathophysiology as well as the factors that affect each one (152). Recognition of the interplay between several factors on the outcome of interest is not only important in the design and conduct of perinatal research studies but it also complicates the interpretation of the findings.

2.7 Genetic Background of Preeclampsia

Most reproductive diseases seem to represent complex genetic disorders as it is thought that no simple correspondence between genotype and phenotype exists and both genetic and environmental factors contribute to the susceptibility risk (173). Undoubtedly, the genetic architecture behind preeclampsia is complex as includes environmental factors, maternal, paternal and fetal genes, and their combined effects (174). Complex diseases occur as the result of numerous common variants at different loci which individually have a small effect but collectively contribute to an individual's susceptibility to disease (175). The degree of genetic influence on preeclampsia has first suggested by the observed incidence of the disease in relatives, and a familial tendency in the nineteenth century (176). The familial aggregation of preeclampsia is often assessed using twin studies that can help to distinguish between environmental and genetic influences on individual traits and behaviors. Despite that very few twin studies have been possible to conducted because of the rarity of the disease, these have revealed that preeclampsia has a higher relative risk compared to controls, but there is no simple mode of inheritance (174,177–180). Taken together, in twin studies, the incidence rate of preeclampsia might be different, and pathogenic

effects of other factors in addition to genetic predispositions play important roles in the onset of this complex disease (181).

In theory, identification of candidate genes for preeclampsia could substantially help the understanding of this central public health problem and provide clues for its prediction, prevention and treatment (182). Several studies have been conducted to date that reported associations between preeclampsia and polymorphisms and mutations of various genes that were selected based on their contribution in cellular pathways linking to the clinical features of preeclampsia. Numerous candidate genes have been proposed as having a role, primarily those with a plausible role in the known underlying pathophysiology of preeclampsia, mainly genes involved in renin-angiotensin system, immune maladaptation, inherited thrombophilias, synthesis, placental ischemia, and increased oxidative stress (175,183,184). However, after two decades of research using the candidate gene approach and linkage analysis, no single genetic susceptibility for preeclampsia has been confirmed or refuted as candidate gene studies have been undermined by conflicting and inconclusive results (182,184). This design requires sample sizes of thousands to have adequate power to detect realistic genotypic relative risks of ~1.1–1.3 and only few studies have been of this size in the preeclampsia field (182,185). In addition, selection of candidate genes for examination is limited by an incomplete understanding of biological mechanisms involved in the pathogenesis of preeclampsia, and often such studies are focused on a limited number of candidate genes and lack of reproducibility, that undermines the reliability of association with preeclampsia. Likewise, inconsistency of clinical diagnosis and ethnic variations within study populations may also had an impact on research findings (182,184,186). Although large research efforts have been devoted to

the analysis of single gene contributions using the candidate gene approach or genetic linkage analysis in families, still, no universally reliable genetic variants have been identified. This might be because such approaches have been much less successful in disentangling the genetic risk for more common and complex diseases like preeclampsia with genetic changes combined with environmental factors and polygenic susceptibility (184,187–191).

The technological advances that allowed for the development of large genotyping arrays have made genome-wide association studies (GWAS) commonplace in disease gene mapping over the past decade (192). Through GWAS it is possible to find single nucleotide polymorphism (SNP) that is associated with a disease and indicates a region of the human genome which influences the risk of the specific disorder. GWAS have evolved over the years into a hypothesis-free, unbiased approach, with the potential for identifying novel genetic variants (186). Recent GWAS in the field have yielded encouraging results, however, given that preeclampsia is a complex disease with great phenotypic diversity, it is apparent that larger studies with adequate statistical power are needed to improve our genetic knowledge base for this complex disease (182,184,186). Three GWAS have been published today that include several genetic loci linked or associated with preeclampsia (193–195). Two of the three GWAS had a smaller number of cases and did not find any genome-wide significant associations (193,195), whereas the third, identified two loci (rs7579169 and rs12711941) near the Inhibin beta B gene that satisfied the genome-wide significance threshold, but they could not be replicated in two cohorts from Norway and Finland (194). Subsequent case-control studies in European and Chinese women have shown a significant ($P<0.05$) association between the SNP rs7579169 and preeclampsia (196,197).

The terms “polygenic scores” (PGS), “genetic risk scores” (GRS) and “polygenic risk scores” (PRS) are used to describe the approaches designed to summarize genome-wide genotype data into a single variable that measures genetic liability to a disorder or a trait (192,198). Technically, such scores are calculated from GWAS summary statistics to explore the genetic contribution to the disease’s etiology and/or to predict of individual disease risk (192,198). The use of a GRS based on GWAS findings as an indicator of risk for a given condition is a novel method of investigating genetic susceptibility to a complex trait (199). Although GRS are easy to calculate and capture important information about an individual’s risk of developing a disease, still is unlikely to have sufficient utility, so it may be more useful when combined with environmental risk factors or with high-risk variants (198). Three studies have been published until today to determine the association between GRS and risk of preeclampsia (199–201). The first study that investigated the association between an established GRS for hypertension (SBP, DBP, and MAP) and preeclampsia in two different study populations did not identify a statistically significant relationship, suggesting that an underlying predisposition to essential hypertension is not on the causal pathway of preeclampsia (199). The association between the genetic predisposition to dyslipidemia, estimated by four GRS (total cholesterol, LDL cholesterol (LDL-C), HDL cholesterol (HDL-C), and triglycerides) based on established loci for blood lipids, and risk of preeclampsia was also examined. The results of this study demonstrate that only the GRS related to lower HDL-C was marginally associated with an increased risk for preeclampsia, suggesting that dyslipidemia may be a component along the causal pathway to preeclampsia (200). Lastly, in a recent study the association of a GRS for elevated levels of C-reactive protein (CRP) and the risk of preeclampsia was also examined, in which an increased

genetic risk for elevated CRP was found to be protective against the development of preeclampsia in two independent populations (201).

In summary, although there seems to be a familial effect on the liability of developing preeclampsia, there is no simple mode of inheritance, due to preeclampsia is considered a complex disease with great phenotypic diversity. Given that the origins of preeclampsia are still not well understood, future research studies should focus on additional loci, particularly rare variants, to explain the etiology of preeclampsia. Reaching common agreements between researchers on the definition and reporting of preeclampsia will contribute to a more efficient translation of future knowledge into public health and medical interventions (182).

2.8 Prediction of Preeclampsia and Gestational Diabetes

To assess a screening tool's predictive ability, sensitivity, specificity, positive predictive value and negative predictive value should be assessed. A perfect screening test would be 100% sensitive and 100% specific, henceforth, would be positive for all those with the disease and negative for all those who did not (202). In clinical terms this means, with a high sensitivity, most patients who are going to develop the disease will screen positive, while with a lower specificity, it means that some patients who will not develop the disease will also screen positive. The test should also be simple, rapid, non-invasive, inexpensive as well as valid, reliable and reproducible (203). The ultimate predictor of preeclampsia and gestational diabetes should presumably identify women with an increased risk early in pregnancy, who could be offered potential treatment to prevent the disorder and thereby reduce its negative consequences. A new screening method has been previously proposed by Kypros Nicolaides, in which

clinicians instead of increasing the number of prenatal care visits towards the end of the pregnancy, effective screening at the beginning of pregnancy was recommended. This would lead to fewer unnecessary visits and more focused prenatal care (204,205). By following this strategy, low-risk pregnancies would attend a standard care program with fewer visits, while a more accurate monitoring of high-risk cases and possible prophylactic treatment (e.g. low-dose aspirin) would possibly lead to early diagnosis, a reduction in the number of complicated pregnancies, better define risk and direct resources to reduce morbidity, but also could lead to fewer long term complications for both mother and child (204,206,207).

Partly due to lack of knowledge of possible underlying pathophysiological mechanisms involved in preeclampsia, there are not yet any reliable and validated predictors to identify most women who will develop the disease. The traditional method for detection and diagnosis of preeclampsia is by routine detection of its signs such as raised blood-pressure and proteinuria during antenatal monitoring that could indicate evolving the disease. Unfortunately, this method is not valuable for early prediction or identification of a high-risk women that are possible to develop the disease (208,209). With one in 10 pregnant women developing symptoms suggestive of preeclampsia (e.g. headache, abdominal pain) but only 20% of these reaching a diagnosis, there is a clear need for improved testing methods (210). In addition, there are certain at-risk groups of patients such as those with chronic hypertension, pre-gestational diabetes, multifetal gestations and previous preeclampsia. Established organizations such as the ACOG and NICE endorses evaluation of risk factors as the best and only recommended screening approach for preeclampsia. Although recognition of clinical and demographic factors can be useful in clinical practice and

might help when selecting a high-risk group, yet, they are neither sensitive enough nor sufficiently specific to be used alone and therefore, they cannot be used reliably for prediction of preeclampsia. Also, this approach of screening is likely to result in classifying many pregnant women as screen-positive and consequently in need of more regular monitoring, which undermines the purpose of screening and creates a significant strain on the healthcare system (209,211).

Even though preeclampsia is a highly clinically relevant topic, no early and reliable first trimester marker is currently available for early prediction of development of this pregnancy-associated disease. An extensive research in the last 20 years, has identified a wide range of potential biophysical and biochemical predictors of preeclampsia based on our knowledge of the pathophysiology of this disease (212–215). Many of these markers are measurable in maternal blood and have therefore been evaluated as biomarkers for the prediction of preeclampsia. These include serum and plasma markers of placental endocrine function, maternal endothelial dysfunction, renal dysfunction, general metabolic status, oxidative stress, and hemolysis and inflammatory markers (216). Recently investigated screening markers for preeclampsia include factors related to angiogenesis, coagulation, lipids, placental hormones, cell adhesion, fetal DNA, inflammation, and growth factors. Despite several years of research in the field, a single test accurate enough to predict preeclampsia sufficiently well has not yet been found (203,217,218). A recent review of different biochemical markers for preeclampsia before the 25th week of gestation in cohort and case control studies revealed no test with a sensitivity and specificity over 90% (217). Another study reviewed 27 different tests for preeclampsia prediction,

but only a few reached specifications above 90%. These were BMI of 34kg/m² or higher, α-fetoprotein and bilateral uterine artery Doppler notching (219).

The absence of a robust, sensitive single marker is not surprising since preeclampsia is characterized by a complex pathophysiology with heterogeneous clinical and laboratory findings. Thus, it is unlikely that a single marker could predict the mixed presentations and potential causes of the disorder (214,215). It is now recognized that combinations of markers that reflect different aspects of disease's pathogenesis are needed to improve the possibility for predicting preeclampsia with a high degree of accuracy (203). Potential components of such a combination could be anamnestic risk factors, angiogenic, inflammatory and other biochemical factors, uterine artery Doppler and mean arterial pressure (MAP). Yet, until to date, there is no general acceptance of these combinations in clinical practice (214,215). A previous large study that combined maternal characteristics, including MAP, uterine artery pulsatility index and the biochemical markers PAPP-A, PlGF, PP13, sEndoglin, Inhibin-A, Activin-A, Pentraxin 3 and P-Selectin has demonstrated 95% specificity, for early-onset 91% sensitivity, intermediate onset 79% sensitivity and late onset preeclampsia 61% sensitivity (220). Another example of such combination is the foetal haemoglobin (HbF)/Haemoglobin ratio and α1-microglobulin, that has demonstrated 90% sensitivity and 77% specificity for prediction of preeclampsia in early pregnancy (221). Findings of systematic reviews and meta-analyses that evaluated the predictive capabilities of combinations of biochemical and ultrasonographic markers showed that such combinations predicted preeclampsia better than a single predictor and this might improve the prediction of preeclampsia, especially in high-risk populations (218,222–224). Overall, no reliable single predictor for preeclampsia exists and the clinical tools

are restricted to subjective symptoms with poor specificity and sensitivity. A combination of biophysical, biochemical and ultrasound markers may provide a more useful predictive tool than a test of either component alone, however, further research is necessary to identify additional combinations of markers that may predict the occurrence of preeclampsia since to date no biomarker combination has performed well enough for clinical application.

Gestational diabetes mellitus (GDM) is a common metabolic condition of pregnancy associated with several pregnancy complications and with established beneficial effect of treatment. However, the controversy and many different approaches to screening and diagnosis present a challenge to scientific advancement in this area (225–227). In current clinical practice, a variety of tests and methods are used in the screening of GDM, including the random glucose measurement, fasting glucose measurement and a glucose challenge test (blood glucose measurement one hour after ingestion of 50 g of glucose) (228). Until today, there is no agreement on which screening test is most appropriate, due to estimates of accuracy and costs of the tests reported in the literature vary. There is also a debate on which women should be tested as there are recommendations for the inclusion of all pregnant women (universal screening), while other recommend the exclusion of all women except those at risk (selective screening). For instance, international bodies such as the American Diabetes Association, advocate the use of selective screening based on clinical factors for GDM such as age>25 years, obesity and previous GDM, to identify women at risk for GDM (229–231). Opponents of this selective strategy criticize the use of risk factors to select women for screening, since this strategy have limited accuracy and fails to identify over one-third of cases of GDM and therefore, universal screening is widely

recommended. Nevertheless, the sensitivity and specificity are both considered to be low, leaving women with GDM undiagnosed on the one hand, and leading to unnecessary testing in healthy women on the other (228,232–235).

The identification of women at high risk of developing GDM who would benefit from targeted preventative measures, has resulted in the investigation of new biomarkers with a possible use of them as predictors. An extensive body of research have investigated potential biomarkers in the prediction of GDM, however they have reported conflicting and inconsistent results, either because many of the factors being interlinked and sharing similar metabolic pathways or because of the lack of consistency in the diagnostic criteria of GDM between studies (236–239). Findings from previous systematic reviews and meta-analyses that evaluated the relationship between various biomarkers, including inflammatory markers, adipokines, and endothelial function, revealed that decreased adiponectin is an independent predictor of GDM. Increased levels of TNF-α and leptin may also be predictive, but further prospective studies are required to firmly establish their role independent of BMI and insulin resistance (238–240). Additionally, previous meta-analysis has exposed that triglyceride levels are markedly elevated throughout the course of pregnancy in women with GDM, however, further research was recommended to establish the potential clinical utility for identifying women at risk for subsequently developing GDM (241).

Due to the lack of predictive ability of a single marker, it is now recognized that combinations of risk factors and maternal or placental markers reflecting pathophysiological pathways implicated in GDM in a multivariate logistic regression

model may have the most potential application to enhance the prediction for GDM (238,242). Early detection and prediction of women at risk of GDM would allow streamlined antenatal care and allocation of targeted dietary and lifestyle interventions to reduce the development of GDM, which consequently would improve pregnancy outcomes (243,244). There are several published predictive models for GDM that combined various biomarkers with maternal clinical risk factors, achieving good sensitivity and specificity for prediction of GDM, however they are not widely used in routine clinical practice (245,246). For instance, a previous simple risk prediction tool based on previous GDM, family history of type 2 diabetes mellitus, high risk ethnicity, age and BMI, achieved a sensitivity of 61.3% and specificity of 71.4% for differentiating women according to their risk of GDM (247). Further multi-parametric risk prediction models that investigated the potential of prediction of GDM by a combination inflammatory and other biomarker have shown incremental sensitivity and specificity and their translation to clinically important improvements in prediction is debatable, with very few implementation studies performed (238,248–250). A recent systematic review evaluating the quality and characteristics of seventeen studies describing first-trimester prediction models for GDM revealed various shortcomings on the model development studies, since only few have been externally validated and most showed moderate to low methodological quality. External validation was recommended to enhance generalizability and assess their true value in clinical practice (246). In summary, no reliable single predictor for GDM exists and the clinical practice is restricted to examination of maternal history with limited specificity and sensitivity. A combination of risk factors and maternal or placental markers may provide a more useful predictive strategy; however, further research is necessary to determine whether predictive models can be further improved with the addition of

novel biomarkers implicated in the pathophysiology of GDM. Early risk stratification by prediction modeling might offer opportunities to improve care for those women at high risk of developing GDM. Such work should therefore be prioritized, especially at a time of rising obesity levels, which will substantially increase the number of women with this condition.

2.9 Goal and Significance of Research

The long-term goal of this research study is to reduce the occurrence of preeclampsia and gestational diabetes. As a step towards this goal, by using an umbrella review methodology, this dissertation aims to systematically assess the evidence across published systematic reviews and meta-analyses on the risk factors and/or interventions for preeclampsia and gestational diabetes, evaluate whether there is evidence for diverse biases in this body of literature, and finally, pinpoints which of the previously studied risk factors or interventions present the strongest consistent evidence. Ultimately, better understanding of the evidence on an entire field across many systematic reviews and meta-analyses, can be important for public health, not only for understanding the reliability of an evidence-base but also serves as the foundation for clinical and public health recommendations.

Three original studies have conducted to examine this goal: an umbrella review of systematic reviews and meta-analyses of observational studies on genetic and non-genetic risk factors for preeclampsia, an umbrella review of meta-analyses and systematic reviews of randomized trials of interventions for preventing preeclampsia, and lastly, an umbrella review of meta-analyses of observational studies on non-genetic risk factors for gestational diabetes. To our knowledge, no previously

published research has attempted such a comprehensive assessment of risk factors and/or interventions for preeclampsia or gestational diabetes. Such studies will be highly appreciated by the scientific community because of the importance of the topic, since preeclampsia and gestational diabetes are considered major causes of maternal and fetal morbidity and mortality worldwide. Furthermore, both diseases are not only increase the risk for maternal and fetal complication in pregnancy but are also associated with long-term risks, such cardiovascular disease in both mother and child.

Findings from this study can provide greater understanding of critical issues related to screening, prediction, prevention and treatment of both pregnancy-related diseases, which can be translated into evidence-based-medicine actions, such as improvement of risk stratification tools, and establishment of modifiable risk factors. In addition, this research could be useful, to illustrate new mechanistic understanding of preeclampsia and gestational diabetes, new clinical ramifications or research needs and guide the design of future preventive interventions measures. Hence, we believe that this study will be the beginning of a long-term initiative that addresses all the scenarios, from the assessment of the burden to the identification of appropriate implementation approaches.

Chapter 3 – Meta-Research Methods

Each year several million new research papers are published whereas at the same time the number of systematic reviews and meta-analyses is growing rapidly. Meta-analysis remains a gold standard for evidence-based decision-making and important research design for guiding medical practice, health policy and health technology assessments (43,44). However, as previously explicated, of the published systematic reviews and meta-analyses, around 1 in 6 have misleading estimates, mostly in genetic literature published by China, and probably another 1 in 3 meta-analyses are unnecessary and/or conflicted of other research types. Of the remaining, approximately half have serious methodological flaws and many others are correct but with weak or insufficient evidence to inform decision making. Only a very small minority are both non-misleading and truly informative meta-analyses (9).

As previously mentioned, systematic reviews and meta-analyses draw strength by combining evidence from many primary studies that have addressed a similar research question. However, even if perfectly done with perfect data, a single meta-analysis, that addresses one treatment comparison for a single outcome may offer a limited view of the evidence. If there is only one choice for treatment, one outcome of interest and faultless results, this meta-analysis may assist for decision making. However, usually there are many treatments options, many outcomes to consider and research is imperfect (9). Meta-analyses of observational associations suffer from limitations too. Most meta-analyses that combine evidence from observational studies focus on studying the association of one or at most a few putative risk factors for a specific

outcome. Hence, there is an enormous volume of published studies on risk factor epidemiology and usually a large number of studied risk factors for a particular outcome (23).

While meta-analysis is considered to provide high quality evidence, it would be extremely important to detect different types of bias in favor of statistically significant results that create associations that do not exist, with the goal to decrease the number of wrong decision making in everyday clinical practice and public health (23). There are now several tools (e.g. funnel-plot asymmetry tests) available with which meta-analysts can assess the potential magnitude of publication bias, however, these tests may be affected by any type of significance-chasing bias and they may also be affected by a wide variety of other issues, including genuine diversity across the study-specific effects such as the presence of heterogeneity or the lack of studies with significant results (251–253). The evaluation of such biases in each single study is difficult as the data are usually limited, unless designs and analysis plans are registered a priori (23). It is easier therefore, to evaluate this type of biases across multiple studies performed on the same question with goal to gain insight into the average bias in the field (70,254).

Meta-research, also known as “research on research”, is an evolving scientific discipline that investigates research practices with the ultimate goal of evaluate and improve evidence-based practices (24). Meta-research uses both theoretical and empirical investigation with analytical and computational methods to study how research is done and where improvements can be made with objective to improve the scientific enterprise. It was categorized into five major areas of interest: Methods,

Reporting, Reproducibility, Evaluation, and Incentives, which these correspond, respectively, with how to perform, communicate, verify, evaluate, and reward research (24). Given the types of questions addressed, meta-research interfaces with many other established disciplines, including but are not limited to, epistemology, psychology, statistics, informatics, evidence-based medicine, research synthesis methods (e.g. meta-analysis), organizational and operations research, ethics, policy research, and behavioural economics (24). The primary remit of meta-research is not a single meta-analysis that synthesizes evidence on multiple studies on a specific question but the combination of evidence from multiple meta-analyses on multiple topics, which offer insights about how common and how consistent certain biases are across a large field or multiple fields. This emphasis on the broader picture is the key characteristic of meta-research (24).

In the era of meta-research, several research studies have generated, with the terminology around these studies yet to be unclear, with various names attributed to many times the same process (25). Overviews or reviews, overviews of systematic reviews, systematic reviews of systematic reviews, umbrella reviews, umbrella reviews of systematic reviews, systematic umbrella reviews, multiple treatments meta-analysis, meta-analysis of meta-analyses and meta-epidemiological studies are some of the terms used to describe certain types of one study which collects and combines studies which in turn have collected and combined studies (25). Irrespective of their name, all of these types of reviews have a defining feature in common: a systematic review is the principal and often sole “study type” that is considered for inclusion. Therefore, we can say that it is a second level stage of combining studies, or a third level stage of analyses (25).

In contrast to a systematic review or a meta-analysis which are limited to one treatment comparison or even one outcome, a meta-research study combines all of the data from all comparisons together can provide an overall picture. This helps to recognize the relative merits of all available interventions, and consequently be more useful for health technology assessments, evidence-based guidelines and medical decision-making. Likewise, this research methodology has parallel application to non-randomized research. Given that for many diseases, there can be hundreds of proposed associations (genetic, nutritional, environmental), such studies can systematize and summarize the totally evidence to keep track of where we stand and what to make of the torrents of data on postulated risk factors. The synthesis of such complex information from many systematic reviews and multiple meta-analyses requires rigorous and systematic methods and is not something that can be performed lightly by a subject-matter expert based on subjective opinion alone (22).

3.1 Overviews of Reviews and Meta-Epidemiologic Studies

Overviews of reviews are a recent development in research synthesis with a developing still methodology (25). Although initially entitled as umbrella reviews, they have been subsequently being referred as meta-reviews, overviews of systematic reviews, reviews of reviews and systematic review of systematic reviews (255). They are defined as reviews that gather information from individual systematic reviews relevant to a single health problem using explicit and systematic methods examining different interventions for the same condition or different outcomes for the same intervention in the same condition or the same intervention for different conditions or populations or finally adverse effects from the same intervention across multiple conditions (35,256). In their majority, overviews of reviews are narrative or qualitative reviews of their

systematic reviews reporting on the findings and summary estimates from the meta-analysis if occurred (257).

Meta-epidemiologic studies can be seen as overviews of reviews with a non-clinical first topic and usually focus on given methodological aspects (e.g. they may focus on finding issues or small study effects). The meta-epidemiology is based on the combination of epidemiology and meta-analysis. Meta-epidemiology attempts to describe the distribution of research evidence for a specific question, examine heterogeneity and associated risk factors, identify and control bias between studies and summarize research evidence (258,259). It is not therefore a simple meta-analysis or narrative review, but a sort of meta-review (25). It has been recognized as another epidemiological research methodology that controls meta-confounders, similar to traditional epidemiological research methodology that controls confounding variables (258,260,261). The interest of meta-epidemiology is to control potential biases in previous quantitative systematic reviews and draw appropriate inferences. With this background, diverse methods, such as meta-regression, imputation, informative missing odds ratio, two statistical models, and others were attempted, and the term meta-epidemiology was introduced (262).

3.2 Umbrella Reviews

It has been suggested that one of the solutions for limited utility of systematic reviews is perform systematic reviews of systematic reviews, also known as umbrella reviews or systematic umbrella reviews. An umbrella review has been reported as one of the four types of “next-generation” systematic reviews that may raise the bar and help shape a new generation of more reliable evidence synthesis (11). They are not

necessarily brand-new ideas, but in the current circumstances of uncontrollable overproduction and unchecked quality of systematic reviews and meta-analyses, they have a fresh opportunity for impact (11). Such reviews emerged only recently, and their number is increasing since their content is an attractive way to distil and translate large amounts of evidence. A simple search of PubMed conducted on June 29, 2017 indicates that there were 239 hits for a phrase "overview of systematic reviews", and 93 hits for "umbrella review", and that number of those studies started increasing in year 2010 (263).

The principle reason for the conduct of an umbrella review is to summarize the evidence from multiple research syntheses. Particularly, umbrella reviews allow a higher-level synthesis of large amount of evidence from all systematic reviews and meta-analyses on a given topic and may be based on outcomes, risk factors or interventions, e.g. all treatments for a condition or set of conditions; or all risk factors assessed for some disease or all associations that a specific risk factor has been evaluated for in relationship to a variety of outcomes/diseases (11). In theory, umbrella reviews may also encompass systematic reviews and meta-analyses on data of diagnostic, prognostic and predictive tests, if these are pertinent to consider in the overall management of a disease, in addition to just treatment decisions (22). It is important to note that the principal aim of an umbrella review is to provide an overview of the range and validity of the reported associations of existing research syntheses related to a given topic or question, and not to re-synthesize, for example, with meta-analysis or meta-synthesis, the results of existing systematic reviews or meta-analyses.

Umbrella reviews are conducted to provide an overall examination of the body of information that is available for a give topic, and to compare the results of published systematic reviews (255). The wide picture obtainable from the conduct of an umbrella review is ideal to highlight whether the evidence base around a topic is consistent or contradictory, and to explore the reasons for the findings (264). Umbrella reviews may permit understanding of the amount and credibility of the evidence, identification of research gaps and weaknesses as well as the main sources of heterogeneity, bias and quality features that affect the credibility of the results in a large research field (11,22). As it brings together comparisons of a large number of existing systematic reviews and meta-analyses into one accessible and usable document, this would ultimately contribute to the improvement of overall healthcare, which this is the background and aim of the meta-research emergence (25,262).

The methods of the umbrella review are standardized, and we will follow the same principles as previously described in published umbrella reviews conducted on various fields of research (265–269). For practical reasons, in the next section we present a stepwise description of the tasks performed to summarize and evaluate the evidence using the umbrella review methodology. Most of the features described here are not unique for the operational conduct of an umbrella review, and researchers familiar with the conduct of a systematic review will immediately identify the similarities in process and methods used. Despite these similarities, there are several important features for researchers undertaking an umbrella review worth noting.

3.3 Umbrella Review Methodology

As previously mentioned, the main goal of this thesis is to systematically overview, analyze and summarize evidence across the published literature of clinical identities with a large impact on the perinatal epidemiology field, namely preeclampsia and gestational diabetes, and map whether any interventions or fields of risk factors include convincing evidence to support their results. As outlined earlier, for this PhD research we follow the same methodology principles as previously described in published umbrella reviews, nevertheless, minor differences may occur among the three umbrella review studies as the methodology and presentation of the results is align to the umbrella review question.

3.3.1 Eligibility Criteria

Depending on the umbrella review question meta-analyses and systematic reviews of observational or international studies were identified and retained if they included at least three studies in which information was provided per included study on a measure of association and its standard error and on the number of cases and the total population. We did not apply any language restrictions in the selection of eligible studies and we included only systematic reviews and meta-analyses of epidemiological studies in humans. If an article presented separated meta-analyses on other medical diseases including the outcome of interest, we only extracted information on the latter. When more than one meta-analysis on the same research question was eligible, the meta-analysis with the largest number of component studies with data on individual studies' effect sizes was retained for the main analysis, but we kept a record of other published meta-analyses focused on the same risk factor. We excluded narrative reviews, letters to the editor, systematic reviews without a quantitative synthesis of

data, and studies in which risk factors were used for screening, diagnostic, or prognostic purposes. We also did not include the older version of two meta-analyses that were published by the same authors on the same intervention or risk factor when there was only a 2–3 years difference between the two versions.

3.3.2 Literature Search and Data Extraction

The search strategy for an umbrella review should aim to identify all research syntheses relevant to the review question. Two researchers independently searched PubMed, ISI Web of Science and Cochrane Library to identify systematic reviews and meta-analyses of observational studies or interventional studies. For example, the search strategy for the “Umbrella review of genetic and non-genetic risk factors for preeclampsia” we used the keywords ("pre-eclampsia" OR "preeclampsia") AND ("systematic review" OR "meta-analysis"). We also systematically searched PubMed and GWAS central to identify genome-wide association studies (GWAS) examining genetic associations with the particular outcome of interest. All identified publications underwent a parallel, three-step review of title, abstract, and full text based on predefined inclusion and exclusion criteria. We also screened the references of the retrieved articles for possible eligible papers.

To minimize risk of bias in the umbrella review process, data extraction was performed independently by two investigators, and in case of discrepancies, the final decision was reached by discussion or a third investigator, when necessary. From each eligible meta-analysis, we extracted information on the first author, year of publication, the examined risk factor or intervention administered, the number of studies included, the number of cases and controls for each study or total number of participants per

treatment arm and events in each arm in case of a clinical trial, and the study-specific relative risk estimates (risk ratio, odds ratio) or standardized mean differences along with the corresponding confidence intervals (CI). Also, we recorded the reported summary meta-analytic estimates using both fixed and random effect methods along with the corresponding 95% confidence intervals. Also, we noted whether the selected systematic reviews and meta-analyses applied any criteria to evaluate the quality of the included studies.

3.3.3 Assessment of Summary Effect and Heterogeneity

For each meta-analysis, we estimate the summary effects and its 95% confidence interval by using both fixed and random effect models (94,270). Additionally, we calculate the 95% prediction intervals (PI) for the summary random effects estimates, which further account for between-study heterogeneity and indicates the uncertainty for the effect that would be expected in the new study observing the same association (271,272). The 95% PI shows where the true effects are for 95% of the studies from the population of studies that are synthesized or similar studies that might be done in the future. For the largest study of each meta-analysis, we calculated the standard error (SE) of the effect size; we examined whether the standard error was less than 0.10 and whether the largest study presented a statistically significant effect. In a study with SE of less than 0.10, the difference between the effect estimate and the upper or lower 95% confidence interval is less than 0.20 (i.e. this uncertainty is less than what is considered a small effect size).

We assessed heterogeneity between studies, and we reported the P value of the χ^2-based Cochran Q test and the I^2 metric for inconsistency, which could reflect either

diversity or bias. I^2 ranges between 0% and 100% and is the ratio of between-study variance over the sum of within and between-study variances (273). Values exceeding 50% or 75% are usually considered to represent large or very large heterogeneity, respectively. Its confidence intervals were calculated as per Ioannidis et al. (274).

3.3.4 Assessment of Small Study Effects

We evaluated whether there is evidence for small study effect (i.e. if small studies tend to give higher risk estimates than large studies). Small study effects can indicate publication and other selective reporting biases, but they can also reflect genuine heterogeneity, chance, or other reasons for differences between small and large studies (275). We used the regression asymmetry test proposed by Egger for this assessment (276). A P value <0.10 accompanied by a more conservative effect in larger studies was considered evidence for the existence small-study effects.

3.3.5 Evaluation of Excess Statistical Significance

The excess significant test was performed to evaluate whether there is a relative excess of formally significant findings in published literature due to any reason (such as publication bias, selective reporting of outcomes or analyses). The number of expected positive studies is estimated and compared against the number of observed number of studies with statistically significant results (P<0.05) by using the χ^2 test (82). A binomial test was used to evaluate whether the number of positive studies in a meta-analysis is too large according to the power that these studies have to detect plausible effects at α=0.05. A comparison between observed vs expected is performed separately for each meta-analysis and it is also extended to groups of many meta-analyses after summing the observed and expected from each meta-analysis. The expected number

of significant studies for each meta-analysis is calculated by the sum of the statistical power estimates for each component study (82). The power of each component study was estimated using the fixed effects summary, the random effects summary, or the effect size of the largest study (smallest SE) as the plausible effect size (277). The power of each study was calculated with an algorithm using a non-central *t* distribution (278). Excess statistical significance for single meta-analyses was claimed at $P<0.10$ (one-sided $P<0.05$, with observed > expected as previously proposed) given that the power to detect a specific excess will be low, especially with few positive studies (82). We classified risk factors or interventions categories based on biological pathways or types of exposures involved. We examined excess statistical of significant separately in each of these categories as selective reporting bias may arise in different domains of research.

3.3.6 Methodological quality

We did not conduct a qualitative assessment of component studies as it was beyond the scope of the umbrella review methodology, and this should be performed in the original systematic reviews and meta-analyses through standardized tools, such as Newcastle–Ottawa scale. However, we recorded whether the authors of the original meta-analyses have performed any quality assessment of the synthesized studies.

3.3.7 Assessment of Epidemiologic Credibility of Non-Genetic Associations

The associations that had the strongest validity and no suggestive information of bias were identified and graded based on a set of methodological criteria, which have been previously applied in other research fields (266–268,279,280). We used a ranking system to grade the evidence from systematic review and meta-analyses in terms of

the significance of the summary effect, the 95% prediction interval, presence of large heterogeneity, small study effects, and excess significance bias. Specifically, we characterized as *convincing* the associations fulfilling the following criteria: had significant effect under the random-effects model at $P<10^{-6}$, were based on evidence from more than 1000 cases, the between-study heterogeneity was not large ($I^2<50\%$), the 95% PI excludes the null value and had no evidence of small-study effects and excess of significance bias. Additionally, the associations with more than 1000 cases, a significant effect at $P<10^{-6}$ and nominally statistically significant effect present at the largest study were characterized as *highly suggestive*. We considered as *suggestive* the associations with significant effect at $P<10^{-3}$ and more than 1000 cases. The rest of statistically significant associations at $P<0.05$ under random-effects model were graded as *weak* associations.

3.3.8 Epidemiological Credibility of Genetic Associations

We used the Venice criteria to evaluate the epidemiological credibility of all significant genetic associations (281). Credibility was defined based upon the grade (A=strong, B=moderate or C=weak) of three categories: amount of evidence, replication of the association, and protection from bias. Amount of evidence was graded by the sum of test alleles or genotypes among both cases and controls in the meta-analysis; ('A' for over 1,000, 'B' for 100 to 1,000, and 'C' for less than 100). Replication of the association was graded as "A" if there was an extensively replicated study supported by at least 1 well conducted meta-analysis, "B" if it was a well-conducted meta-analysis with some methodological limitations and "C", if there was no independent replication, failed replication or flawed meta-analysis. Assessment of protection from bias included consideration of the magnitude of the association,

heterogeneity statistic and findings from tests for selective reporting biases (test for small-study effects and excess statistical significance). According to these criteria, the credibility level of the cumulative evidence was defined as high (A grades only), low (one or more C grades) or intermediate (all other combinations) (281).

3.3.9 Epidemiological Credibility of Interventional Evidence

We used a ranking system to grade the evidence from meta-analyses of RCTs in terms of the significance of the summary effect ($p<0.001$, $0.001\leq p<0.05$, $p\geq0.05$), 95% prediction interval (excluding the null or not), and presence of large heterogeneity (I^2 >50%), small study effects ($p>0.10$), and excess significance ($p<0.05$). Studies that reported a p-value of less than 0.001, had a 95% prediction interval not including the null, had no evidence of small-study effects or no evidence of excess significance, and did not have large heterogeneity were considered as representing robust evidence of effectiveness of interventions (Class I). Meta-analyses that had a p-value less than 0.001 and the largest study reporting a significant effect were considered to have the next best quality of evidence (Class II). Finally, meta-analyses with only a p-value of less than 0.05 were classified as quality of evidence Class III.

3.3.10 Presentation of the Results

In this section we provide context to the results and sufficient descriptive detail for the reader about the inclusion of the research syntheses into the umbrella review, the relevance of included research syntheses to the umbrella review question and the evidence base they offer to the research question. As the aim of the umbrella review is to present a summary of existing research syntheses relevant to a particular topic or question and not any further "synthesis" of the results of these publications, to this, the

results of all included studies are presented to the reader to allow ready and easily interpretable overview of the findings and gain a clear understanding of a broad topic area. Construction of multiple tables is often necessary to clearly present all the data collected from reviews (282,283), therefore, well-constructed tables will facilitate analysis as they make patterns in the data easier to detect.

3.4 Limitations of Umbrella Review Methodology

Umbrella reviews provide an up-to-date overview on a specific research topic by considering systematic reviews and meta-analyses, which represent the highest level of evidence to inform decision-making. However, the umbrella review approach has some limitations that should be considered when interpreting their findings. As with other forms of evidence synthesis, the utility of umbrella reviews will be largely dependent on the availability of published systematic reviews and meta-analyses. Hence, this approach may favour the selection of more commonly and readily studied risk factors or interventions, since they are more likely to be included in a systematic review or a meta-analysis. In addition, for some factors that are difficult or uncommon to study, the current standardized methodological assessment of the epidemiological credibility using a wide range of tests and criteria is unlikely to be remarkable, given the limited data. Even though, umbrella reviews adopted credibility assessment criteria, which were based on already established tools, still, none of the components of these criteria provides definitive proof of lack of reliability, but they cumulatively suggest that the results are susceptible to bias and uncertainty.

Several limitations of the umbrella review approach are largely reflected by limitations in the original studies. Because systematic reviews and meta-analyses included

primary studies with differences in design, population, outcome or exposure definitions and other basic characteristics, large heterogeneity may be worrisome. In addition, it is possible that the results of studies included in a meta-analysis to have previously been standardized (e.g. cleaned or made to follow consistent definitions and adjustments). Such standardization efforts are likely to reduce, if anything, inconsistency and selective reporting bias, whereas the last, may be more prominent in the primary study reports. Another limitation of an umbrella review is the use of existing published systematic-reviews and meta-analyses, and their results may depend on choices made about what estimates to select from each primary study and how to represent them in the meta-analysis. Likewise, because umbrella reviews depended on the original meta-analyses quality assessment, and ultimately the studies that they include, it is possible that deficiencies in the methodological quality at each level can compromise the results and conclusions of an umbrella review.

Chapter 4 – Genetic and non-genetic risk factors for preeclampsia: An umbrella review of systematic reviews and meta-analyses of observational studies

4.1 Abstract

Objective: To summarize evidence from the literature on the genetic and non-genetic risk factors associated with preeclampsia, assess the presence of statistical biases and identify risk factors with robust evidence.

Methods: We searched PubMed and ISI Web of Science from inception to October 2016, to identify systematic reviews and meta-analyses of observational studies examining associations between genetic and non-genetic risk factors for preeclampsia. For each meta-analysis we estimated the summary effect size by random-effects and fixed-effects models, the 95% confidence interval and the 95% prediction interval. We estimated the between-study heterogeneity expressed by I^2 (considering above 75% as very large), evidence of small-study effects (large studies had significantly more conservative results than smaller studies and evidence of excess significance bias (too many studies with statistically significant results).

Results: Fifty-eight eligible meta-analyses were identified, which included 1466 primary studies and provided data on 130 risk factors associated with preeclampsia, covering a very wide range of risk factors: co-morbid diseases, genetic factors, exposure to environmental agents and a range of biomarkers. Sixty-five (50%) associations had nominally statistically significant findings at $P<0.05$, while sixteen (12%) were significant at $P<10^{-6}$. Sixty-five (50%) associations had large or very large heterogeneity. Evidence for small-study effects and excess significance bias was found in ten (8%) and twenty-six (20%) associations, respectively. Oocyte donation vs spontaneous conception (OR 4.33, 95% CI: 3.11-6.03) had >1000 cases, 95% prediction intervals excluding the null, not suggestive of large heterogeneity ($I^2<50\%$), small-study effects (P for Egger's test>0.10), or excess of significance (P>0.05).

Across the statistically significant genetic risk factors (P<0.05), only PAI-1 4G/5G (recessive model) polymorphism was supported with strong evidence for a contribution to the pathogenesis of preeclampsia. Eleven factors (serum iron level, PAPP-A, chronic kidney disease, polycystic ovary syndrome, mental stress, bacterial & viral infections, cigarette smoking, oocyte donation vs assisted reproductive technology, obese vs normal weight women, severe obese vs normal weight women and primiparity) presented highly suggestive evidence for preeclampsia.

Conclusions: A large proportion of meta-analyses of genetic and non-genetic risk factors for preeclampsia have caveats, which threaten their validity. Oocyte donation vs spontaneous conception and PAI-1 4G/5G polymorphism (recessive model) show the strongest consistent evidence.

4.2 Introduction

Preeclampsia (PE) is a severe pregnancy-associated disease, characterized by the occurrence of hypertension and proteinuria in previously healthy women after the 20th weeks of gestation. PE affects approximately 2-8% of all pregnancies and is associated with substantially higher maternal and fetal morbidity and mortality worldwide (284,285). The clinical spectrum of PE varies, from mild, which is characterized by a moderate increase in blood pressure and proteinuria, to the most severe outcome of eclampsia, described by seizures as a sign of damage of the cerebral vessels, and HELLP syndrome (Hemolysis, Elevated Liver enzyme, Low platelets), which significantly threatens the lives of pregnant women and their fetuses (286). The true etiology of PE remains an issue of debate, and generates uncertainty on prediction, prevention and treatment., occurring as interplay between genetic and non-genetic factors (287,288).

Numerous meta-analyses and systematic reviews have claimed that several environmental, biological and genetic risk factors are associated with PE risk. If causal, these associations might be useful for the accurate prediction and diagnosis of this condition in early pregnancy, which would allow a timely allocation of screening resources and prevention of maternal and fetal complications (289–291). In addition, preventive measures such as aspirin administration in high risk women appear more likely to be beneficial if started earlier in pregnancy during the first trimester or even preconception (292,293). Nevertheless, there is a possibility that some observed associations in the literature do not reflect a genuine association but include different types of bias in favor of positive statistically significant associations (20). The pursuit of positive results may be generated with several different mechanisms, such as

selective analyses, outcome bias and fabrication bias (19,82). These biases can cause either false published findings (19) or inflated effects (18).

To our knowledge, this is the first attempt to summarize the evidence from existing meta-analyses on genetic and non-genetic risk factors for PE. We aim to summarize evidence from meta-analyses on the risk factors that have been associated with PE, evaluate whether there are hints of biases in this literature and how they manifest, and finally identify which of the previously studied associations represent robust epidemiologic evidence.

4.3 Methods

The concept of umbrella review

We conducted an umbrella review, which is a systematic collection and evaluation of multiple systematic reviews and meta-analyses performed on a specific research topic (22). An umbrella review brings together comparisons of a large number of existing systematic reviews and meta-analyses on risk factors into one accessible and usable document (22,25). The methods of the umbrella review are standardized and in this work we follow state-of-the-art approaches as previously published umbrella reviews on risk factors and various outcomes (265–268).

Literature search

Two researchers (KG and SP) independently searched PubMed and ISI Web of Science from inspection to October 8, 2016, to identify systematic reviews and meta-analyses of observational studies examining associations between risk factors and PE. The search strategy used the keywords ("pre-eclampsia" OR "preeclampsia") AND ("systematic review" OR "meta-analysis"). Initially, the title and abstract of each these articles were examined and then we retrieved potentially eligible articles for full text evaluation. We also systematically searched PubMed to identify genome-wide association studies (GWAS) examining genetic associations with PE. Any discrepancies were resolved with discussion.

Eligibility criteria and data extraction

Articles were eligible if the authors had performed a systematic search to identify pertinent studies that examined the association between various risk factors and PE. The full text of potentially eligible articles was scrutinized independently by two

investigators (KG, SP). Meta-analyses or systematic reviews were retained if they included at least three studies in which information was provided per included study on a measure of association and its standard error between the risk factor and PE and on the number of cases/population. We excluded studies in which risk factors were used for screening, diagnostic, or prognostic purposes or meta-analyses that examined PE as a risk factor for other medical conditions. We did not apply any language restrictions in the selection of eligible studies. When more than one meta-analysis on the same research question was eligible, the meta-analysis with the largest number of component studies with data on individual studies' effect sizes was retained for the main analysis.

Data extraction was performed independently by two investigators (KG, SP), and in case of discrepancies, the final decision was reached by discussion or a third investigator, when necessary (EE). From each eligible meta-analysis, we extracted information on the first author, year of publication, the examined risk factors, the number of studies included, the study-specific relative risk estimates (risk ratio, odds ratio) or standardized mean differences along with the corresponding confidence intervals (CI). Also, we recorded the reported summary meta-analytic estimates using both fixed and random effect methods along with the corresponding confidence intervals and the number of cases and controls for each study. We noted whether the selected meta-analyses applied any criteria to evaluate the quality of the included observational studies.

Assessment of summary effect and heterogeneity

For each meta-analysis, we estimated the summary effects and its 95% confidence interval by using both fixed and random effect models (94,270). Additionally, we calculated the 95% prediction intervals (PI) for the summary random effects estimates, which further account for between-study heterogeneity and indicates the uncertainty for the effect that would be expected in the new study observing the same association (271,294). For the largest study of each meta-analysis, we calculated the standard error (SE) of the effect size, and we examined whether the standard error was less than 0.10 and whether the largest study presented a statistically significant effect. In a study with SE of less than 0.10, the difference between the effect estimate and the upper or lower 95% confidence interval is less than 0.20 (i.e. this uncertainty is less than what is considered a small effect size).

We assessed heterogeneity between studies, and we reported the P value of the χ^2-based Cochran Q test and the I^2 metric for inconsistency, which could reflect either diversity or bias. I^2 ranges between 0% and 100% and is the ratio of between-study variance over the sum of within and between-study variances (273). Values exceeding 50% or 75% are usually considered to represent large or very large heterogeneity, respectively. Confidence intervals were calculated as per Ioannidis et al. (274).

Assessment of small study effects

We evaluated whether there is evidence for small study effect (i.e. if small studies tend to give higher risk estimates than large studies). Small study effects can indicate publication and other selective reporting biases, but they can also reflect genuine heterogeneity, chance, or other reasons for differences between small and large studies

(275). We used the regression asymmetry test proposed by Egger for this assessment (276). A P-value <0.10 accompanied by a more conservative effect in larger studies was considered evidence for the existence small-study effects.

Evaluation of excess statistical significance

The excess of statistical significance test was performed to evaluate whether there is a relative excess of formally significant findings in published literature due to any reason. The number of expected positive studies is estimated and compared against the number of observed number of studies with statistically significant results (P<0.05) (82). A binomial test was used to evaluate whether the number of positive studies in a meta-analysis is too large according to the power that these studies have to detect plausible effects at α=0.05. A comparison between observed vs expected is performed separately for each meta-analysis and it is also extended to groups of many meta-analyses after summing the observed and expected from each meta-analysis. The power of each component study was estimated using the fixed effects summary, the random effects summary, or the effect size of the largest study (smallest SE) as the plausible effect size (277). The power of each study was calculated with an algorithm using a non-central *t* distribution (278). Excess statistical significance for single meta-analyses was claimed at P<0.10 (one-sided P<0.05, with observed > expected as previously proposed) given that the power to detect a specific excess will be low, especially with few positive studies (82).

We classified risk factors into categories based on biological pathways or types of exposures involved: biomarkers, environmental factors, genetic markers, diseases and disorders, supplementation, infections and other risk factors. We examined excess

statistical of significance separately in each of these categories as selective reporting bias may arise in different domains of research. The excess of statistical significance test was conducted separately for meta-analyses with I^2 values less than or equal to 50% and greater than 50%, because values above 50% are typically reflected evidence of large heterogeneity beyond chance (295).

Grading of non-genetic and genetic associations

We characterized as convincing the non-genetic associations fulfilling the following criteria: had significant effect under the random-effects model at $P<10^{-6}$, were based on evidence from more than 1000 cases, the between-study heterogeneity was not large ($I^2<50\%$), the 95% PI excludes the null value and had no evidence of small-study effects and excess of significance bias. Additionally, associations with more than 1000 cases, a significant effect at $P<10^{-6}$ and nominally statistically significant effect present at the largest study were characterized as highly suggestive. We considered as suggestive the associations with significant effect at $P<10^{-3}$ and more than 1000 cases. The rest of statistically significant associations at $P<0.05$ under random-effects model were graded as weak associations.

We used the Venice criteria to evaluate the epidemiological credibility of all significant genetic associations (281). Credibility was defined based upon the grade (A=strong, B=moderate or C=weak) of three categories: amount of evidence, replication of the association, and protection from bias. Amount of evidence was graded by the sum of test alleles or genotypes among both cases and controls in the meta-analysis; ('A' for over 1,000, 'B' for 100 to 1,000, and 'C' for less than 100). Replication of the association was graded as "A" if there was an extensively replicated

study supported by at least 1 well conducted meta-analysis, “B” if it was a well-conducted meta-analysis with some methodological limitations and “C”, if there was no independent replication, failed replication or flawed meta-analysis. Assessment of protection from bias included consideration of the magnitude of the association, heterogeneity statistic and findings from tests for selective reporting biases (test for small-study effects and excess statistical significance). According to these criteria, the credibility level of the cumulative evidence was defined as high (A grades only), low (one or more C grades) or intermediate (all other combinations) (281).

All authors had full access to all of the data in the study. Statistical analysis and the power calculations were performed in STATA version 14 (STATA Corp, College Station, TX).

4.4 Results

Description of eligible meta-analyses

The search identified 634 items, of which 535 were excluded after the title and abstract review. Of the remaining 99 articles that entered the full-text review, 8 articles did not report the appropriate information for the calculation of excess of statistical significance (either because the total sample size was missing or the study-specific relative risk estimates were missing), one article was a pooled analyses of cohort studies, two articles included only 2 component studies, and 18 articles excluded because a larger systematic review or meta-analysis investigating the same risk factor was available (Figure 4.2). Therefore, 71 articles were analyzed, of which 13 were systematic reviews without any quantitative component and 58 were meta-analyses. The 58 eligible meta-analyses (288,296–322,182,323–351), included data on 130 comparisons in seven broad areas (biomarkers [n=27 comparisons], environmental factors [n=6 comparisons], genetic markers [n=66 comparisons], diseases and disorders [n=8 comparisons], supplementation [n=1 comparisons], infections [n=3] and other risk factors [n=19 comparisons]).

The characteristics of the included meta-analyses are shown in Table 4.1. Based on the study design of the synthesized studies that examined non-genetic associations, we had 7 (20%) meta-analyses synthesizing retrospective case-control data only, 3 (9%) meta-analyses that included prospective data (cohort studies) and 25 (71%) of studies including both types of data, noted as mixed. Regarding the genetic association studies, 15 (65%) meta-analyses synthesized case-control data, 7 (30%) of studies used both types of data (case-control and cohort data), and 1 (4%) meta-analysis that included only cohorts.

There were 3 to 51 studies per meta-analysis, with a median of eight studies. The median number of case and control subjects in each study was 96 and 161, respectively. The median number of case and control subjects in each meta-analysis was 1123 and 3598, respectively. The number of cases was greater than 1000 in 70 meta-analyses. Overall, 441 (30%) individual studies observed nominally statistically significant results. Twenty-one (36%) meta-analyses used the Newcastle–Ottawa Scale to assess qualitatively the included primary studies. Two articles used assessment criteria for non-randomized observational studies adapted from Duckitt & Harrington, two articles used the Methodological Index for Non-Randomized Studies (MINORS) and nine articles used other assessment tools. Twenty-four papers (42%) did not perform any quality assessment. Supplementary Table 4.5 summarizes these 130 meta-analyses that included 1466 individual study estimates.

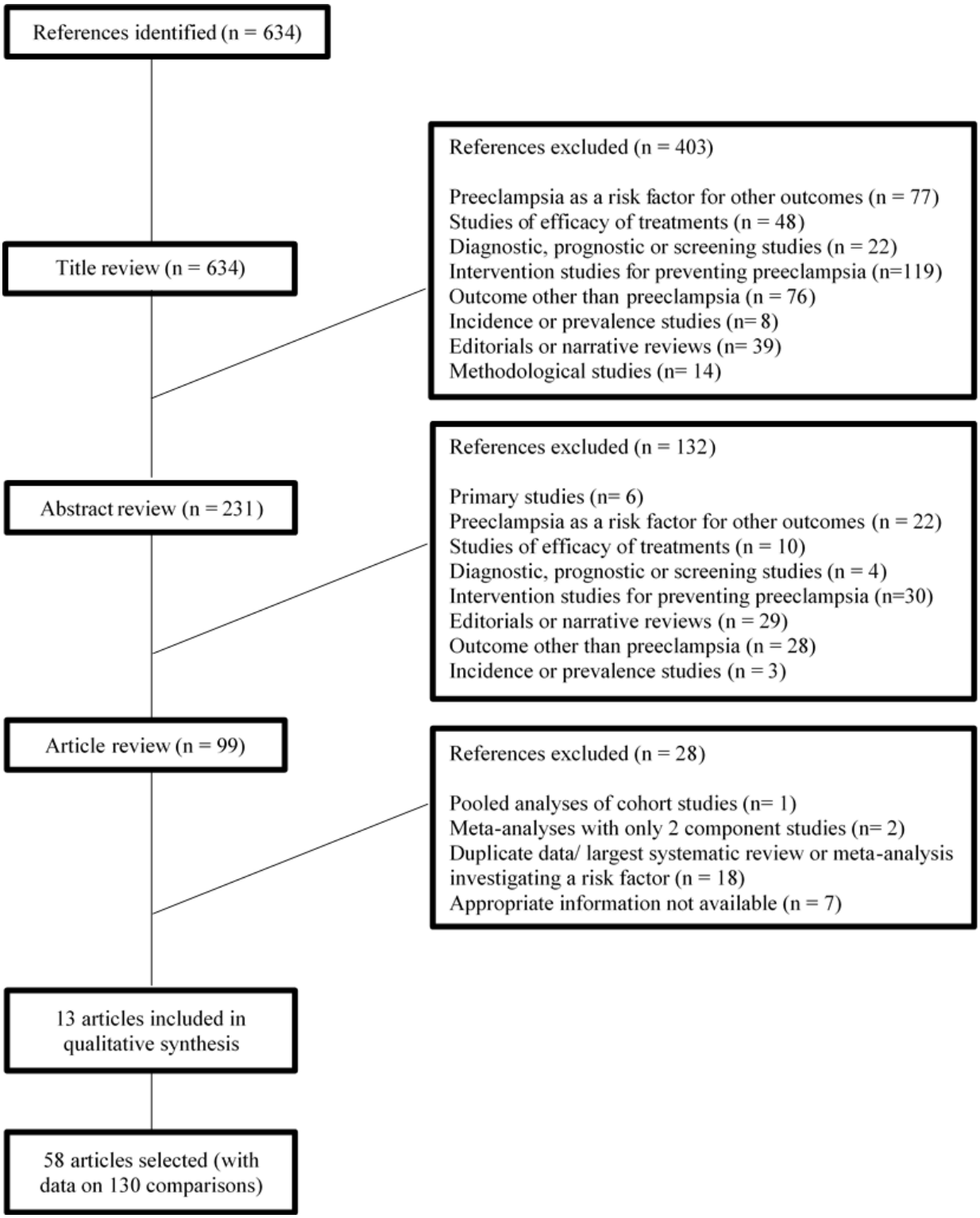

Figure 4.1. Flowchart of the included studies

Table 4.1. Quantitative synthesis and assessment of bias across the 130 associations of genetic and non-genetic risk factors and preeclampsia

Area	Author, year	Comparison	Study design	Studies	Cases/controls	Random effects*	Largest effect‡	P Random	Egger§	I^2 (P)‖	95% PI ≠
Biomarker	Fan Y, 2016	Copper level	Retrospective	12	442/463	1.86 (0.41-8.51)	1.22 (0.64-2.34)	.4217606	0.26	97 (<0.01)	0.00-835.6
Biomarker	Song QY, 2015	Serum iron level	Mixed	23	1023/889	9.97 (4.00-24.9)	38.02 (17.6-82.1)	8.22×10^{-7}	<0.01	96 (<0.01)	0.09-1101
Biomarker	Cohen MJ, 2015	Serum Vitamin E	Mixed	34	1578/1820	0.46 (0.27-0.79)	1.11 (0.61-2.04)	.46495506	<0.01	93 (<0.01)	0.02-10.3
Biomarker	Cohen MJ, 2015	Serum Vitamin C	Mixed	29	1362/1415	0.37 (0.22-0.61)	0.65 (0.48-0.87)	1.170×10^{-4}	0.08	91 (<0.01)	0.02-5.69
Biomarker	Liu HQ, 2015	β-hCG	Retrospective	12	702/8233	88.7 (4.31-1824)	NA	3.655×10^{-3}	0.75	100 (<0.01)	NA
Biomarker	Ma Y, 2015	Serum zinc level	Retrospective	14	541/550	0.35 (0.17-0.68)	0.10 (0.05-0.21)	2.230×10^{-3}	0.63	88 (<0.01)	0.02-5.43
Biomarker	Allen RE, 2014	PAPP-A	Mixed	9	1147/52208	2.05 (1.62-2.59)	1.52 (1.16-2.00)	2.53×10^{-9}	0.04	45 (0.07)	1.13-3.71
Biomarker	Allen RE, 2014	PLGF	Mixed	4	147/840	1.94 (0.81-4.66)	1.57 (0.81-3.05)	.13891351	0.08	83 (<0.01)	0.04-105
Biomarker	Allen RE, 2014	PP13	Mixed	4	210/3851	4.43 (2.86-6.85)	3.32 (1.77-6.22)	2.832×10^{-11}	0.48	49 (0.11)	0.85-23
Biomarker	Allen RE, 2014	betaHCG	Mixed	4	654/11669	1.09 (0.86-1.39)	1.58 (0.64-3.90)	.47136751	0.04	0 (0.45)	0.64-1.85
Biomarker	Allen RE, 2014	Inhibin A	Mixed	3	63/1152	3.57 (1.68-7.61)	8.94 (2.31-34.5)	9.516×10^{-4}	0.78	21 (0.28)	0.01-2184
Biomarker	Yang Y, 2014	IL-18	Mixed	10	351/421	1.13 (0.49-2.60)	1.02 (0.53-1.95)	.78202462	0.75	89 (<0.01)	0.05-24.3
Biomarker	Yang Y, 2014	IFN-γ	Mixed	12	567/701	5.42 (1.14-25.7)	45.6 (30.6-67.9)	.03330384	0.55	97 (<0.01)	0.01-2713
Biomarker	Lashley EE, 2013	HLA antibodies	Retrospective	3	64/273	0.93 (0.09-9.77)	1.40 (0.58-3.39)	.94851452	0.82	66 (0.05)	0-2.65
Biomarker	Dai B, 2013	Serum concentration of NO	Retrospective	9	297/303	0.17 (0.04-0.81)	2.56 (1.41-4.66)	.02535206	0.14	95 (<0.01)	0.00-50.9
Biomarker	Wei SQ, 2013	25 (OH) D <50 mmol/l	Mixed	6	209/1799	2.11 (1.52-2.94)	1.40 (0.69-2.85)	8.658×10^{-6}	0.66	0 (0.49)	1.32-3.37
Biomarker	Wei SQ, 2013	25 (OH) D <75 mmol/l	Mixed	5	177/1134	1.72 (1.11-2.69)	1.39 (0.27-7.24)	.01610334	0.48	27 (0.24)	0.57-5.21
Biomarker	Kleinrouweler CE 2012	PIGF	Mixed	26	787/3638	0.36 (0.25-0.54)	0.64 (0.33-1.23)	3.207×10^{-7}	0.01	84 (<0.01)	0.06-2.4
Biomarker	Kleinrouweler CE 2012	VEGF	Mixed	4	80/185	0.10 (0.01-1.53)	0.22 (0.08-0.57)	.09872404	0.19	96 (<0.01)	0-42370
Biomarker	Kleinrouweler CE 2012	sFlt-1	Mixed	32	1111/4119	2.38 (1.47-3.86)	1.24 (0.65-2.38)	4.517×10^{-4}	0.12	93 (<0.01)	0.15-37
Biomarker	Kleinrouweler CE 2012	sENG	Mixed	19	739/2402	2.66 (1.53-4.63)	1.20 (0.62-2.30)	5.063×10^{-4}	0.54	91 (<0.01)	0.22-32.3
Biomarker	Hausvater A, 2012	Arterial stiffness	Mixed	9	212/633	18.6 (3.72-93.0)	NA	3.697×10^{-4}	0.26	93 (<0.01)	0.05-6658
Biomarker	do Prado AD, 2010	Anticardiolipin antibodies	Mixed	12	1636/5111	2.85 (1.37-5.95)	1.88 (1.23-2.85)	5.208×10^{-3}	0.36	69 (<0.01)	0.29-28.1
Biomarker	Clark P, 2008	AB blood group	Mixed	13	5710/49069	1.02 (0.86-1.22)	0.82 (0.45-1.50)	.81449562	0.46	18 (0.26)	0.72-1.45
Biomarker	Clark P, 2008	A blood group	Mixed	14	5047/44743	0.96 (0.85-1.07)	1.00 (0.81-1.24)	.43608716	0.82	57 (<0.01)	0.68-1.35
Biomarker	Clark P, 2008	B blood group	Mixed	12	5324/48911	1.05 (0.94-1.18)	1.01 (0.72-1.42)	.40009776	0.71	23 (0.21)	0.82-1.35
Biomarker	Clark P, 2008	O blood group	Mixed	18	5945/54609	1.01 (0.91-1.12)	0.98 (0.80-1.21)	.85278952	0.52	49 (0.01)	0.73-1.39
Environmental	Hu H, 2014	NO_2	Mixed	5	3629/117497	1.10 (1.03-1.17)	1.06 (0.96-1.17)	4.565×10^{-3}	0.12	0 (0.73)	0.99-1.21
Environmental	Pedersen M, 2014	Air pollution	Mixed	4	4905/165789	1.05 (0.99-1.13)	1.13 (1.07-1.19)	.14465134	0.19	65 (0.03)	0.79-1.40
Environmental	Pedersen M, 2014	NOx	Mixed	3	1385/48725	1.03 (0.91-1.17)	1.00 (0.87-1.15)	.63256347	0.08	0 (0.54)	0.46-2.28
Environmental	Pedersen M, 2014	PM_{10}	Mixed	4	4656/201197	0.95 (0.86-1.05)	0.83 (0.77-0.89)	.31586644	0.73	83 (<0.01)	0.60-1.50
Environmental	Pedersen M, 2014	CO	Mixed	3	3583/112308	1.10 (0.99-1.22)	1.18 (1.03-1.35)	.09113282	0.94	24 (0.27)	0.44-2.76
Environmental	Pedersen M, 2014	O_3	Mixed	4	4943/164360	1.03 (1.00-1.06)	1.10 (0.94-1.30)	9.954×10^{-3}	0.07	0 (0.85)	0.98-1.09
Genetic markers	Zeng F, 2016	G894T	Retrospective	26	3241/6419	1.45 (1.09-1.94)	1.37 (0.92-2.04)	.01179173	0.65	41 (0.02)	0.55-3.86
Genetic markers	Zeng F, 2016	T-786C	Retrospective	15	2268/3100	1.25 (0.94-1.68)	2.57 (1.27-5.19)	.1302688	0.14	46 (0.02)	0.52-3.00
Genetic markers	Zhang G, 2016	rs4762 in AGT gene	Retrospective	3	790/2492	0.95 (0.66-1.38)	1.07 (0.62-1.84)	.78438216	0.20	26 (0.26)	0.04-23.9
Genetic markers	Zhang G, 2016	rs18001133 in MTHFR	Retrospective	49	13356/23082	1.17 (1.05-1.31)	1.26 (1.04-1.53)	5.889×10^{-3}	0.32	75 (<0.01)	0.60-2.29
Genetic markers	Zhang G, 2016	rs6025 in F5 gene	Retrospective	28	8210/9834	1.53 (1.06-2.21)	1.73 (0.78-3.83)	.02393371	0.61	74 (<0.01)	0.28-8.41
Genetic markers	Zhang G, 2016	rs1800896 in IL-10 gene	Retrospective	9	3020/3786	0.91 (0.75-1.11)	1.15 (0.98-1.35)	.36360487	0.04	70 (<0.01)	0.50-1.68
Genetic markers	Zhang G, 2016	rs1800871 in IL-10 gene	Retrospective	4	978/2074	0.79 (0.59-1.07)	0.84 (0.63-1.11)	.12511238	0.87	65 (0.04)	0.23-2.75
Genetic markers	Zhang G, 2016	rs1137101 in LEPR gene	Retrospective	28	8210/9834	1.53 (1.06-2.21)	1.73 (0.78-3.83)	.02393371	0.61	74 (<0.01)	0.28-8.41

Genetic markers	Zhang G, 2016	rs18001131 in MTHFR gene	Retrospective	9	2780/3636	1.15 (0.93-1.40)	0.91 (0.64-1.29)	.1917049	0.21	59 (0.01)	0.63-2.07
Genetic markers	Li Y, 2015	A1675G of AT2R	Retrospective	5	972/3072	1.58 (1.05-2.37)	1.25 (0.82-1.90)	.02686257	0.47	50 (0.09)	0.47-5.35
Genetic markers	Yang W, 2014	IL-10 -1082 A/G	Mixed	11	1741/3560	0.93 (0.77-1.13)	1.38 (0.62-3.09)	.48667154	0.30	63 (<0.01)	0.51-1.70
Genetic markers	Yang W, 2014	IL-10 -819 C/T	Mixed	5	729/1146	1.28 (1.03-1.59)	1.19 (0.88-1.62)	.02483578	0.86	41 (0.15)	0.70-2.35
Genetic markers	Yang W, 2014	IL-10 -592 C/A	Mixed	3	459/926	1.28 (1.03-1.59)	1.55 (1.04-2.30)	.02641458	0.39	0 (0.46)	0.31-5.26
Genetic markers	Wang X, 2014	G20210A SNP	Mixed	16	2296 /3262	1.79 (1.23-2.61)	1.84 (0.51-6.57)	2.545×10^{-3}	0.96	0 (0.92)	1.18-2.71
Genetic markers	Wang X, 2014	V G1691A SNP	Mixed	23	3131/4036	1.60 (1.25-2.06)	1.74 (0.78-3.89)	2.435×10^{-4}	<0.01	15 (0.25)	0.91-2.82
Genetic markers	Li X, 2014	MTHFR C677T	Mixed	47	6238/11771	1.12 (1.04-1.22)	1.28 (0.98-1.66)	5.188×10^{-3}	0.16	14 (0.21)	0.90-1.41
Genetic markers	Li X, 2014	TGF-β 1 869 T >C	Mixed	4	466/618	0.70 (0.57-0.86)	0.64 (0.39-1.03)	6.052×10^{-4}	0.93	0 (0.84)	0.45-1.09
Genetic markers	Gong LL, 2014	MMP9-1562C>T	Mixed	5	712/766	0.93 (0.61-1.42)	0.82 (0.53-1.27)	.7431311	0.34	72 (<0.01)	0.22-3.97
Genetic markers	Buurma AJ, 2013	AGT rs4762	Retrospective	5	497/1395	1.24 (0.67-2.30)	1.07 (0.62-1.84)	.4899227	0.31	80 (<0.01)	0.13-11.49
Genetic markers	Buurma AJ, 2013	APOE rs429358, rs7412	Retrospective	7	554/712	0.86 (0.65-1.13)	0.96 (0.60-1.55)	.27662924	0.04	4 (0.40)	0.57-1.29
Genetic markers	Buurma AJ, 2013	AT1R rs5186	Retrospective	9	886/1230	1.12 (0.95-1.33)	0.96 (0.69-1.34)	.18747175	0.33	0 0.78)	0.91-1.37
Genetic markers	Buurma AJ, 2013	CTLA4 rs231775	Retrospective	4	353/536	1.25 (1.01-1.56)	1.14 (0.80-1.61)	.04341501	0.82	1 (0.32)	0.68-2.29
Genetic markers	Buurma AJ, 2013	LPL rs1800590	Retrospective	3	395/579	2.27 (0.63-8.21)	0.81 (0.36-1.80)	.21122561	0.12	71 (0.03)	0-5626855
Genetic markers	Buurma AJ, 2013	LPL rs268	Retrospective	4	530/933	2.43 (1.26-4.68)	1.34 (0.51-3.50)	8.119×10^{-3}	0.66	20 (0.29)	0.35-17.1
Genetic markers	Buurma AJ, 2013	NOS3 27 bp-VNTR in intron 4	Retrospective	14	1593/2239	1.14 (0.90-1.43)	0.96 (0.71-1.30)	.2710968	0.03	63 (<0.01)	0.53-2.47
Genetic markers	Buurma AJ, 2013	NOS3 rs2070744	Retrospective	11	1571/2202	1.08 (0.95-1.23)	1.21 (0.96-1.52)	.25571731	0.10	28 (0.18)	0.80-1.46
Genetic markers	Buurma AJ, 2013	NOS3 rs1799983	Retrospective	24	2825/4048	1.19 (1.00-1.42)	1.79 (1.37-2.34)	.05650903	0.55	68 (<0.01)	0.56-2.52
Genetic markers	Buurma AJ, 2013	TLR4 rs4986790	Retrospective	4	723/614	1.07 (0.48-2.39)	3.03 (1.36-6.72)	.87139332	0.92	78 (<0.01)	0.03-38.2
Genetic markers	Buurma AJ, 2013	TLR4 rs4986791	Retrospective	3	614/461	1.20 (0.45-3.17)	2.92 (1.31-6.49)	.71483564	0.59	79 (<0.01)	0-123082
Genetic markers	Buurma AJ, 2013	TNF-alpha rs1800629	Retrospective	12	1592/1837	1.17 (0.91-1.49)	1.61 (1.17-2.22)	.21952434	0.48	54 (0.01)	0.56-2.41
Genetic markers	Buurma AJ, 2013	TNF-alpha rs1799724	Retrospective	3	390/385	0.66 (0.34-1.30)	1.18 (0.84-1.66)	.23144996	0.51	84 (<0.01)	0-2313
Genetic markers	Buurma AJ, 2013	VEGF rs3025039	Retrospective	3	377/514	1.36 (0.64-2.90)	0.73 (0.51-1.03)	.42048284	0.69	87(<0.01)	0-13603
Genetic markers	Cheng D, 2013	VEGF +936 C/T	Retrospective	8	805/1033	1.52 (1.09-2.12)	0.73 (0.51-1.03)	.0144147	0.58	69 (<0.01)	0.54-4.23
Genetic markers	Song GG, 2013	VEGF - 634 C/G	Retrospective	6	408/479	1.35 (1.09-1.67)	2.04 (1.33-3.13)	6.668×10^{-3}	0.86	12 (0.34)	0.90-2.01
Genetic markers	Song GG, 2013	VEGF -2578 A/ C	Retrospective	8	617/672	0.93 (0.78-1.10)	1.05 (0.78-1.41)	.39203909	0.99	13 (0.33)	0.68-1.26
Genetic markers	Song GG, 2013	VEGF -1154 A/G	Retrospective	3	159/161	1.14 (0.83-1.56)	1.06 (0.69-1.64)	.41612914	0.45	0 (0.89)	0.15-8.86
Genetic markers	Morgan JA, 2013	PAI-1 (4G/4G)	Mixed	12	1511/ 3492	1.28 (1.09-1.50)	1.19 (0.77-1.84)	2.646×10^{-3}	0.56	0 (0.63)	1.07-1.53
Genetic markers	Dai B, 2013	eNOS 4 b/a	Retrospective	10	1374/1376	1.43 (0.87-2.37)	1.77 (0.80-3.92)	.16052581	0.37	30 (0.17)	0.45-4.55
Genetic markers	Zhao L, 2013	SERPINE1 -675 4G/5G	Retrospective	11	1297/1791	1.37 (1.10-1.71)	1.66 (1.10-2.51)	5.112×10^{-3}	0.42	20 (0.25)	0.88-2.15
Genetic markers	Staines-Urias E, 2012	F5 rs6025	Mixed	41	4499/15188	1.74 (1.50-2.02)	1.67 (0.61-4.61)	2.902×10^{-13}	0.56	0 (0.53)	1.49-2.03
Genetic markers	Staines-Urias E, 2012	F2 rs1799963	Mixed	30	3546/11712	1.72 (1.40-2.12)	1.45 (0.67-3.14)	3.211×10^{-7}	0.03	0 (0.55)	1.38-2.14
Genetic markers	Staines-Urias E, 2012	ACE rs4646994	Mixed	30	3101/5134	1.17 (1.03-1.34)	1.03 (0.86-1.22)	.01714227	0.06	68 (<0.01)	0.65-2.13
Genetic markers	Staines-Urias E, 2012	AGT rs699	Mixed	27	2329/4896	1.26 (1.05-1.51)	1.31 (0.70-2.45)	.0110987	0.32	70 (<0.01)	0.57-2.79
Genetic markers	Staines-Urias E, 2012	MTHFR rs1801133	Mixed	51	5160/10151	1.06 (0.99-1.15)	1.21 (0.68-2.13)	.10516551	0.03	38 (<0.01)	0.79-1.49
Genetic markers	Staines-Urias E, 2012	SERPINE1 rs1799889	Mixed	12	1194/1757	0.89 (0.77-1.04)	0.90 (0.64-1.27)	.13240358	0.42	40 (0.76)	0.59-1.33
Genetic markers	Staines-Urias E, 2012	EPHX1 rs1051740	Mixed	4	562/462	0.85 (0.72-1.00)	0.94 (0.72-1.23)	.06194903	0.87	0 (0.51)	0.59-1.24
Genetic markers	Staines-Urias E, 2012	EPHX1 rs2234922	Mixed	3	425/427	1.28 (0.83-1.96)	1.87 (1.23-2.83)	.26470006	0.26	60 (0.08)	0.01-134
Genetic markers	Staines-Urias E, 2012	PPARG rs1801282	Mixed	3	390/449	0.80 (0.57-1.12)	0.81 (0.43-1.51)	.19441149	0.07	0 (0.90)	0.09-7.35
Genetic markers	Staines-Urias E, 2012	THBD C1418T	Mixed	3	260/268	0.71 (0.49-1.03)	0.78 (0.52-1.15)	.07266551	0.30	0 (0.50)	0.07-7.73
Genetic markers	Staines-Urias E, 2012	IL-6 rs1800795	Mixed	3	248/1575	0.91 (0.70-1.19)	0.91 (0.42-1.94)	.49809512	0.76	0 (0.90)	0.16-5.13
Genetic markers	Staines-Urias E, 2012	VEGFA rs699947	Mixed	3	225/269	0.88 (0.69-1.14)	0.92 (0.61-1.38)	.3352699	0.69	0 (0.90)	0.17-4.52
Genetic markers	Staines-Urias E, 2012	HLA-G -14 bp	Mixed	3	219/334	1.42 (0.68-2.98)	0.97 (0.68-1.38)	.35665444	0.90	85 (<0.01)	0-11540
Genetic markers	Staines-Urias E, 2012	LEP rs7799039	Mixed	3	198/326	1.51 (0.92-2.49)	1.20 (0.85-1.71)	.10567967	0.43	68 (0.05)	0.01-412
Genetic markers	Staines-Urias E, 2012	LEP TTTC	Mixed	3	141/227	0.86 (0.53-1.38)	1.01 (0.68-1.51)	.53082544	0.42	56 (0.10)	0.01-135
Genetic markers	Lin R, 2012	AGT M235T	Retrospective	29	5053/11578	1.61 (1.21-2.14)	1.40 (0.32-6.06)	9.986×10^{-4}	0.47	45 (<0.01)	0.57-4.52
Genetic markers	Lin R, 2012	AGT T174M	Retrospective	6	1362/4159	1.09 (0.76-1.57)	0.97 (0.54-1.74)	.63402843	0.35	48 (0.09)	0.40-2.95

Genetic markers	Zhao L, 2012	AGTR1 +1166A>C	Retrospective	10	845/1150	1.19 (0.96-1.47)	1.15 (0.67-1.99)	.11145683	0.42	27 (0.20)	0.74-1.91
Genetic markers	Zhong WG, 2012	ACE D/I	Retrospective	11	1600/1898	1.93 (1.19-3.12)	0.87 (0.59-1.28)	7.830×10^{-3}	0.26	91 (<0.01)	0.31-12.1
Genetic markers	Shaik AP, 2011	ACE (II genotype)	Retrospective	16	1620/2158	0.99 (0.70-1.40)	0.94 (0.57-1.54)	.93826151	0.79	73 (<0.01)	0.27-3.56
Genetic markers	Xie C, 2011	TNF-α 308 G/A	Retrospective	18	1888/2497	0.98 (0.76-1.25)	0.56 (0.36-0.87)	.85141826	0.56	52 (<0.01)	0.43-2.21
Genetic markers	Xie C, 2011	IL-6 -174 G/C	Retrospective	4	396/507	1.23 (0.93-1.61)	1.44 (0.89-2.33)	.14226516	0.44	0 (0.81)	0.67-2.24
Genetic markers	Rodger MA, 2010	FVL	Retrospective	9	1060/20773	1.26 (0.91-1.74)	1.27 (0.51-3.14)	.16965123	0.27	0 (0.99)	0.85-1.86
Genetic markers	Rodger MA, 2010	PGM	Prospective	6	549/13705	1.27 (0.80-2.03)	1.03 (0.41-2.56)	.31766677	0.30	0 (0.99)	0.65-2.46
Genetic markers	Medica I, 2007	AGT/T704C (Met235Thr)	Retrospective	15	1146/2276	1.66 (1.20-2.29)	0.29 (0.03-2.58)	2.242×10^{-3}	0.77	6 (0.38)	1.00-2.73
Genetic markers	Serrano NC, 2006	ACE-I/D	Mixed	22	2596/3828	1.23 (1.04-1.45)	0.90 (0.73-1.11)	.01737599	0.01	57 (<0.01)	0.66-2.26
Genetic markers	Lin J, 2005	FLV (1691 G-A)	Retrospective	11	1135/1471	2.25 (1.28-3.94)	2.21 (1.06-4.59)	4.609×10^{-3}	0.43	57 (<0.01)	0.42-12.2
Diseases/disorders	Saccone G, 2015	Celiac disease	Mixed	5	14618/507559	2.05 (0.89-4.74)	1.19 (0.79-1.78)	.09218346	0.66	90 (<0.01)	0.11-40.1
Diseases/disorders	Zhang JJ, 2015	Chronic kidney disease	Mixed	9	14993/504700	10.4 (6.28-17.1)	22.3 (15.6-31.9)	5.179×10^{-20}	0.71	77 (<0.01)	2.12-50.7
Diseases/disorders	Hu R, 2015	Depression	Mixed	5	1104/2874	1.66 (1.29-2.13)	1.12 (0.64-1.96)	6.521×10^{-5}	0.34	16 (0.32)	0.96-2.86
Diseases/disorders	Qin JZ, 2013	Polycystic ovary syndrome	Mixed	15	1866/1194098	3.26 (2.06-5.16)	2.04 (1.78-2.34)	4.327×10^{-7}	<0.01	41 (0.05)	1.02-10.43
Diseases/disorders	Zhang S, 2013	Mental stress	Mixed	12	16705/649188	1.49 (1.27-1.74)	1.14 (1.05-1.24)	5.169×10^{-7}	0.02	68 (<0.01)	0.97-2.29
Diseases/disorders	Zhang S, 2013	Work stress	Mixed	4	496/8246	1.50 (1.15-1.97)	1.51 (0.99-2.31)	3.197×10^{-3}	0.98	0 (0.75)	0.83-2.72
Diseases/disorders	Zhang S, 2013	Depression and anxiety	Mixed	5	753/7489	1.88 (1.08-3.25)	0.93 (0.55-1.59)	.0250717	0.44	73 (<0.01)	0.28-12.65
Diseases/disorders	Grigoriadis S, 2013	Maternal depression	Prospective	4	227/8843	1.35 (0.95-1.92)	1.24 (0.77-2.00)	.08895785	0.46	7 (0.36)	0.56-3.26
Supplementation	Schoenaker DA, 2014	Calcium intake	Mixed	3	387/1100	0.88 (0.60-1.29)	0.89 (0.53-1.52)	.51002502	0.87	0 (0.99)	0.07-10.82
Infections	Huang QT, 2016	Chronic hepatitis B infection	Retrospective	11	14298/423216	0.79 (0.63-1.00)	1.13 (0.78-1.63)	.04574222	0.90	20 (0.25)	0.51-1.25
Infections	Sgolastra F, 2013	Periodontal disease	Mixed	15	1040/3983	2.17 (1.38-3.41)	2.05 (1.47-2.86)	8.433×10^{-4}	0.50	78 (<0.01)	0.42-11.29
Infections	Rustveld LO, 2008	Bacterial & viral infections	Mixed	21	2390/11556	2.08 (1.63-2.66)	1.78 (1.18-2.67)	4.143×10^{-9}	0.65	56 (<0.01)	0.92-4.72
Other	Xu Y, 2016	Isolated single umbilical artery	Mixed	3	783/64443	0.82 (0.56-1.21)	0.84 (0.56-1.26)	.32120883	0.50	0 (0.85)	0.07-9.96
Other	Basaran A, 2016	CVS vs no invasive	Mixed	6	1189/46410	0.83 (0.42-1.66)	0.83 (0.61-1.13)	.60295188	0.29	92 (<0.01)	0.07-9.29
Other	Basaran A, 2016	CVS vs no invasive & amniocentesis	Mixed	7	1320/56266	1.00 (0.46-2.17)	0.83 (0.61-1.13)	.99506932	0.49	96 (<0.01)	0.06-16
Other	Wei J, 2015	Cigarette smoking	Prospective	17	62089/1784382	0.67 (0.60-0.75)	0.87 (0.83-0.91)	2.122×10^{-12}	0.36	92 (<0.01)	0.43-1.05
Other	Masoudian P, 2015	Oocyte donation vs ART	Retrospective	13	1499/25299	2.54 (1.98-3.24)	3.15 (2.27-4.37)	1.095×10^{-13}	0.90	14 (0.31)	1.61-4.00
Other	Masoudian P, 2015	Oocyte donation vs NC	Retrospective	4	2712/54816	4.33 (3.11-6.03)	3.35 (2.42-4.63)	3.477×10^{-18}	0.26	26 (0.26)	1.52-12.4
Other	Aune D, 2014	Pre-pregnancy PA high vs low activity	Mixed	5	621/9696	0.65 (0.45-0.94)	0.60 (0.30-1.20)	.02352111	0.63	0 (0.91)	0.36-1.19
Other	Aune D, 2014	Pre-pregnancy PA per 1hr per day	Mixed	3	479/6002	0.73 (0.53-0.99)	0.36 (0.07-1.88)	.04374593	0.09	0 (0.69)	0.10-5.42
Other	Aune D, 2014	Early pregnancy PA high vs low activity	Mixed	11	5702/162900	0.79 (0.70-0.91)	1.03 (0.74-1.44)	6.099×10^{-4}	0.90	0 (0.55)	0.68-0.92
Other	Aune D, 2014	Early pregnancy PA per 20 MET hrs/week	Mixed	3	2576/85388	0.86 (0.70-1.07)	0.98 (0.89-1.09)	.16690052	0.30	68 (0.04)	0.07-9.95
Other	Aune D, 2014	Early pregnancy PA per 1hr per day	Mixed	7	5293/151083	0.83 (0.73-0.95)	0.95 (0.80-1.14)	6.473×10^{-3}	0.66	20 (0.28)	0.63-1.09
Other	Aune D, 2014	Early pregnancy walking	Mixed	4	535/9674	0.68 (0.51-0.89)	1.00 (0.43-2.33)	5.549×10^{-3}	0.09	0 (0.75)	0.37-1.24
Other	Aune D, 2014	Early pregnancy occupational PA	Mixed	6	620/18119	0.82 (0.66-1.03)	0.75 (0.52-1.07)	.08838791	0.78	0 (0.68)	0.60-1.13
Other	González CM, 2014	Donor insemination	Mixed	7	2342/8556	1.57 (1.01-2.42)	1.69 (1.38-2.08)	.04326553	0.82	49 (0.07)	0.52-4.70
Other	Wang Z, 2013	Obese vs normal weight (adjusted)	Prospective	10	34340/1685991	2.93 (2.58-3.33)	3.64 (2.54-5.21)	0	0.11	67 (<0.01)	2.07-4.15
Other	Wang Z, 2013	Severe obese vs normal weight women	Prospective	6	19976/877162	3.12 (2.24-4.37)	2.53 (2.32-2.76)	2.581×10^{-11}	0.60	97 (<0.01)	0.96-10.2
Other	Kasawara KT, 2012	Physical activity (case-control)	Mixed	6	923/8481	0.77 (0.53-1.11)	1.16 (0.72-1.86)	.15938804	0.93	76 (<0.01)	0.23-2.60
Other	Kasawara KT, 2012	Physical activity (cohort studies)	Mixed	10	5547/178680	0.94 (0.83-1.07)	1.10 (1.01-1.19)	.33829233	0.17	60 (<0.01)	0.67-1.32
Other	Luo ZC, 2007	Primiparity	Mixed	23	54462/1966490	2.42 (2.16-2.71)	2.27 (2.22-2.32)	0	0.58	92 (0)	1.47-3.97

Abbreviations: Random effects, summary odds ratio (95% CI) using random effects model; Largest effect, odds ratio (95% CI) of the largest study in the meta-analysis; Egger, p-value from Egger's regression asymmetry test for evaluation of publication bias; P, p-value; β-hCG, Human chorionic gonadotropin; PAPP-A, pregnancy-associated plasma protein A; PLGF, Placental growth factor; PP13, Placental Protein 13; sFlt-1, Soluble fms-like tyrosine kinase-1; HLA, Human leukocyte antigen; PIGF, placental growth factor;

VEGF, vascular endothelial growth factor; sENG, soluble endoglin; NO_2, Nitrogen dioxide; NOx, Mono-nitrogen oxides; PM_{10}, Particulate matter 10 micrometers; CO, Carbon Monoxide; O_3, Ozone; IL-6, Interleukin 6; LEPR, leptin receptor; IL-18, Interleukin-18; IFN-γ, Interferon gamma; AT2R, Angiotensin type 2 receptor; IL-10, Interleukin 10; SNP, Single-nucleotide polymorphisms; MTHFR, Methylene tetrahydrofolate reductase; MMP-9, Matrix metallopeptidase 9; PAI-1, Plasminogen activator inhibitor-1; AGT, Angiotensinogen; AGTR1, Angiotensin II Receptor Type 1; ACE, Angiotensin; eNOS, Endothelial nitric oxide synthase; TNF, Tumor necrosis factor; FVL, Factor V Leiden; PGM, Prothrombin Gene Mutation; CVS, chorionic villus sampling; ART, assisted reproductive technology; NC, natural conception; PA, physical activity; NP, not pertinent, because the estimated is larger than the observed, and there is no evidence of excess of statistical significance based on the assumption made for the plausible effect size.

* Summary random effects odds ratio (95% CI) of each meta-analysis, except for two meta-analyses (Wei J 2015 and Aune D, 2014) where the RR was used.
‡ Odds ratio (95% CI) of the largest study in each meta-analysis, except for two meta-analyses (Wei J 2015 and Aune D, 2014) where the RR was used.
§ P-value from the Egger regression asymmetry test for evaluation of publication bias
|| I^2 metric of inconsistency and P-value of the Cochran Q test for evaluation of heterogeneity
≠ 95% Prediction Interval

Systematic reviews with qualitative synthesis

We have also summarized the evidence of the published systematic reviews without any quantitative component. According to their findings, serum calprotectin and cardiac troponin I levels were elevated in women with PE compared to healthy controls, where cell-free fetal DNA quantification has been shown to be a promising marker for PE prediction, especially for the development of early-onset or severe PE (352–354). PE was more prevalent in cold and humid seasons (355), whereas long inter-pregnancy intervals, possibly longer than 5 years, are also independently associated with an increased risk of PE (356). Psychotropic drugs such as lithium for the management of antenatal psychiatric disorders have been also associated with PE (357). Pregnant women with systemic lupus erythematosus and Cushing's syndrome are at higher risk of developing PE in contrast to healthy pregnancies (358,359). Laparoscopic adjustable gastric band (LAGB) surgery seems to improve pregnancy outcomes such as PE in obese women compared to pregnancies in obese women without LAGB (360,361). Limited evidence was found on whether shift work, HIV infection, or antiretroviral therapy and thrombophilic disorders are associated with an increased risk for PE (362–364).

Summary effect sizes and significant findings

Of the 130 meta-analyses, 65 (50%) had nominally statistically significant findings at $P<0.05$ using the random effects model, of which 53 reported increased risks and 12 showed decreased risks of PE. Out of these, a total of 28 (22%) associations presented statistically significant effect at $P<0.001$, while only 16 (12%) survived after the application of a more stringent p-value threshold of $P<10^{-6}$ (Table 4.1). The sixteen risk factors that presented a

significant effect at $P<10^{-6}$ for an association with PE were; the serum iron level, PAPP-A, PP13, PlGF, F5 rs6025, F2 rs1799963, chronic kidney disease, polycystic ovary syndrome, mental stress, bacterial & viral infections, cigarette smoking, oocyte donation vs ART, oocyte donation vs normal conception, obese vs normal weight, severe obese vs normal weight and primiparity. Additional information on all 130 meta-analyses is available on Supplementary Table 4.5.

Across the seven areas of risk factors there were differences in the proportion of associations that had nominally statistically significant summary effects. Based on the random effects calculations at $P<0.05$, 100%, 75%, 63% and 59% of the meta-analyses on infections, diseases and disorders, other risk factors and biomarkers respectively, found nominally statistically significant summary effects. On the contrary, this was seen only in 39% and 33% of the meta-analyses on genetic markers and environmental factors, respectively.

Between-study heterogeneity and prediction intervals

33 (25%) meta-analyses had large heterogeneity estimates ($I^2 \geq 50\%$ and $I^2 \leq 75\%$) and 32 (25%) meta-analyses had very large heterogeneity estimates ($I^2 > 75\%$) (Table 4.1). The highest proportion (56%) of I^2 exceeding 75% was observed in meta-analyses of biomarkers. When we calculated the 95% prediction intervals, in only 14 (11%) meta-analyses the null value was excluded. This included two meta-analyses on biomarkers (PAPP-A and Vitamin D <50 mmol/l), five on genetic markers (G20210A SNP, PAI-1 4G/5G, F5 rs6025, F2 rs1799963, AGT/T704C-Met235Thr), two on diseases and disorders (chronic kidney disease and polycystic ovary syndrome), and five on other risk factors (oocyte donation vs ART,

oocyte donation vs spontaneous conception, high vs low levels of physical activity in early pregnancy, obese vs normal weight and primiparity) (Table 4.1).

Small-study effects

Evidence for statistically significant small-study effects (Egger test P<0.10 and the random effects summary estimate was larger compared to the point estimate of the largest study in the meta-analysis) was identified in 10 of 130 (8%) meta-analyses (Supplementary Table 4.5). These included two meta-analyses on biomarkers (PAPP-A, PlGF), one on environmental factors (NOx), four on genetic markers (NOS3 27 bp-VNTR in intron 4, F2 rs1799963, ACE rs4646994, ACE-I/D), two on diseases and disorders category (polycystic ovary syndrome and mental stress) and one on other risk factors (Pre-pregnancy physical activity per 1hr per day).

Test of excess statistical significance

Twenty-six (20%) associations had hints for excess statistical significance bias with statistically significant (P<0.05) excess of positive studies under any of the three assumptions for the plausible effect size; the fixed effects summary, the random effects summary or the results of the largest study (Supplementary Table 4.5). Ten (38%) of them pertained to the biomarkers, nine (35%) pertained to genetic markers, three (12%) pertained to diseases and disorders, and four (15%) pertained to other risk factors. Also, the observed and expected number of positive studies shows that overall the excess of positive results was driven by meta-analyses with small estimates of heterogeneity ($I^2 \leq 50\%$). Table 4.2 shows the results of excess of statistical significance bias according to category of risk factor.

Table 4.2. Observed and expected number of positive studies by type of risk factor*

Area	No. of studies	Observed positive	Expected positive (fixed) †	P‡ (fixed)	Expected positive (random)§	P‡ (random)	Expected positive (largest)‖	P‡ (largest)	Expected positive (composite) ¶	P‡ (composite)
All	1466	479	560.3	0.00	605.9	0.00	601.3	0.00	560.3	0.00
Biomarkers	353	178	166	0.20	200	0.02	133	0.00	133	0.00
Environmental	23	4	4.9	0.80	4.4	NP	10.5	0.01	4.4	NP
Genetic markers	830	162	229.6	0.00	235.5	0.00	323.4	0.00	229.6	0.00
Diseases & disorders	59	29	37.6	0.03	45	0.00	27.4	0.70	27.4	0.70
Supplementation	3	0	0.32	NP	0.32	NP	0.3	NP	0.3	NP
Infections	47	21	27.3	0.08	28.9	0.02	23	0.66	23	0.66
Other	151	85	95	0.09	92.2	0.24	84	0.93	84	0.93

* NP, not pertinent, because the estimated is larger than the observed, and there is no evidence of excess of statistical significance based on the assumption made for the plausible effect size.

† Expected number of statistically significant studies using the summary fixed effects estimate of each meta-analysis as the plausible effect size.

‡ P value of the excess of statistically significant test. All statistical tests were two-sided.

§ Expected number of statistically significant studies using the summary random effects estimate of each meta-analysis as the plausible effect size.

‖ Expected number of statistically significant studies using the effect of the largest study of each meta-analysis as the plausible effect size.

¶ Expected number of statistically significant studies using the most conservative of the three estimates (fixed effects summary, random effects summary, largest study) of each meta-analysis as the plausible effect size.

Risk factors with strong evidence of association

After applying our credibility criteria, only one non-genetic risk factor, oocyte donation vs spontaneous conception, presented convincing evidence for an association with PE, supported by more than 1000 cases, $P<10^{-6}$ under the random effect model, no hints for small-study effects and for excess statistical significance, not large heterogeneity ($I^2<50\%$) and a 95% PI excluding the null value. This association had a summary OR of 4.33 (95% CI: 3.11-6.03; $p=3.48 \times 10^{-18}$) with small heterogeneity ($I^2=26\%$) and supported by 2712 cases. Eleven risk factors (serum iron level, PAPP-A, chronic kidney disease, polycystic ovary syndrome, mental stress, bacterial & viral infections, cigarette smoking, oocyte donation vs ART, obese vs normal weight women, severe obese vs normal weight women and primiparity) presented highly suggestive evidence for PE. Five risk factors were supported by suggestive evidence, and 22 associations presented weak evidence. An overall assessment of statistically significant associations for PE is presented in Table 4.3.

Assessment of the cumulative epidemiologic evidence for genetic associations was also conducted and evidence was scored as strong, moderate, or weak based on grades of 'A', 'B', or 'C', as specified by the Venice criteria. Of the 26 variants with significant associations with PE risk with $P<0.05$ using the random effects model, only the PAI-1 4G/5G polymorphism (recessive model) was supported by strong evidence for a contribution to the pathogenesis of PE (Table 4.4).

Independent tool-based quality assessment of the primary studies

We have also assessed the quality of the included studies of the meta-analysis of the non-genetic risk factor that presented convincing evidence for an association with PE

using the Newcastle Ottawa Scale (365), in addition to the MINORS scale that the authors used in the original assessment. The methodological quality ranged from 3 points to 8 points maximally, with a median of 6 points, which implies a fair quality of the majority of studies. A quality assessment was also performed among the included studies of meta-analysis of the PAI-1 4G/5G polymorphism using the Q-Genie tool (366). Among the reviewed studies, 8 (67%) studies were rated to have high quality (>45) and 4 (33%) were rated to have moderate quality (>35 and ≤45).

Table 4.3. Assessment across the statistically significant non-genetic associations for preeclampsia

Level of evidence	**Criteria**
Convincing	>1000 cases, [a] $P<10^{-6}$, not large heterogeneity (I^2 <50%), 95% prediction interval excluding the null value, no evidence for small-study effects [b] and excess significance bias [c]
Risk factors supported by convincing evidence	Oocyte donation vs spontaneous conception
Highly suggestive	>1000 cases, [a] $P<10^{-6}$ and nominally statistically significant effect present at the largest study
Risk factors supported by highly suggestive evidence	Serum iron level, PAPP-A, Chronic kidney disease, Polycystic ovary syndrome, Mental stress, Bacterial & viral infections, Cigarette smoking*, Oocyte donation vs ART, Obese vs normal weight women, Severe obese vs normal weight women, Primiparity
Suggestive	>1000 cases, [a] $P<10^{-3}$
Risk factors supported by suggestive evidence	Serum Vitamin C*, sFLT1, Depression, Periodontal disease, Early pregnancy PA high vs low activity*
Weak	The rest associations with [a] $P < 0.05$
Risk factors supported by weak evidence	β-hCG, Serum zinc level*, PP13, Inhibin A, IFN-γ, Serum concentration of NO*, PlGF*, sENG, Arterial stiffness, Anticardiolipin antibodies, NO_2, O_3, Work stress, Depression and anxiety, 25 (OH) D <75 mmol/l, 25 (OH) D <50 mmol/l, Chronic hepatitis B infection*, Pre-pregnancy PA high vs low activity*, Pre-pregnancy PA per 1hr per day*, Early pregnancy PA per 1hr per day*, Early pregnancy walking*, Donor insemination

Abbreviations: β-hCG, Human chorionic gonadotropin; PAPP-A, pregnancy-associated plasma protein A; IFN-γ, Interferon gamma; PP13, Placental Protein 13; sFlt-1, Soluble fms-like tyrosine kinase-1; PlGF, placental growth factor; sENG, soluble endoglin; NO_2, Nitrogen dioxide; O_3, Ozone; ART, assisted reproductive technology; PA, physical activity

[a] P indicates the P-values of the meta-analysis random effects model.
[b] Small study effect is based on the P-value from the Egger's regression asymmetry test (P< 0.10).
[c] Based on the P-value (P<0.05) of the excess significance test using the largest study (smallest standard error) in a meta-analysis as the plausible effect size.
* Factors that show a protective effect against developing pre-eclampsia.

Table 4.4. Assessment of cumulative evidence on 26 significant (P<0.05) genetic associations with preeclampsia risk

Author, year	Gene or variant	Comparison	Studies	Cases/controls	Random effects*	P Random	Egger§	I^2 (P)‖	Excess statistical significance ≠	Venice Criteria†	Cumulative Evidence of Association ¥
Zeng F, 2016	G894T	TT vs GT + GG	26	3241/6419	1.45 (1.09-1.94)	.0118	0.65	41 (0.02)	No	BAA	++
Zhang G, 2016	rs18001133 in MTHFR	Carriers vs non-carriers	49	13356/23082	1.17 (1.05-1.31)	5.89×10^{-3}	0.32	75 (<0.01)	No	AAB	++
Zhang G, 2016	rs6025 in F5 gene	Carriers vs non-carriers	28	8210/9834	1.53 (1.06-2.21)	.0239	0.61	74 (<0.01)	No	BAB	++
Zhang G, 2016	rs1137101 in LEPR	Carriers vs non-carriers	28	8210/9834	1.53 (1.06-2.21)	.0239	0.61	74 (<0.01)	No	BAB	++
Li Y, 2015	A1675G of AT2R	GG vs AG + AA	5	972/3072	1.58 (1.05-2.37)	.0269	0.47	50 (0.09)	No	BAB	++
Yang W, 2014	IL-10 -819 C/T	C vs T	5	729/1146	1.28 (1.03-1.51)	.0248	0.86	41 (0.15)	No	AAB	++
Yang W, 2014	IL-10 -592 C/A	C vs A	3	459/926	1.28 (1.03-1.59)	.0264	0.39	0 (0.46)	No	BAA	++
Wang X, 2014	G20210A SNP	GG vs GA/AA	16	2296 /3262	1.79 (1.23-2.61)	2.55×10^{-3}	0.96	0 (0.92)	No	AAB	++
Wang X, 2014	V G1691A SNP	GG vs GA/AA	23	3131/4036	1.60 (1.25-2.06)	2.44×10^{-4}	<0.01	15 (0.25)	No	AAB	++
Li X, 2014	MTHFR C677T	CT + TT vs CC	47	6238/11771	1.12 (1.04-1.22)	5.19×10^{-3}	0.16	14 (0.21)	Yes	AAB	++
Li X, 2014	TGF-β 1 869 T >C	TT vs TC + CC	4	466/618	0.70 (0.57-0.86)	6.05×10^{-4}	0.93	0 (0.84)	No	BAA	++
Buurma AJ, 2013	CTLA4 rs231775	Carriers vs non-carriers	4	353/536	1.25 (1.01-1.56)	.0434	0.82	14 (0.32)	No	BAA	++
Buurma AJ, 2013	LPL rs268	Carriers vs non-carriers	4	530/933	2.43 (1.26-4.68)	.0081	0.66	20 (0.29)	No	BAA	++
Cheng D, 2013	VEGF +936 C/T	T vs C	8	805/1033	1.52 (1.09-2.12)	.0144	0.58	69 (<0.01)	No	BAC	+
Song GG, 2013	VEGF - 634 C/G	C vs G	6	408/479	1.35 (1.09-1.67)	6.67×10^{-3}	0.86	12 (0.34)	No	BAB	++
Morgan JA, 2013	PAI-1	4G/4G	12	1511/ 3492	1.28 (1.09-1.50)	2.65×10^{-3}	0.56	0 (0.63)	No	AAA	+++
Zhao L, 2013	SERPINE1 -675	4G/4G vs 4G/5G + 5G/5G	11	1297/1791	1.37 (1.10-1.71)	5.11×10^{-3}	0.42	20 (0.25)	No	BAB	++
Staines-Urias E, 2012	F5 rs6025	Carriers vs non-carriers	41	4499/15188	1.74 (1.50-2.02)	2.90×10^{-13}	0.56	0 (0.53)	Yes	AAB	++
Staines-Urias E, 2012	F2 rs1799963	Carriers vs non-carriers	30	3546/11712	1.72 (1.40-2.12)	3.21×10^{-7}	0.03	0 (0.55)	Yes	BAB	++
Staines-Urias E, 2012	ACE rs4646994	Carriers vs non-carriers	30	3101/5134	1.17 (1.03-1.34)	.0171	0.06	68 (<0.01)	Yes	AAC	+
Staines-Urias E, 2012	AGT rs699	Carriers vs non-carriers	27	2329/4896	1.26 (1.05-1.51)	.0111	0.32	70 (<0.01)	No	AAB	++
Lin R, 2012	AGT M235T	TT vs MM	29	5053/11578	1.61 (1.21-2.14)	9.99×10^{-4}	0.47	45 (<0.01)	Yes	AAC	+
Zhong WG, 2012	ACE D/I	D vs I	11	1600/1898	1.93 (1.19-3.12)	7.83×10^{-3}	0.26	91 (<0.01)	Yes	AAC	+
Medica I, 2007	AGT/T704C (Met235Thr)	CC + TT vs TT	15	1146/2276	1.66 (1.20-2.29)	2.24×10^{-3}	0.77	6 (0.38)	No	BAB	++
Serrano NC, 2006	ACE-I/D	Carriers vs non-carriers	22	2596/3828	1.23 (1.04-1.45)	.0174	0.01	57 (<0.01)	No	AAC	+
Lin J, 2005	FLV (1691 G-A)	Carriers vs non-carriers	11	1135/1471	2.25 (1.28-3.94)	4.61×10^{-3}	0.43	57 (<0.01)	No	BAA	++

Abbreviations: Random effects, summary odds ratio (95% CI) using random effects model; Largest effect, odds ratio (95% CI) of the largest study in the meta-analysis; Egger, p-value from Egger's regression asymmetry test for evaluation of publication bias; P, p-value; PIGF, placental growth factor; VEGF, vascular endothelial growth factor; IL-6, Interleukin 6; LEPR, leptin receptor; AT2R, Angiotensin type 2 receptor; IL-10, Interleukin 10; SNP, Single-nucleotide polymorphisms; MTHFR, Methylene tetrahydrofolate reductase; MMP-9, Matrix metallopeptidase 9; PAI-1, Plasminogen activator inhibitor-1; AGT, Angiotensinogen; AGTR1, Angiotensin II Receptor Type 1; ACE, Angiotensin; eNOS, Endothelial nitric oxide synthase; TNF, Tumor necrosis factor; FVL, Factor V Leiden; PGM, Prothrombin Gene Mutation.

* Summary random effects odds ratio (95% CI) of each meta-analysis.
‡ Odds ratio (95% CI) of the largest study in each meta-analysis
§ P-value from the Egger regression asymmetry test for evaluation of publication bias
‖ I^2 metric of inconsistency and P-value of the Cochran Q test for evaluation of heterogeneity
≠ Based on the P-value (P<0.05) of the excess significance test using the largest study (smallest standard error) in a meta-analysis as the plausible effect size.
†Venice Criteria grades are in the order of amount of evidence, replication of the association and protection from bias
¥ Cumulative epidemiological evidence as graded by the Venice criteria as strong (+++), moderate (++), or weak (+) for association with preeclampsia risk

4.5 Discussion

Main Findings

Overall, 130 associations have been studied as risk factors for PE, including biomarkers, genetic markers, environmental factors, supplementation, diseases and disorders, infections and other risk factors. Of those, oocyte donation vs spontaneous conception provided convincing evidence. PAI-1 4G/5G (recessive model) polymorphism had strong evidence for a contribution to the pathogenesis of PE, as specified by the Venice criteria. Eleven risk factors from various fields achieved highly suggestive evidence for an association with PE.

Interpretation

PE remains a disease of theories, as a large number of factors and a genetic component is likely to be involved, but none have been clearly established to date. From biological standpoint, oocyte donation can act as an independent risk factor for development of PE. During normal pregnancy, various immunosuppressive mechanisms maintain to diminished innate immune response in order to prevent fetal rejection as the fetal tissue is directly exposed to the maternal blood and hence, at risk of being attacked by components of both the innate and acquired immune system (367,368). A fetus conceived spontaneously is a semi-allograft, in which both maternal and paternal genes are expressed, whereas a fetus conceived through oocyte donation is an absolute allograft and this could lead to an altered or inadequate immune-protection of placentation and eventually resulting in PE (369–372). This theory is further supported from the fact that oocyte donation versus other assisted reproduction techniques had highly suggestive evidence of epidemiological credibility. Moreover, immune dysregulation may interpret the highly suggestive evidence in the risk of pre-eclampsia

among primiparous women because the first successful (non-preeclamptic) pregnancy may induce adaptive changes in favor to immune tolerance in subsequent pregnancies (351).

The genetic architecture behind PE is complex (174). To date, most research in this field has been focused on candidate genes, primarily those for which a plausible role in the known underlying pathophysiology (175). Only three genome-wide association studies were identified that include several genetic loci associated with PE (193–195). One study, identified two loci (rs7579169 and rs12711941) near the Inhibin beta B gene that satisfied the genome-wide significance threshold (194), but the results could not be replicated in two cohorts from Norway and Finland. Subsequent case-control studies in European and Chinese women have shown a significant association between the SNP rs7579169 and PE (196,197).

Eleven factors (serum iron level, PAPP-A, chronic kidney disease, polycystic ovary syndrome, mental stress, bacterial & viral infections, cigarette smoking, oocyte donation vs ART, obese/severe vs normal weight women, primiparity), achieved highly suggestive evidence for an association with PE. There are several mechanisms that support these findings. Regarding biomarkers, iron is considered a significant etiologic factor in the endothelial cell damage in PE cases because of its effects on the formation of oxygen free radicals and subsequent lipid peroxidation (373–375). Reduced PAPP-A, being an important regulator of insulin-like growth factor, can play a role in the development of PE in normal karyotype pregnancies (376).

Renal insufficiency, maternal hypertension, proteinuria, and recurrent urinary tract infection which are often coexist in women with chronic kidney disease, may contribute individually and cumulatively to PE (377–379). Insulin resistance and/or associated hyperglycemia that often exist in polycystic ovary syndrome (PCOS) and obese patients could be a possible explanation of a higher risk for PE, since it possibly directly predispose women to hypertension by increased renal sodium re-absorption and stimulation of the sympathetic nervous system and/or may impair endothelial function (380). Increased levels of androgens in PCOS pregnancies have also been associated with the development of PE (381).

Cigarette smoking during pregnancy seems protective against developing PE. Experimental studies have demonstrated that carbon monoxide decrease the levels of antiangiogenic factors such as sFlt1 and soluble endoglin, or increase the levels of angiogenic factors like VEGF, (382) which are thought to be involved in the pathogenesis of PE (383–385). Infection may be important in the pathogenesis of PE, either through initiation by increasing the risk of acute uteroplacental atherosclerosis and/or its enhancement by magnifying the maternal systemic inflammatory response (386) or through direct effect on trophoblast cells by destruction or impairment of trophoblast cells, resulting in shallow invasion of maternal spiral arteries (387).

Strengths and limitations

Both Egger and excess of significance test offer hints of bias, not definitive proof thereof, while the Egger test is difficult to interpret when the between-study heterogeneity is large. The frequency of meta-analyses with small-study asymmetry effects was not high (8%), and this rate is commensurate with chance. Nevertheless,

our estimates are likely to be conservative as a negative test result does not exclude the potential for bias.

The majority of the included studies for non-genetic associations were retrospective which is indicative of a higher potential for bias inherent in the included studies. However, by performing a standardized methodological process for the assessment of the epidemiological credibility of the findings using a variety of test we accomplish to incorporate all these biases together and provide a complete picture of the totality of evidence as it stands today. The interpretation of excess of statistical significance test for the results of a single meta-analysis, especially one with few studies, should be cautious because a negative test does not exclude the potential for bias (82). Furthermore, quality assessment of the primary studies was very heterogeneous, reflecting the lack of standardized quality assessment methodologies.

4.6 Conclusion

Oocyte donation vs spontaneous conception was supported by convincing evidence for an association with PE, and 11 risk factors achieved highly suggestive evidence for an association with PE. PAI-1 4G/5G (recessive model) polymorphism was supported by strong evidence for a contribution to the pathogenesis of PE. The use of standardized definitions and protocols for exposures, outcomes, and statistical analyses (388,389), the adoption of reporting guidelines (e.g. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) and STrengthening the REporting of Genetic Association Studies (STREGA)) (390,391) and registration of hypothesis-testing observational studies, (392,393) may help improve the evidence

in the future, diminish the threat of biases and improve the reliability of this important literature.

Contribution to authorship: KG and SP were involved in formulating the hypothesis and the design of the study protocol. KG and SP performed the literature search, the selection of eligible articles and the data extraction. KG analysed the data. All authors (KG, EE, SP) were involved in data interpretation. KG and SP wrote the first draft of the manuscript and EE was involved in the revision of the manuscript. All authors (KG, EE, SP) approved the final version of the submitted manuscript. KG and SP are guarantors.

Chapter 4: Supplemental material

Supplemental Table 4.5. Analytical description of the 130 selected meta-analyses with observed and expected number of "positive" study datasets

Area	Author, year	Comparison	Studies	Cases/controls	Random effects*	Fixed effects†	Largest effect‡	Egger§	I^2 (95% CI) (P)‖	95% PI ≠	O¶	E #	P** (fixed)	E ¥	P** (random)	E ϑ	P** (largest)	P** (largest)
Biomarker	Fan Y, 2016	Copper level	12	442/463	1.86 (0.41-8.51)	3.53 (2.69-4.64)	1.22 (0.64-2.34)	0.26	97 (96-97) (<0.01)	0.00-835.6	9	9.7	0.71	3.7	0.00	0.9	0.00	0.00
Biomarker	Song QY, 2015	Serum iron level	23	1023/889	9.97 (4.00-24.9)	5.28 (4.36-6.38)	38.02 (17.6-82.1)	<0.01	96 (95-96) (<0.01)	0.09-1101	20	22.6	0.05	23.8	NP	24	NP	NP
Biomarker	Cohen MJ, 2015	Serum Vitamin E	34	1578/1820	0.46 (0.27-0.79)	0.76 (0.66-0.87)	1.11 (0.61-2.04)	<0.01	93 (92-94) (<0.01)	0.02-10.3	20	3.8	0.00	15.1	0.12	2.0	0.00	0.00
Biomarker	Cohen MJ, 2015	Serum Vitamin C	29	1362/1415	0.37 (0.22-0.61)	0.52 (0.45-0.61)	0.65 (0.48-0.87)	0.08	91 (89-93) (<0.01)	0.02-5.69	19	9.8	0.00	17.7	0.71	5.6	0.00	0.00
Biomarker	Liu HQ, 2015	β-hCG	12	702/8233	88.7 (4.31-1824)	30.5 (25.8-35.9)	NA	0.75	100 (NA) (<0.01)	NA	7	12	NP	12	NP	12	NP	NP
Biomarker	Ma Y, 2015	Serum zinc level	14	541/550	0.35 (0.17-0.68)	0.37 (0.29-0.46)	0.10 (0.05-0.21)	0.63	88 (83-92) (<0.01)	0.02-5.43	8	9	0.59	9.6	0.39	13.9	NP	NP
Biomarker	Allen RE, 2014	PAPP-A	9	3340/52208	2.05 (1.62-2.59)	1.85 (1.60-2.15)	1.52 (1.16-2.00)	0.04	45 (0-73) (0.07)	1.13-3.71	7	7.5	0.66	7.9	0.31	6.3	0.73	0.73
Biomarker	Allen RE, 2014	PlGF	4	147/840	1.94 (0.81-4.66)	1.61 (1.13-2.30)	1.57 (0.81-3.05)	0.08	83 (38-92) (<0.01)	0.04-105	1	1.0	NP	1.8	0.63	1.0	NP	NP
Biomarker	Allen RE, 2014	PP13	4	210/3851	4.43 (2.86-6.85)	4.33 (3.19-5.89)	3.32 (1.77-6.22)	0.48	49 (0-82) (0.11)	0.85-23	4	4.0	NP	4.0	NP	3.9	NP	NP
Biomarker	Allen RE, 2014	β-hCG	4	654/11669	1.09 (0.86-1.39)	1.09 (0.86-1.39)	1.58 (0.64-3.90)	0.04	0 (0-68) (0.45)	0.64-1.85	0	0.5	NP	0.5	NP	3.6	NP	NP
Biomarker	Allen RE, 2014	Inhibin A	3	63/1152	3.57 (1.68-7.61)	3.41 (1.84-6.30)	8.94 (2.31-34.5)	0.78	21 (0-78) (0.28)	0.01-2184	2	2.3	NP	2.3	0.55	2.8	0.16	0.16
Biomarker	Yang Y, 2014	IL-18	10	351/421	1.13 (0.49-2.60)	1.17 (0.89-1.53)	1.02 (0.53-1.95)	0.75	89 (82-92) (<0.01)	0.05-24.3	6	0.7	0.00	0.6	0.00	0.5	0.00	0.00
Biomarker	Yang Y, 2014	IFN-γ	12	567/701	5.42 (1.14-25.7)	4.82 (3.78-6.14)	45.6 (30.6-67.9)	0.55	97 (97-98) (<0.01)	0.01-2713	7	10	NP	10.5	NP	12	NP	NP
Biomarker	Lashley EE, 2013	HLA antibodies	3	64/273	0.93 (0.09-9.77)	1.27 (0.56-2.89)	1.40 (0.58-3.39)	0.82	66 (0-82) (0.05)	0-2.65	0	0.2	NP	0.2	NP	0.3	NP	NP
Biomarker	Dai B, 2013	Serum concentration of NO	9	297/303	0.17 (0.04-0.81)	0.32 (0.23-0.43)	2.56 (1.41-4.66)	0.14	95 (94-97) (<0.01)	0.00-50.9	9	6.1	0.04	8.2	NP	4.7	0.00	0.00
Biomarker	Wei SQ, 2013	25 (OH) D <50 mmol/l	6	209/1799	2.11 (1.52-2.94)	2.11 (1.52-2.94)	1.40 (0.69-2.85)	0.66	0 (0-61) (0.49)	1.32-3.37	3	3.1	NP	3.1	NP	0.9	0.05	0.05
Biomarker	Wei SQ, 2013	25 (OH) D <75 mmol/l	5	177/1134	1.72 (1.11-2.69)	1.77 (1.23-2.55)	1.39 (0.27-7.24)	0.48	27 (0-73) (0.24)	0.57-5.21	2	1.6	0.66	1.5	0.60	0.7	0.15	0.15
Biomarker	Kleinrouweler CE 2012	PlGF	26	787/3638	0.36 (0.25-0.54)	0.48 (0.42-0.56)	0.64 (0.33-1.23)	0.01	84 (78-88) (<0.01)	0.06-2.4	12	11.9	NP	17.2	0.04	5.8	0.01	0.01
Biomarker	Kleinrouweler CE 2012	VEGF	4	80/185	0.10 (0.01-1.53)	0.23 (0.12-0.39)	0.22 (0.08-0.57)	0.19	96 (93-97) (<0.01)	0-42370	2	3.3	0.13	4.0	NP	3.4	0.11	0.11
Biomarker	Kleinrouweler CE 2012	sFlt-1	32	1111/4119	2.38 (1.47-3.86)	1.88 (1.66-2.14)	1.24 (0.65-2.38)	0.12	93 (91-94) (<0.01)	0.15-37	12	12.7	0.86	19.5	0.01	2.9	0.00	0.00
Biomarker	Kleinrouweler CE 2012	sENG	19	739/2402	2.66 (1.53-4.63)	2.46 (2.09-2.90)	1.20 (0.62-2.30)	0.54	91 (88-93) (<0.01)	0.22-32.3	9	12.7	0.09	13.8	NP	1.5	0.00	0.00
Biomarker	Hausvater A, 2012	Arterial stiffness	9	212/633	18.6 (3.72-93.0)	10.2 (6.76-15.3)	NA	0.26	93 (90-95) (<0.01)	0.05-6658	8	8.5	0.39	8.9	0.13	9.0	0.00	0.00
Biomarker	do Prado AD, 2010	Anticardiolipin antibodies	12	1636/5111	2.85 (1.37-5.95)	2.25 (1.65-3.01)	1.88 (1.23-2.85)	0.36	69 (33-81) (<0.01)	0.29-28.1	4	8.8	NP	10.1	NP	7.1	0.08	0.08
Biomarker	Clark P, 2008	AB blood group	13	5710/49069	1.02 (0.86-1.22)	1.00 (0.86-1.17)	0.82 (0.45-1.50)	0.46	18 (0-57) (0.26)	0.72-1.45	1	0.7	0.49	0.7	0.52	5.4	NP	NP
Biomarker	Clark P, 2008	A blood group	14	5047/44743	0.96 (0.85-1.07)	0.97 (0.91-1.03)	1.00 (0.81-1.24)	0.82	57 (5-75) (<0.01)	0.68-1.35	3	0.8	0.05	1.0	0.07	0.7	0.03	0.03
Biomarker	Clark P, 2008	B blood group	12	5324/48911	1.05 (0.94-1.18)	1.05 (0.95-1.15)	1.01 (0.72-1.42)	0.71	23 (0-61) (0.21)	0.82-1.35	0	0.9	NP	0.9	NP	0.6	NP	NP
Biomarker	Clark P, 2008	O blood group	18	5945/54609	1.01 (0.91-1.12)	1.00 (0.93-1.06)	0.98 (0.80-1.21)	0.52	49 (0-69) (0.01)	0.73-1.39	3	0.9	0.06	0.9	0.06	1.0	0.07	0.07
Environmental	Hu H, 2014	NO_2	5	3629/117497	1.10 (1.03-1.17)	1.10 (1.03-1.17)	1.06 (0.96-1.17)	0.12	0 (0-64) (0.73)	0.99-1.21	1	1.3	NP	1.3	NP	0.7	0.52	0.52
Environmental	Pedersen M, 2014	Air pollution	4	4905/165789	1.05 (0.99-1.13)	1.07 (1.03.1.11)	1.13 (1.07-1.19)	0.19	65 (0-86) (0.03)	0.79-1.40	1	1.0	NP	0.7	0.51	2.3	0.32	0.32
Environmental	Pedersen M, 2014	NOx	3	1385/48725	1.03 (0.91-1.17)	1.03 (0.91-1.17)	1.00 (0.87-1.15)	0.08	0 (0-73) (0.54)	0.46-2.28	0	0.2	NP	0.2	NP	0.2	NP	NP
Environmental	Pedersen M, 2014	PM_{10}	4	4656/201197	0.95 (0.86-1.05)	0.94 (0.91-0.98)	0.83 (0.77-0.89)	0.73	83 (41-92) (<0.01)	0.60-1.50	1	0.8	0.56	0.6	0.50	3.1	NP	NP
Environmental	Pedersen M, 2014	CO	3	3583/112308	1.10 (0.99-1.22)	1.10 (1.00-1.21)	1.18 (1.03-1.35)	0.94	24 (0-79) (0.27)	0.44-2.76	1	1.3	NP	1.3	NP	2.6	NP	NP
Environmental	Pedersen M, 2014	O_3	4	4943/164360	1.03 (1.00-1.06)	1.03 (1.00-1.06)	1.10 (0.94-1.30)	0.07	0 (0-68) (0.85)	0.98-1.09	0	0.4	NP	0.4	NP	1.7	0.14	0.14
Genetic markers	Zeng F, 2016	G894T	26	3241/6419	1.45 (1.09-1.94)	1.42 (1.17-1.74)	1.37 (0.92-2.04)	0.65	41 (0-62) (0.02)	0.55-3.86	4	9.5	NP	10.3	NP	8.0	0.13	0.13
Genetic markers	Zeng F, 2016	T-786C	15	2268/3100	1.25 (0.94-1.68)	1.33 (1.09-1.63)	2.57 (1.27-5.19)	0.14	46 (0-69) (0.02)	0.52-3.00	4	4.6	NP	3.1	0.53	14.6	NP	NP
Genetic markers	Zhang G, 2016	rs4762 in AGT gene	3	790/2492	0.95 (0.66-1.38)	0.93 (0.68-1.27)	1.07 (0.62-1.84)	0.20	26 (0-79) (0.26)	0.04-23.9	0	0.2	NP	0.2	NP	0.2	NP	NP
Genetic markers	Zhang G, 2016	rs18001133 in MTHFR	49	13356/23082	1.17 (1.05-1.31)	1.21 (1.14-1.27)	1.26 (1.04-1.53)	0.32	75 (67-81) (<0.01)	0.60-2.29	12	11.6	0.87	9.0	0.27	48.5	NP	NP
Genetic markers	Zhang G, 2016	rs6025 in F5 gene	28	8210/9834	1.53 (1.06-2.21)	1.60 (1.35-1.91)	1.73 (0.78-3.83)	0.61	74 (61-81) (<0.01)	0.28-8.41	9	20.4	NP	18.8	NP	22.4	NP	NP
Genetic markers	Zhang G, 2016	rs1800896 in IL-10 gene	9	3020/3786	0.91 (0.75-1.11)	1.00 (0.91-1.10)	1.15 (0.98-1.35)	0.04	70 (28-83) (<0.01)	0.50-1.68	4	0.5	0.00	1.0	0.01	1.8	0.09	0.09
Genetic markers	Zhang G, 2016	rs1800871 in IL-10 gene	4	978/2074	0.79 (0.59-1.07)	0.79 (0.66-0.94)	0.84 (0.63-1.11)	0.87	65 (0-86) (0.04)	0.23-2.75	1	1.5	NP	1.4	NP	0.9	NP	NP
Genetic markers	Zhang G, 2016	rs1137101 in LEPR gene	28	8210/9834	1.53 (1.06-2.21)	1.60 (1.35-1.91)	1.73 (0.78-3.83)	0.61	74 (61-81) (<0.01)	0.28-8.41	9	20.4	NP	18.8	NP	22.4	NP	NP
Genetic markers	Zhang G, 2016	rs18001131 in MTHFR gene	9	2780/3636	1.15 (0.93-1.40)	1.10 (0.97-1.24)	0.91 (0.64-1.29)	0.21	59 (0-78) (0.01)	0.63-2.07	2	0.9	0.24	1.5	0.65	0.9	0.24	0.24
Genetic markers	Li Y, 2015	A1675G of AT2R	5	972/3072	1.58 (1.05-2.37)	1.51 (1.15-1.98)	1.25 (0.82-1.90)	0.47	50 (0-80) (0.09)	0.47-5.35	1	2.0	0.65	2.4	0.38	0.8	0.58	0.58
Genetic markers	Yang W, 2014	IL-10 -1082 A/G	11	1741/3560	0.93 (0.77-1.13)	0.96 (0.86-1.07)	1.38 (0.62-3.09)	0.30	63 (13-79) (<0.01)	0.51-1.70	4	0.6	0.00	0.7	0.00	5.3	NP	NP
Genetic markers	Yang W, 2014	IL-10 -819 C/T	5	729/1146	1.28 (1.03-1.59)	1.28 (1.09-1.51)	1.19 (0.88-1.62)	0.86	41 (0-77) (0.15)	0.70-2.35	2	1.8	NP	1.8	NP	1.0	0.27	0.27
Genetic markers	Yang W, 2014	IL-10 -592 C/A	3	459/926	1.28 (1.03-1.59)	1.28 (1.03-1.59)	1.55 (1.04-2.30)	0.39	0 (0-73) (0.46)	0.31-5.26	1	1.0	NP	1.0	NP	0.9	NP	NP
Genetic markers	Wang X, 2014	G20210A SNP	16	2296 /3262	1.79 (1.23-2.61)	1.79 (1.23-2.61)	1.84 (0.51-6.57)	0.96	0 (0-45) (0.92)	1.18-2.71	2	11	NP	11	NP	15.7	NP	NP
Genetic markers	Wang X, 2014	V G1691A SNP	23	3131/4036	1.60 (1.25-2.06)	1.56 (1.24-1.95)	1.74 (0.78-3.89)	<0.01	15 (0-49) (0.25)	0.91-2.82	4	11	NP	11.8	NP	14.1	NP	NP
Genetic markers	Li X, 2014	MTHFR C677T	47	6238/11771	1.12 (1.04-1.22)	1.12 (1.04-1.21)	1.28 (0.98-1.66)	0.16	14 (0-40) (0.21)	0.90-1.41	1	3.8	0.18	3.9	0.18	9.4	NP	NP

Genetic markers	Gong LL, 2014	MMP9-1562C>T	5	712/766	0.93 (0.61-1.42)	0.98 (0.79-1.22)	0.82 (0.53-1.27)	0.34	72 (0-87) (<0.01)	0.22-3.97	1	0.3	0.23	0.3	0.28	0.8	0.57	0.57
Genetic markers	Li X, 2014	TGF-β 1 869 T >C	4	466/618	0.70 (0.57-0.86)	0.70 (0.57-0.86)	0.64 (0.39-1.03)	0.93	0 (0-68) (0.84)	0.45-1.09	2	1.4	0.61	1.4	0.61	2.0	NP	NP
Genetic markers	Buurma AJ, 2013	AGT rs4762	5	497/1395	1.24 (0.67-2.30)	1.11 (0.85-1.46)	1.07 (0.62-1.84)	0.31	80 (37-90) (<0.01)	0.13-11.49	1	0.4	0.32	0.8	0.57	0.3	0.27	0.27
Genetic markers	Buurma AJ, 2013	APOE rs429358, rs7412	7	554/712	0.86 (0.65-1.13)	0.86 (0.66-1.12)	0.96 (0.60-1.55)	0.04	4 (0-60) (0.40)	0.57-1.29	0	0.6	NP	0.6	NP	0.4	NP	NP
Genetic markers	Buurma AJ, 2013	AT1R rs5186	9	886/1230	1.12 (0.95-1.33)	1.12 (0.95-1.33)	0.96 (0.69-1.34)	0.33	0 (0-54) 0.78)	0.91-1.37	0	0.7	NP	0.7	NP	0.5	NP	NP
Genetic markers	Buurma AJ, 2013	CTLA4 rs231775	4	353/536	1.25 (1.01-1.56)	1.25 (1.02-1.53)	1.14 (0.80-1.61)	0.82	14 (0.72) (0.32)	0.68-2.29	1	0.6	0.45	0.6	0.53	0.32	0.28	0.28
Genetic markers	Buurma AJ, 2013	LPL rs1800590	3	395/579	2.27 (0.63-8.21)	1.60 (0.85-2.99)	0.81 (0.36-1.80)	0.12	71 (0-89) (0.03)	0-5626855	1	1.7	0.58	2.7	NP	0.52	0.43	0.43
Genetic markers	Buurma AJ, 2013	LPL rs268	4	530/933	2.43 (1.26-4.68)	2.44 (1.38-4.32)	1.34 (0.51-3.50)	0.66	20 (0-74) (0.29)	0.35-17.1	2	3.9	NP	3.9	NP	1.2	0.59	0.59
Genetic markers	Buurma AJ, 2013	NOS3 27 bp-VNTR in intron 4	14	1593/2239	1.14 (0.90-1.43)	1.06 (0.93-1.21)	0.96 (0.71-1.30)	0.03	63 (23-78) (<0.01)	0.53-2.47	4	0.8	0.01	1.2	0.03	0.8	0.01	0.01
Genetic markers	Buurma AJ, 2013	NOS3 rs2070744	11	1571/2202	1.08 (0.95-1.23)	1.11 (0.99-1.23)	1.21 (0.96-1.52)	0.10	28 (0-64) (0.18)	0.80-1.46	2	0.9	0.21	0.7	0.17	1.7	0.69	0.69
Genetic markers	Buurma AJ, 2013	NOS3 rs1799983	24	2825/4048	1.19 (1.00-1.42)	1.21 (1.10-1.34)	1.79 (1.37-2.34)	0.55	68 (49-78) (<0.01)	0.56-2.52	7	3.3	0.04	2.9	0.02	15.9	NP	NP
Genetic markers	Buurma AJ, 2013	TLR4 rs4986790	4	723/614	1.07 (0.48-2.39)	1.06 (0.73-1.54)	3.03 (1.36-6.72)	0.92	78 (0-90) (<0.01)	0.03-38.2	1	0.2	0.21	0.3	0.22	4	NP	NP
Genetic markers	Buurma AJ, 2013	TLR4 rs4986791	3	614/461	1.20 (0.45-3.17)	1.10 (0.71-1.72)	2.92 (1.31-6.49)	0.59	79 (0-91) (<0.01)	0-123082	1	0.2	0.21	0.4	0.37	3	NP	NP
Genetic markers	Buurma AJ, 2013	TNF-alpha rs1800629	12	1592/1837	1.17 (0.91-1.49)	1.19 (1.02-1.39)	1.61 (1.17-2.22)	0.48	54 (0-74) (0.01)	0.56-2.41	3	1.4	0.15	1.2	0.12	6.3	0.08	0.08
Genetic markers	Buurma AJ, 2013	TNF-alpha rs1799724	3	390/385	0.66 (0.34-1.30)	0.77 (0.60-0.99)	1.18 (0.84-1.66)	0.51	84 (15-93) (<0.01)	0-2313	2	0.6	0.10	1.3	0.57	0.3	0.03	0.03
Genetic markers	Buurma AJ, 2013	VEGF rs3025039	3	377/514	1.36 (0.64-2.90)	1.21 (0.94-1.55)	0.73 (0.51-1.03)	0.69	87 (47-94) (<0.01)	0-13603	1	0.4	0.37	0.9	NP	0.9	NP	NP
Genetic markers	Cheng D, 2013	VEGF +936 C/T	8	805/1033	1.52 (1.09-2.12)	1.45 (1.22-1.72)	0.73 (0.51-1.03)	0.58	69 (15-83) (<0.01)	0.54-4.23	3	2.5	0.72	3.0	NP	2.0	0.42	0.42
Genetic markers	Song GG, 2013	VEGF - 634 C/G	6	408/479	1.35 (1.09-1.67)	1.35 (1.10-1.65)	2.04 (1.33-3.13)	0.86	12 (0-66) (0.34)	0.90-2.01	1	1.0	NP	1.0	NP	3.8	NP	NP
Genetic markers	Song GG, 2013	VEGF -2578 A/ C	8	617/672	0.93 (0.78-1.10)	0.93 (0.79-1.09)	1.05 (0.78-1.41)	0.99	13 (0-62) (0.33)	0.68-1.26	0	0.5	NP	0.5	NP	0.4	NP	NP
Genetic markers	Song GG, 2013	VEGF -1154 A/G	3	159/161	1.14 (0.83-1.56)	1.14 (0.83-1.56)	1.06 (0.69-1.64)	0.45	0 (0-73) (0.89)	0.15-8.86	0	0.2	NP	0.2	NP	0.2	NP	NP
Genetic markers	Morgan JA, 2013	PAI-1 (-675 4G/4G)	12	1511/ 3492	1.28 (1.09-1.50)	1.28 (1.09-1.50)	1.19 (0.77-1.84)	0.56	0 (0-50) (0.63)	1.07-1.53	2	2.4	NP	2.4	NP	1.5	0.65	0.65
Genetic markers	Dai B, 2013	eNOS 4 b/a	10	1374/1376	1.43 (0.87-2.37)	1.37 (0.93-2.03)	1.77 (0.80-3.92)	0.37	30 (0-66) (0.17)	0.45-4.55	1	2.8	0.30	3.4	0.18	6.1	NP	NP
Genetic markers	Zhao L, 2013	SERPINE1 -675 4G/5G	11	1297/1791	1.37 (1.10-1.71)	1.36 (1.13-1.64)	1.66 (1.10-2.51)	0.42	20 (0-60) (0.25)	0.88-2.15	2	2.8	0.74	2.9	0.74	5.9	NP	NP
Genetic markers	Staines-Urias E, 2012	F5 rs6025	41	4499/15188	1.74 (1.50-2.02)	1.74 (1.50-2.02)	1.67 (0.61-4.61)	0.56	0 (0-33) (0.53)	1.49-2.03	6	21.9	0.00	21.9	0.00	20	0.00	0.00
Genetic markers	Staines-Urias E, 2012	F2 rs1799963	30	3546/11712	1.72 (1.40-2.12)	1.72 (1.40-2.12)	1.45 (0.67-3.14)	0.03	0 (0-37) (0.55)	1.38-2.14	2	16.6	NP	16.6	NP	10.5	NP	NP
Genetic markers	Staines-Urias E, 2012	ACE rs4646994	30	3101/5134	1.17 (1.03-1.34)	1.10 (1.03-1.18)	1.03 (0.86-1.22)	0.06	68 (51-77) (<0.01)	0.65-2.13	6	2.1	0.01	3.1	0.12	1.6	0.00	0.00
Genetic markers	Staines-Urias E, 2012	AGT rs699	27	2329/4896	1.26 (1.05-1.51)	1.21 (1.10-1.32)	1.31 (0.70-2.45)	0.32	70 (53-79) (<0.01)	0.57-2.79	6	3	0.11	3.9	0.27	4.8	0.61	0.61
Genetic markers	Staines-Urias E, 2012	MTHFR rs1801133	51	5160/10151	1.06 (0.99-1.15)	1.04 (0.99-1.10)	1.21 (0.68-2.13)	0.03	38 (7-55) (<0.01)	0.79-1.49	7	2.7	0.02	2.9	0.03	6.4	0.68	0.68
Genetic markers	Staines-Urias E, 2012	SERPINE1 rs1799889	12	1194/1757	0.89 (0.77-1.04)	0.90 (0.80-1.00)	0.90 (0.64-1.27)	0.42	40 (0-68) (0.76)	0.59-1.33	2	0.9	0.22	0.9	0.24	0.9	0.21	0.21
Genetic markers	Staines-Urias E, 2012	EPHX1 rs1051740	4	562/462	0.85 (0.72-1.00)	0.85 (0.72-1.00)	0.94 (0.72-1.23)	0.87	0 (0-68) (0.51)	0.59-1.24	0	0.4	NP	0.4	NP	0.2	NP	NP
Genetic markers	Staines-Urias E, 2012	EPHX1 rs2234922	3	425/427	1.28 (0.83-1.96)	1.32 (1.01-1.73)	1.87 (1.23-2.83)	0.26	60 (0-87) (0.08)	0.01-134	1	0.7	0.57	0.6	0.49	2.4	0.10	0.10
Genetic markers	Staines-Urias E, 2012	PPARG rs1801282	3	390/449	0.80 (0.57-1.12)	0.80 (0.57-1.12)	0.81 (0.43-1.51)	0.07	0 (0-73) (0.90)	0.09-7.35	0	0.5	NP	0.5	NP	0.5	NP	NP
Genetic markers	Staines-Urias E, 2012	THBD C1418T	3	260/268	0.71 (0.49-1.03)	0.71 (0.49-1.03)	0.78 (0.52-1.15)	0.30	0 (0-73) (0.50)	0.07-7.73	0	0.7	NP	0.7	NP	0.5	NP	NP
Genetic markers	Staines-Urias E, 2012	IL-6 rs1800795	3	248/1575	0.91 (0.70-1.19)	0.91 (0.70-1.19)	0.91 (0.42-1.94)	0.76	0 (0-73) (0.90)	0.16-5.13	0	0.2	NP	0.2	NP	0.2	NP	NP
Genetic markers	Staines-Urias E, 2012	VEGFA rs699947	3	225/269	0.88 (0.69-1.14)	0.88 (0.69-1.14)	0.92 (0.61-1.38)	0.69	0 (0-73) (0.90)	0.17-4.52	0	0.2	NP	0.2	NP	0.2	NP	NP
Genetic markers	Staines-Urias E, 2012	HLA-G -14 bp	3	219/334	1.42 (0.68-2.98)	1.37 (1.06-1.79)	0.97 (0.68-1.38)	0.90	85 (28-93) (<0.01)	0-11540	1	0.6	0.51	0.7	0.57	0.2	0.15	0.15
Genetic markers	Staines-Urias E, 2012	LEP rs7799039	3	198/326	1.51 (0.92-2.49)	1.38 (1.06-1.79)	1.20 (0.85-1.71)	0.43	68 (0-89) (0.05)	0.01-412	1	0.6	0.49	0.9	NP	0.3	0.26	0.26
Genetic markers	Staines-Urias E, 2012	LEP TTTC	3	141/227	0.86 (0.53-1.38)	0.92 (0.68-1.24)	1.01 (0.68-1.51)	0.42	56 (0-86) (0.10)	0.01-135	1	0.2	0.16	0.2	0.20	0.2	0.14	0.14
Genetic markers	Lin R, 2012	AGT M235T	29	5053/11578	1.61 (1.21-2.14)	1.53 (1.26-1.84)	1.40 (0.32-6.06)	0.47	45 (6-64) (<0.01)	0.57-4.52	4	16	NP	18.2	NP	8.5	NP	NP
Genetic markers	Lin R, 2012	AGT T174M	6	1362/4159	1.09 (0.76-1.57)	1.16 (0.91-1.49)	0.97 (0.54-1.74)	0.35	48 (0-78) (0.09)	0.40-2.95	1	1.0	NP	0.6	0.44	0.3	0.29	0.29
Genetic markers	Zhao L, 2012	AGTR1 +1166A>C	10	845/1150	1.19 (0.96-1.47)	1.19 (1.00-1.42)	1.15 (0.67-1.99)	0.42	27 (0-64) (0.20)	0.74-1.91	1	1.5	NP	1.5	NP	1.2	NP	NP
Genetic markers	Zhong WG, 2012	ACE D/I	11	1600/1898	1.93 (1.19-3.12)	1.72 (1.49-1.99)	0.87 (0.59-1.28)	0.26	91 (86-93) (<0.01)	0.31-12.1	8	7.8	NP	9.2	0.41	1.1	0.00	0.00
Genetic markers	Shaik AP, 2011	ACE (II genotype)	16	1620/2158	0.99 (0.70-1.40)	0.95 (0.80-1.12)	0.94 (0.57-1.54)	0.79	73 (52-82) (<0.01)	0.27-3.56	3	0.9	0.06	0.8	0.04	0.9	0.06	0.06
Genetic markers	Xie C, 2011	TNF-α 308 G/A	18	1888/2497	0.98 (0.76-1.25)	0.93 (0.80-1.09)	0.56 (0.36-0.87)	0.56	52 (5-71) (<0.01)	0.43-2.21	3	1.1	0.09	0.9	0.06	11	NP	NP
Genetic markers	Xie C, 2011	IL-6 -174 G/C	4	396/507	1.23 (0.93-1.61)	1.23 (0.93-1.61)	1.44 (0.89-2.33)	0.44	0 (0-68) (0.81)	0.67-2.24	0	0.5	NP	0.5	NP	1.3	0.32	0.32
Genetic markers	Rodger MA, 2010	FVL	9	1060/20773	1.26 (0.91-1.74)	1.26 (0.91-1.74)	1.27 (0.51-3.14)	0.27	0 (0-54) (0.99)	0.85-1.86	0	2.3	0.12	2.3	0.12	2.5	0.07	0.07
Genetic markers	Rodger MA, 2010	PGM	6	549/13705	1.27 (0.80-2.03)	1.27 (0.80-2.03)	1.03 (0.41-2.56)	0.30	0 (0-61) (0.99)	0.65-2.46	0	1.4	0.35	1.4	0.35	0.32	NP	NP
Genetic markers	Medica I, 2007	AGT/T704C (Met235Thr)	15	1146/2276	1.66 (1.20-2.29)	1.66 (1.23-2.25)	0.29 (0.03-2.58)	0.77	6 (0-50) (0.38)	1.00-2.73	2	5.6	0.06	5.6	0.06	12.9	NP	NP
Genetic markers	Serrano NC, 2006	ACE-I/D	22	2596/3828	1.23 (1.04-1.45)	1.12 (1.01-1.23)	0.90 (0.73-1.11)	0.01	57 (23-72) (<0.01)	0.66-2.26	4	1.7	0.09	3.3	0.56	1.7	0.08	0.08
Genetic markers	Lin J, 2005	FLV (1691 G-A)	11	1135/1471	2.25 (1.28-3.94)	2.40 (1.70-3.39)	2.21 (1.06-4.59)	0.43	57 (0-76) (<0.01)	0.42-12.2	5	9.2	NP	8.7	NP	8.5	NP	NP
Diseases/disorders	Saccone G, 2015	Celiac disease	5	14618/507559	2.05 (0.89-4.74)	1.80 (1.44-2.24)	1.19 (0.79-1.78)	0.66	90 (79-94) (<0.01)	0.11-40.1	3	3.7	0.61	3.9	0.31	2.7	NP	NP
Diseases/disorders	Zhang JJ, 2015	Chronic kidney disease	9	14993/504700	10.4 (6.28-17.1)	11.1 (9.00-13.7)	22.3 (15.6-31.9)	0.71	77 (50-87) (<0.01)	2.12-50.7	8	8.9	0.08	8.9	0.10	9.0	NP	NP
Diseases/disorders	Hu R, 2015	Depression	5	1104/2874	1.66 (1.29-2.13)	1.64 (1.32-2.03)	1.12 (0.64-1.96)	0.34	16 (0-69) (0.32)	0.96-2.86	4	3.7	NP	3.8	NP	0.6	0.00	0.00
Diseases/disorders	Qin JZ, 2013	Polycystic ovary syndrome	15	1866/1194098	3.26 (2.06-5.16)	2.14 (1.88-2.43)	2.04 (1.78-2.34)	<0.01	41 (0-66) (0.05)	1.02-10.43	5	10	NP	13.7	NP	9.4	NP	NP
Diseases/disorders	Zhang S, 2013	Mental stress	12	16705/649188	1.49 (1.27-1.74)	1.28 (1.20-1.35)	1.14 (1.05-1.24)	0.02	68 (32-81) (<0.01)	0.97-2.29	6	4.2	0.37	7.3	0.56	2.4	0.02	0.02
Diseases/disorders	Zhang S, 2013	Work stress	4	496/8246	1.50 (1.15-1.97)	1.50 (1.15-1.97)	1.51 (0.99-2.31)	0.98	0 (0-68) (0.75)	0.83-2.72	0	2.4	NP	2.4	NP	2.4	NP	NP
Diseases/disorders	Zhang S, 2013	Depression and anxiety	5	753/7489	1.88 (1.08-3.25)	1.73 (1.32-2.28)	0.93 (0.55-1.59)	0.44	73 (0-87) (<0.01)	0.28-12.65	2	3.8	0.10	4.1	NP	0.3	0.04	0.04
Diseases/disorders	Grigoriadis S, 2013	Maternal depression	4	227/8843	1.35 (0.95-1.92)	1.34 (0.97-1.86)	1.24 (0.77-2.00)	0.46	7 (0-70) (0.36)	0.56-3.26	0.9	NP	0.9	NP	0.6	0.46	0.9	0.9
Supplementation	Schoenaker DA, 2014	Calcium intake	3	387/1100	0.88 (0.60-1.29)	0.88 (0.60-1.29)	0.89 (0.53-1.52)	0.87	0 (0-73) (0.99)	0.07-10.82	0	0.3	NP	0.3	NP	0.3	NP	NP
Infections	Huang QT, 2016	Chronic hepatitis B infection	11	14298/423216	0.79 (0.63-1.00)	0.79 (0.67-0.93)	1.13 (0.78-1.63)	0.90	20 (0-61) (0.25)	0.51-1.25	3	7.0	NP	6.9	NP	4.1	0.76	0.76
Infections	Sgolastra F, 2013	Periodontal disease	15	1040/3983	2.17 (1.38-3.41)	1.93 (1.59-2.33)	2.05 (1.47-2.86)	0.50	78 (61-85) (<0.01)	0.42-11.29	8	7.9	NP	9.2	0.60	8.6	0.80	0.80
Infections	Rustveld LO, 2008	Bacterial & viral infections	21	2390/11556	2.08 (1.63-2.66)	2.03 (1.78-2.31)	1.78 (1.18-2.67)	0.65	56 (20-72) (<0.01)	0.92-4.72	10	12.4	0.38	12.8	0.26	10.3	NP	NP

Other	Xu Y, 2016	Isolated single umbilical artery	3	783/64443	0.82 (0.56-1.21)	0.82 (0.56-1.21)	0.84 (0.56-1.26)	0.50	0 (0-73) (0.85)	0.07-9.96	0	1.0	0.56	1.0	0.56	0.85	0.56	0.56
Other	Wei J, 2015	Cigarette smoking	17	62089/1784382	0.67 (0.60-0.75)	0.71 (0.69-0.72)	0.87 (0.83-0.91)	0.36	92 (89-94) (<0.01)	0.43-1.05	14	15.1	0.42	15.5	0.19	10.7	0.13	0.13
Other	Masoudian P, 2015	Oocyte donation vs ART	13	1499/25299	2.54 (1.98-3.24)	2.61 (2.12-3.22)	3.15 (2.27-4.37)	0.90	14 (0-55) (0.31)	1.61-4.00	6	8.5	0.15	8.3	0.25	9.8	0.02	0.02
Other	Masoudian P, 2015	Oocyte donation vs NC	4	2712/54816	4.33 (3.11-6.03)	4.08 (3.16-5.26)	3.35 (2.42-4.63)	0.26	26 (0-75) 0.26)	1.52-12.4	4	3.9	NP	4.0	NP	3.8	NP	NP
Other	Aune D, 2014	Pre-pregnancy PA high vs low activity	5	621/9696	0.65 (0.45-0.94)	0.65 (0.45-0.94)	0.60 (0.30-1.20)	0.63	0 (0-64) (0.91)	0.36-1.19	0	3.2	NP	3.2	NP	3.8	NP	NP
Other	Aune D, 2014	Pre-pregnancy PA per 1hr per day	3	479/6002	0.73 (0.53-0.99)	0.73 (0.53-0.99)	0.36 (0.07-1.88)	0.09	0 (0-73) (0.69)	0.10-5.42	0	1.6	0.11	1.6	0.11	3.0	NP	NP
Other	Aune D, 2014	Early pregnancy PA high vs low activity	11	5702/162900	0.79 (0.70-0.91)	0.79 (0.70-0.91)	1.03 (0.74-1.44)	0.90	0 (0-51) (0.55)	0.68-0.92	2	4.3	0.22	4.3	0.22	0.7	0.16	0.16
Other	Aune D, 2014	Early pregnancy PA per 20 MET hrs/week	3	2576/85388	0.86 (0.70-1.07)	0.93 (0.85-1.02)	0.98 (0.89-1.09)	0.30	68 (0-89) (0.04)	0.07-9.95	1	0.6	0.47	1.3	NP	0.2	0.17	0.17
Other	Aune D, 2014	Early pregnancy PA per 1hr per day	7	5293/151083	0.83 (0.73-0.95)	0.84 (0.75-0.94)	0.95 (0.80-1.14)	0.66	20 (0-66) (0.28)	0.63-1.09	3	2.9	NP	3.1	NP	0.8	0.04	0.04
Other	Aune D, 2014	Early pregnancy walking	4	535/9674	0.68 (0.51-0.89)	0.68 (0.51-0.89)	1.00 (0.43-2.33)	0.09	0 (0-68) (0.75)	0.37-1.24	1	2.4	0.31	2.4	0.31	0.2	0.19	0.19
Other	Aune D, 2014	Early pregnancy occupational PA	6	620/18119	0.82 (0.66-1.03)	0.82 (0.66-1.03)	0.75 (0.52-1.07)	0.78	0 (0-61) (0.68)	0.60-1.13	0	1.1	0.60	1.1	0.60	2.0	0.19	0.19
Other	González CM, 2014	Donor insemination	7	2342/8556	1.57 (1.01-2.42)	1.65 (1.38-1.98)	1.69 (1.38-2.08)	0.82	49 (0-77) (0.07)	0.52-4.70	3	3.9	0.71	3.5	0.72	4.0	0.47	0.47
Other	Wang Z, 2013	Obese vs normal weight women (adjusted)	10	34340/1685991	2.93 (2.58-3.33)	3.26 (3.16-3.37)	3.64 (2.54-5.21)	0.11	67 (20-81) (<0.01)	2.07-4.15	10	9.9	NP	9.9	NP	10	NP	NP
Other	Wang Z, 2013	Severe obese vs normal weight women	6	19976/877162	3.12 (2.24-4.37)	2.86 (2.71-3.01)	2.53 (2.32-2.76)	0.60	97 (95-97) (<0.01)	0.96-10.2	6	6	NP	6	NP	6	NP	NP
Other	Kasawara KT, 2012	Physical activity (case-control)	6	923/8481	0.77 (0.53-1.11)	0.76 (0.64-0.91)	1.16 (0.72-1.86)	0.93	76 (30-88) (<0.01)	0.23-2.60	3	2.1	0.42	2.0	0.41	0.8	0.04	0.04
Other	Kasawara KT, 2012	Physical activity (cohort studies)	10	5547/178680	0.94 (0.83-1.07)	0.99 (0.94-1.05)	1.10 (1.01-1.19)	0.17	60 (0-78) (<0.01)	0.67-1.32	3	0.5	0.01	1.3	0.13	2.1	0.45	0.45
Other	Basaran A, 2011	CVS vs no invasive	6	1189/46410	0.83 (0.42-1.66)	0.59 (0.50-0.70)	0.83 (0.61-1.13)	0.29	92 (87-95) (<0.01)	0.07-9.29	3	4.0	0.41	1.6	0.19	1.62	0.35	0.35
Other	Basaran A, 2011	CVS vs no invasive & amniocentesis	7	1320/56266	1.00 (0.46-2.17)	0.81 (0.69-0.94)	0.83 (0.61-1.13)	0.49	96 (94-97) (<0.01)	0.06-16	4	2.2	0.21	0.4	0.00	1.8	0.08	0.08
Other	Luo ZC, 2007	Primiparity	23	54462/1966490	2.42 (2.16-2.71)	2.33 (2.28-2.37)	2.27 (2.22-2.32)	0.58	92 (90-94) (0)	1.47-3.97	22	21.9	NP	22	NP	21.7	NP	NP

Abbreviations: Random effects, summary odds ratio (95% CI) using random effects model; Fixed effects, summary odds ratio (95% CI) using fixed effects model; Largest effect, odds ratio (95% CI) of the largest study in the meta-analysis; Egger, p-value from Egger's regression asymmetry test for evaluation of publication bias; P, p-value; O, observed number of "positive" studies; E, expected number of "positive" studies; NP, not pertinent, because the estimated E is larger than the O, thus there is no evidence of excess statistical significance based on the assumption made for the plausible effect size; β-hCG, Human chorionic gonadotropin; PAPP-A, pregnancy-associated plasma protein A; PLGF, Placental growth factor; PP13, Placental Protein 13; sFlt-1, Soluble fms-like tyrosine kinase-1; HLA, Human leukocyte antigen; PIGF, placental growth factor; VEGF, vascular endothelial growth factor; sENG, soluble endoglin; NO_2, Nitrogen dioxide; NOx, Mono-nitrogen oxides; PM_{10}, Particulate matter 10 micrometers; CO, Carbon Monoxide; O_3, Ozone; IL-6, Interleukin 6; LEPR, leptin receptor; IL-18, Interleukin-18; IFN-γ, Interferon gamma; AT2R, Angiotensin type 2 receptor; IL-10, Interleukin 10; SNP, Single-nucleotide polymorphisms; MTHFR, Methylene tetrahydrofolate reductase; MMP-9, Matrix metallopeptidase 9; PAI-1, Plasminogen activator inhibitor-1; AGT, Angiotensinogen; AGTR1, Angiotensin II Receptor Type 1; ACE, Angiotensin; eNOS, Endothelial nitric oxide synthase; TNF-a, Tumor necrosis factor alpha; FVL, Factor V Leiden; PGM, Prothrombin Gene Mutation; CVS, chorionic villus sampling; ART, assisted reproductive technology; NC, natural conception; PA, physical activity; NP, not pertinent, because the estimated is larger than the observed, and there is no evidence of excess of statistical significance based on the assumption made for the plausible effect size.

* Summary random effects odds ratio (95% CI) of each meta-analysis, except for three meta-analyses (Wei J 2015, Aune D, 2014 and Wang Z, 2013) where the RR was used.
† Summary fixed effects odds ratio (95% CI) of each meta-analysis, except for three meta-analyses (Wei J 2015, Aune D, 2014 and Wang Z, 2013) where the RR was used.
‡ Odds ratio (95% CI) of the largest study in each meta-analysis, except for three meta-analyses (Wei J 2015, Aune D, 2014 and Wang Z, 2013) where the RR was used.
§ P-value from the Egger regression asymmetry test for evaluation of publication bias
|| I^2 metric of inconsistency (95% confidence intervals of I^2) and P-value of the Cochran Q test for evaluation of heterogeneity
≠ 95% Prediction Interval
¶ Observed number of statistically significant studies
# Expected number of statistically significant studies using the summary fixed effects estimate of each meta-analysis as the plausible effect size
** P-value of the excess statistical significance test
¥ Expected number of statistically significant studies using the summary random effects estimate of each meta-analysis as the plausible effect size
8 Expected number of statistically significant studies using the effect of the largest study of each meta-analysis as the plausible effect size

Chapter 5 – Randomized clinical trials for preventing preeclampsia: an umbrella review of the literature

5.1 Abstract

Background: Preeclampsia is a severe pregnancy-associated disease, which is characterized by the occurrence of hypertension and proteinuria in previously healthy women after the 20th weeks of gestation. Although numerous systematic reviews and meta-analyses have been published examining the association between various pharmacologic and non-pharmacologic interventions for the prevention of preeclampsia, the epidemiological credibility of these associations has not been thoroughly assessed. The objective of this study is to summaries evidence and evaluates the strength and validity in the reported associations for preeclampsia prevention.

Methods: An umbrella review was performed to identify systematic reviews and meta-analyses of randomized controlled trials evaluating the association of various interventions for preeclampsia prevention. For each association, we estimated the summary effect size by random-effects and fixed-effects models, the 95% confidence interval, and the 95% prediction interval. We also assessed the between-study heterogeneity, evidence for small study effects and excess significance bias. We further applied standardized methodological criteria to evaluate the epidemiological credibility of the statistically significant associations.

Results: Twenty-nine eligible meta-analyses were identified that included 456 primary studies, providing data on 57 associations. Twenty-four (42%) associations had nominally statistically significant findings at $p<0.05$, while only 10 (18%) were significant at $p<10^{-3}$ under the random-effects model. Sixteen (28%) associations had large or very large heterogeneity. Evidence of excess significance bias was found in 15 (26%) associations. After applying our classification criteria, the following three

interventions were classified as Class I level of evidence including low dose aspirin ≤16 weeks of gestation for preterm preeclampsia, diet and nutrition counselling and dietary interventions.

Conclusions: Early administration of aspirin in women with preterm preeclampsia, diet and nutrition counselling, and dietary interventions present the strongest consistent evidence. The findings from our study highlight the importance of patient education on diet and lifestyle modifications in reducing the risk of preeclampsia, as well as the recommendation for early administration of aspirin in women at high risk pregnancies.

5.2 Introduction

Preeclampsia (PE) is a severe pregnancy-associated disease, which is characterized by the occurrence of hypertension and proteinuria after the 20th weeks of gestation in previously healthy women. Based on recent data, PE affects approximately 2-8% of all pregnancies and is associated with substantially higher maternal and fetal morbidity and mortality worldwide (284,285). The clinical spectrum of PE varies from a mild form of the disease, characterized by a moderate increase in blood pressure and proteinuria, to the most severe characterized by seizures as a sign of damage of the cerebral vessels and HELLP (Hemolysis, Elevated Liver enzyme, Low Platelets) syndrome, a life-threating condition for the pregnant women and their fetuses (286). Until today, the true etiology of PE remains unclear, which generates uncertainty on prediction, prevention and treatment.

Many pharmacologic and non-pharmacologic interventions have been studied for the prevention of PE, including antioxidants, vitamins, dietary salt restrictions, diuretics for fluid control, fish oil, calcium supplementation, aspirin, and heparin. However, in some cases, evidence was not sufficient to support their recommendation (394,395). In contrast, administration of low-dose aspirin before 16 weeks of gestation in high-risk pregnancies and calcium, especially in low-calcium intake populations remain the only strategies associated with a definitive reduction in risk (289,396). On the other hand, the available screening tools for risk stratification for PE are sub-optimal. The development of effective prevention strategies of PE has proved difficult due to the uncertainty in disease etiology and its multifactorial and complex nature (289,397) as well as due to the limitations of the current predictive tests (398,399,213). Since the prevalence of early-onset PE is relatively low, screening tests are required to perform

better in terms of their sensitivity and specificity in order to produce meaningful positive predictive values.

In view of the importance of guidelines for PE prevention, the assessment of the credibility of the available evidence could have significant implications both for clinical practice and for public health in more general. Instead of looking at limited indication-specific data, it much more useful to have a wider view of the evidence across many indications where the effects of interventions for PE prevention have been assessed. This can be performed in the setting of an umbrella review which collects and evaluates evidence from multiple resources systematically and gives an overview of the strengths, weaknesses, and biases of this literature at-large (22,400). We performed an umbrella review of the evidence across published meta-analyses or systematic reviews of randomized controlled trials (RCTs) for PE prevention in order to provide a comprehensive summary of the range of interventions, present the magnitude, direction, significance of the reported associations and effects, assess the potential biases, and identify the associations and effects that present the most convincing epidemiological evidence.

5.3 Methods

Literature search and selection criteria

We conducted an umbrella review, defined as a comprehensive and systematic collection and evaluation of published systematic reviews and meta-analyses performed on a specific research topic (22). The methods of the umbrella review are standardized and for this work we followed the same principles as previously described in published umbrella reviews conducted on various fields of research (265–269). We

used a ranking system to grade the evidence from meta-analyses of RCTs in terms of the significance of the summary effect, the 95% prediction interval, presence of large heterogeneity, small study effects, and excess significance bias.

We systematically searched PubMed, the Cochrane Database of Systematic Reviews, and ISI Web of Science up to April 7, 2017, to identify systematic reviews and meta-analyses of randomized trials of interventions the prevention of PE. We searched for the keywords ("pre-eclampsia" OR "preeclampsia") AND ("systematic review" OR "meta-analysis"). All identified publications underwent a parallel, three-step review of title, abstract, and full text (performed by KG and SP) based on predefined inclusion and exclusion criteria. We also screened the references of the retrieved articles for possible eligible papers.

We included systematic reviews and meta-analyses of randomized controlled trials that examined the association of a respective intervention related to PE prevention. Meta-analyses and systematic reviews were retained if they included at least 2 studies in which information was provided on a measure of association and its standard error and on the number of events and the number of participants. We did not apply any language restrictions in the selection of eligible studies. We included only meta-analyses and systematic reviews of randomized clinical trials in humans. If an article presented separate meta-analyses on other medical conditions, in addition to PE, we only extracted information on the latter. When more than one meta-analysis on the same research question was eligible, the meta-analysis with the largest number of component studies with data on individual studies' effect sizes was retained for the main analysis. We excluded narrative reviews, letters to the editor, meta-analyses of

non-RCTs and systematic reviews without a quantitative synthesis of data. We also did not include the older version of two meta-analyses that were published by the same authors on the same intervention when there was only a 2–3 years difference between the two versions.

Data extraction

Data extraction was performed independently by two investigators (KG and SP), and in case of discrepancies, the final decision was reached by discussion or by having a third investigator (EE) review the study, when necessary. For each article we extracted data regarding the first author's name, publication year, number of studies included, intervention administered, total number of participants per treatment arm, events in each arm, the reported summary risk estimates (risk ratio and odds ratio) with 95% confidence intervals (CI), model used for analysis (fixed or random), and the heterogeneity statistic (I^2).

Statistical analysis

For each unique meta-analysis, we estimated the summary effect and its 95% CI by using both fixed and random effect models (94,270). We also estimated the 95% prediction intervals (PI) for the summary random effects estimates, which further account for between-study heterogeneity and indicates the uncertainty for the effect that would be expected in the new study examining the same association (271,294). For the largest study of each meta-analysis, we calculated the standard error (SE) of the effect size, and we examined whether the SE was less than 0.10 and whether the largest study presented a statistically significant effect. In a study with SE of less than 0.10, the difference between the effect estimate and the upper or lower 95% confidence

interval is less than 0.20 (i.e. this uncertainty is less than what is considered a small effect size).

We assessed heterogeneity among studies and we reported the P value of the χ^2-based Cochran Q test and the I^2 metric for inconsistency, which could reflect either diversity or bias. I^2 ranges between 0% and 100% and quantifies the variability in effect estimates that is due to heterogeneity rather than the sampling error (273). Values exceeding 50% or 75% are usually considered representing large or very large heterogeneity, respectively. Its confidence intervals were calculated as per Ioannidis et al. (2007) (274).

We assessed whether there is evidence for a small-study effect (i.e. whether smaller studies tend to give substantially larger estimates of effect size compared to larger studies). We used the regression asymmetry test proposed by Egger for this assessment (276). A P value <0.10 accompanied by a more conservative effect in larger studies was considered evidence of the presence of small-study effects.

We further applied the excess significant test to evaluate whether there is a relative excess of formally significant findings in published literature due to any reason (such as publication bias, selective reporting of outcomes or analyses). This is a chi-squared based test, in which the expected number of positive studies is estimated and compared against the observed number of studies with statistically significant results (P<0.05) (68). A binomial test was used to evaluate whether the number of positive studies in a meta-analysis is too large according to the power that these studies have to detect plausible effects at α=0.05. A comparison between the number of observed vs expected

is performed separately for each meta-analysis and it is also extended to groups of many meta-analyses after summing the observed and expected from each meta-analysis. The expected number of significant studies for each meta-analysis is calculated by the sum of the statistical power estimates for each component study (68). The power of each component study was estimated using the fixed effects summary, the random effects summary, or the effect size of the largest study (smallest SE) as the plausible effect size (72). The power of each study was calculated with an algorithm using a non-central t distribution.(401) Excess statistical significance for single meta-analyses was set at $p<0.10$ (one-sided $p<0.05$, with observed > expected as previously proposed) (68).

Assessment of epidemiologic credibility

We used a ranking system to grade the evidence. Evidence from meta-analyses of RCTs was assessed in terms of the significance of the summary effect ($P<0.001$, $0.001\leq P<0.05$, $p\geq0.05$), 95% prediction interval (excluding the null or not), and presence of large heterogeneity (I^2 >50%), small study effects ($P>0.10$), and excess significance ($P<0.05$). We also noted the conclusions from any evidence classification with the use of GRADE (Grading of Recommendations Assessment, Development and Evaluation) or any other equivalent system applied by the authors of the original meta-analyses.

Studies that reported a p-value of less than 0.001, had a 95% prediction interval not including the null, had no evidence of small-study effects or no evidence of excess significance, and did not have large heterogeneity were considered as representing robust evidence of effectiveness of interventions (Class I). Meta-analyses that had a p-

value less than 0.001 and the largest study reporting a significant effect were considered to have the next best quality of evidence (Class II). Finally, meta-analyses with only a p-value of less than 0.05 were classified as quality of evidence Class III. The statistical analysis and the power calculations were performed in STATA version 14 (STATA Corp, College Station, TX).

5.4 Results

Description of Eligible Meta-analyses

Overall, the literature search identified 683 publications of which, 29 articles were deemed eligible (Figure 5.1). The publication date of the eligible articles ranged between 2007 and 2017. The 29 eligible papers (291,402–429), included data on 57 different meta-analyses (comparisons) in five broad areas (antiplatelets [n=16 comparisons], vitamins supplements [n=6 comparisons], diet and life-style interventions [n=12 comparisons], calcium supplementation [n=14 comparisons], and other drugs [n=9 comparisons]). There were between 2 to 43 studies per meta-analysis, with a median of five studies. The median number of participants in each study was 195, while the median number of events and participants in each meta-analysis was 336 and 4358, respectively. The number of events was greater than 1000 in 10 meta-analyses (Table 5.1). Supplementary Table 5.4 summarizes these 57 meta-analyses that included 456 individual study estimates.

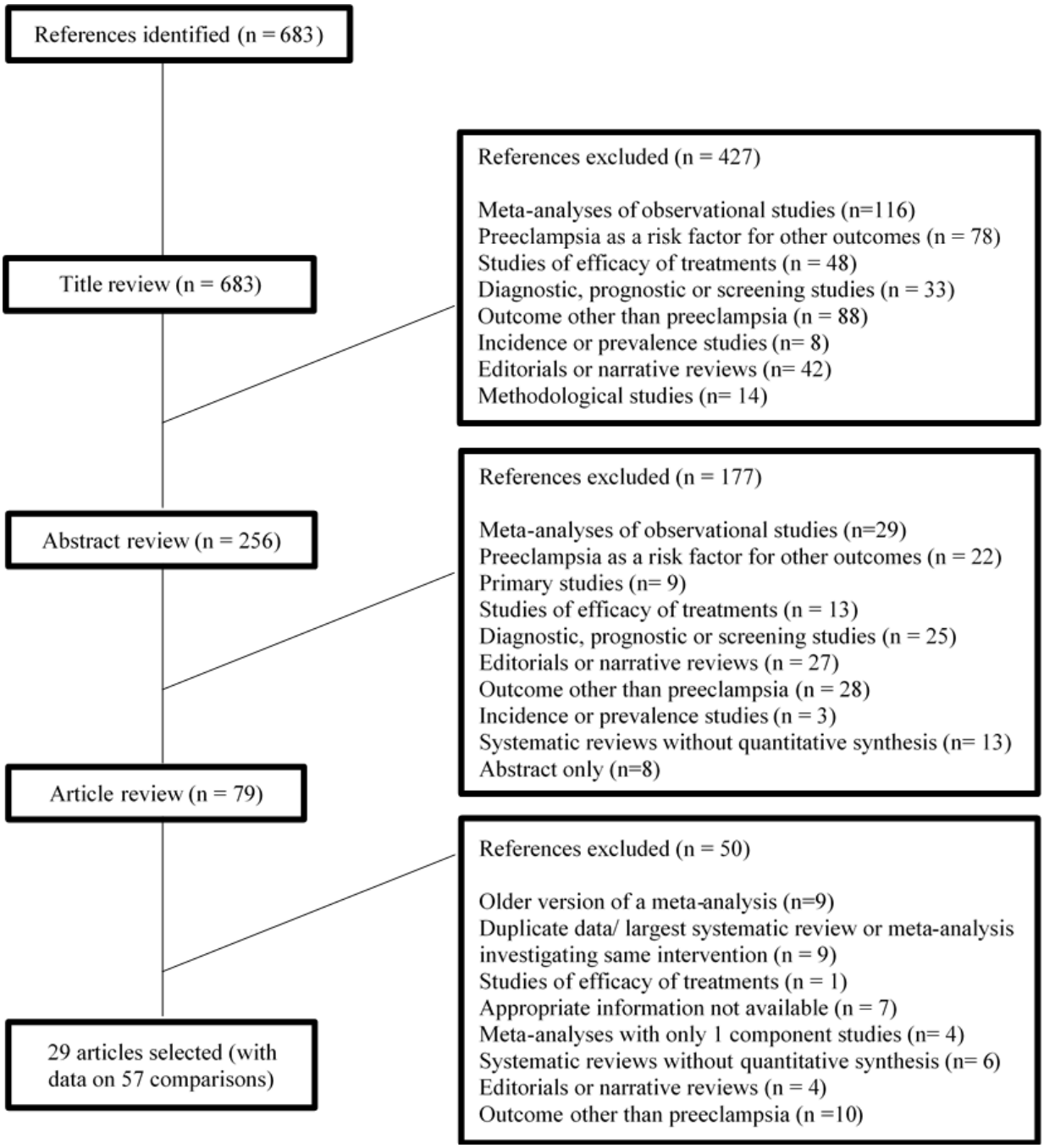

Figure 5.1. Flowchart of the included studies

Table 5.1. Quantitative synthesis and assessment of bias across the 57 associations of interventions for preeclampsia prevention

Area	Author, year	Comparison	Studies	Events/participants	Random effects*	Largest effect‡	P Random	Egger§	I^2 (P)‖	95% PI ≠
Antiplatelets	Roberge S 2017	Aspirin < 16 weeks (severe PE)	9	231/4194	0.50 (0.29-0.86)	0.96 (0.67-1.37)	0.014	0.001	18 (0.03)	0.12-2.09
Antiplatelets	Roberge S 2017	Aspirin >16 weeks	21	1103/15571	0.83 (0.68-1.01)	1.23 (0.90-1.68)	0.056	0.006	42 (0.02)	0.48-1.43
Antiplatelets	Roberge S 2016	Low dose aspirin ≤ 16 weeks (60mg) vs pl	3	281/3293	0.93 (0.75-1.15)	1.05 (0.69-1.60)	0.508	0.606	0 (0.79)	0.23-3.77
Antiplatelets	Roberge S 2016	Low dose aspirin > 16 weeks (60 mg) vs pl	3	601/8483	0.93 (0.70-1.23)	1.23 (0.90-1.68)	0.622	0.895	66 (0.05)	0.04-22.40
Antiplatelets	Roberge S 2016	Low dose aspirin vs placebo	6	882/11776	0.94 (0.81-1.09)	1.23 (0.90-1.68)	0.386	0.980	22 (0.27)	0.68-1.29
Antiplatelets	Roberge S 2016	LMWH and low-dose aspirin or aspirin alone	5	54/590	0.54 (0.32-0.92)	0.35 (0.14-0.86)	0.023	0.649	0 (0.68)	0.23-1.28
Antiplatelets	Henderson JT 2014	Aspirin vs placebo	13	1977/21865	0.78 (0.64-0.95)	0.88 (0.75-1.03)	0.015	0.002	36 (0.09)	0.50-1.21
Antiplatelets	Villa PM 2013	Aspirin ≤16 weeks (abnormal uterine artery flow)	3	97/346	0.55 (0.36-0.83)	0.57 (0.40-0.82)	5×10^{-3}	0.631	16 (0.31)	0.02-17.67
Antiplatelets	Dodd JM 2013	Heparin (alone or with other) vs no treatment	7	91/761	0.47 (0.210.1.01)	0.35 (0.14-0.86)	0.052	0.957	58 (0.03)	0.05-4.27
Antiplatelets	Roberge S 2012	Low-dose aspirin ≤16 weeks for preterm PE	5	45/556	0.11 (0.03-0.33)	0.11 (0.01-0.86)	1×10^{-4}	0.850	0 (0.72)	0.02-0.68
Antiplatelets	Trivedi NA 2011	Low-dose aspirin in low risk women	5	729/16550	0.87 (0.64-1.17)	1.14 (0.94-1.38)	0.349	0.170	67 (0.02)	0.33-2.29
Antiplatelets	Trivedi NA 2011	Low-dose aspirin in high risk women	14	1365/11687	0.79 (0.65-0.97)	0.88 (0.75-1.04)	0.024	0.059	50 (0.02)	0.47-1.33
Antiplatelets	Bujold E 2009	Aspirin vs placebo in women with AUAD	9	245/1317	0.67 (0.47-0.94)	0.95 (0.67-1.35)	0.021	0.450	36 (0.13)	0.30-1.47
Antiplatelets	Duley L 2007	Antiplatelet agents vs pl (moderate risk women)	25	1625/28509	0.77 (0.64-0.92)	0.88 (0.74-1.03)	5×10^{-3}	0.000	43 (0.01)	0.46-1.30
Antiplatelets	Duley L 2007	Antiplatelet agents vs pl (high risk women)	18	748/4121	0.60 (0.45-0.81)	0.91 (0.77-1.06)	6×10^{-4}	0.002	42 (0.03)	0.29-1.27
Antiplatelets	Duley L 2007	Antiplatelet agents vs placebo	43	2373/32590	0.72 (0.62-0.83)	0.88 (0.75-1.03)	1×10^{-5}	0.000	43 (<0.01)	0.43-1.19
Other drugs	Chen B 2015	Fish oil vs control (low risk)	7	155/3720	0.82 (0.53-1.26)	0.87 (0.60-1.25)	0.367	0.617	13 (0.33)	0.36-1.85
Other drugs	Chen B 2015	Fish oil vs control (high risk)	5	129/1965	1.04 (0.72-1.50)	0.96 (0.53-1.76)	0.850	0.226	15 (0.32)	0.47-2.30
Other drugs	Chen B 2015	Fish oil vs control	12	413/7650	0.93 (0.72-1.21)	0.87 (0.60-1.25)	0.586	0.979	10 (0.34)	0.60-1.45
Other drugs	Makrides M 2014	Magnesium supplementation vs no magnesium	3	78/1042	0.88 (0.58-1.34)	1.04 (0.15-7.35)	0.557	0.613	0 (0.60)	0.06-12.83
Other drugs	Gui S 2014	L-arginine vs placebo	2	125/524	0.38 (0.25-0.60)	0.34 (0.21-0.55)	2×10^{-5}	NA	4 (0.31)	NA
Other drugs	Rumbold A 2008	Antioxidants versus control	9	586/5446	0.72 (0.49-1.04)	0.97 (0.80-1.17)	0.083	0.057	58 (0.02)	0.27-1.90
Other drugs	Meher S 2007	Nitric oxide vs placebo/control	4	42/170	0.78 (0.37-1.66)	1.35 (0.61-3.01)	0.521	0.341	37 (0.19)	0.06-10.48
Other drugs	Imhoff-Kunsch 2012	n-3 LCPUFA supplementation vs control	4	93/1683	0.80 (0.44-1.46)	1.15 (0.66-1.99)	0.473	0.027	38 (0.18)	0.10-6.29
Other drugs	Allen R 2014	Essential fatty acids supplementation	6	226/4579	0.88 (0.63-1.24)	1.04 (0.73-1.48)	0.474	0.007	24 (0.25)	0.42-1.86
Vitamins	Rumbold A 2015	Any vitamin E supplementation vs pl	14	1965/20878	0.91 (0.79-1.06)	1.07 (0.93-1.24)	0.213	0.004	47 (0.02)	0.62-1.35
Vitamins	Rumbold A 2015	Vitamin C alone or with other supplements vs pl	16	2003/21956	0.92 (0.80-1.05)	1.07 (0.93-1.24)	0.213	0.010	41 (0.04)	0.64-1.31
Vitamins	De-Regil LM 2015	Vitamin D alone versus no treatment/placebo	2	25/219	0.52 (0.25-1.07)	0.53 (0.25-1.10)	0.075	NA	0 (0.79)	NA
Vitamins	De-Regil LM 2015	Vitamin D + calcium vs no treatment/placebo	3	78/1114	0.50 (0.32-0.80)	0.39 (0.21-0.73)	4×10^{-3}	0.658	0 (0.47)	0.03-9.95
Vitamins	Pérez-López FR 2015	Vitamin D intervention vs pl	3	47/654	0.92 (0.45-1.87)	0.67 (0.33-1.35)	0.815	0.943	22 (0.28)	0.00-400.2
Vitamins	Conde-Agudelo A 2011	Supplementation with vitamins C and E	9	1903/19810	0.99 (0.90-1.09)	1.07 (0.93-1.24)	0.860	0.113	13 (0.32)	0.83-1.18
Diet & life-style	Zheng J 2017	Exercise	2	35/1009	1.05 (0.53-2.08)	1.00 (0.49-2.03)	0.883	NA	0 (0.62)	NA
Diet & life-style	Muktabhant B 2015	Diet and exercise counselling	8	177/3139	0.99 (0.74-1.31)	1.05 (0.73-1.51)	0.936	0.018	0 (0.95)	0.69-1.41
Diet & life-style	Muktabhant B 2015	Supervised exercise	3	47/1024	0.91 (0.52-1.60)	1.00 (0.51-1.97)	0.754	0.463	0 (0.76)	0.02-34.2
Diet & life-style	Muktabhant B 2015	Unsupervised exercise	2	8/229	1.60 (0.38-6.70)	1.34 (0.27-6.72)	0.518	NA	0 (0.63)	NA
Diet & life-style	Muktabhant B 2015	Diet counselling/other	4	54/634	0.90 (0.54-1.47)	2.69 (0.55-13.0)	0.664	0.878	0 (0.44)	0.30-2.67
Diet & life-style	Muktabhant B 2015	All diet and/or exercise vs standard/other care	18	336/5280	0.95 (0.77-1.16)	1.05 (0.73-1.51)	0.596	0.337	0 (0.99)	0.76-1.18
Diet & life-style	Allen R, 2014	Diet and nutrition counseling	6	249/2695	0.68 (0.54-0.86)	0.65 (0.48-0.88)	1×10^{-3}	0.699	0 (0.61)	0.49-0.95
Diet & life-style	Allen R, 2014	Mixed interventi (diet, physical activity, lifestyle)	6	113/1438	0.92 (0.64-1.31)	1.00 (0.55-1.79)	0.625	0.691	0 (0.59)	0.55-1.51
Diet & life-style	Allen R, 2014	All type of interventions	18	588/8712	0.81 (0.69-0.96)	1.04 (0.73-1.48)	0.015	0.583	4 (0.41)	0.64-1.03

Diet & life-style	Thangaratinam S 2012	Dietary interventions	6	249/2624	0.68 (0.54-0.86)	0.65 (0.48-0.88)	1×10^{-3}	0.788	0 (0.55)	0.48-0.95
Diet & life-style	Thangaratinam S 2012	Mixed approach	3	16/369	1.40 (0.49-3.95)	2.69 (0.55-13.03)	0.529	0.948	0 (0.37)	0.00-1179
Diet & life-style	Thangaratinam S 2012	All interventions (diet, mixed, physical activity)	10	272/3072	0.78 (0.56-1.09)	0.65 (0.48-0.88)	0.143	0.125	22 (0.24)	0.40-1.52
Calcium	An LB 2015	Calcium supplementation vs placebo	4	754/1452	0.86 (0.69-1.05)	0.92 (0.75-1.13)	0.134	0.045	37 (0.19)	0.42-1.74
Calcium	Tang R, 2015	Calcium supplementation vs placebo	10	1513/2478	0.62 (0.48-0.81)	0.94 (0.81-1.09)	5×10^{-4}	0.000	72 (<0.01)	0.29-1.32
Calcium	Tang R, 2015	Low baseline calcium	6	494/1053	0.42 (0.23-0.76)	0.92 (0.75-1.12)	4×10^{-3}	0.002	77 (<0.01)	0.06-2.75
Calcium	Tang R, 2015	High baseline calcium	2	359/5045	0.70 (0.34-1.44)	0.94 (0.76-1.16)	0.333	NA	74 (0.05)	NA
Calcium	Tang R, 2015	Unknown baseline calcium	2	660/9208	0.47 (0.08-2.84)	0.94 (0.81-1.09)	0.412	NA	72 (0.06)	NA
Calcium	Tang R, 2015	Calcium supplements vs pl (High risk of PE)	4	410/8665	0.36 (0.14-0.98)	0.92 (0.75-1.12)	0.045	0.021	79 (<0.01)	0.01-25.6
Calcium	Tang R, 2015	Calcium supplements vs pl (Normal risk of PE)	6	1103/16122	0.67 (0.48-0.92)	0.94 (0.81-1.09)	0.012	0.004	72 (<0.01)	0.27-1.67
Calcium	Hofmeyr GJ 2014	Calcium supplements vs pl (Adequate calcium diet)	4	366/5022	0.61 (0.32-1.19)	0.94 (0.77-1.16)	0.148	0.097	51 (0.10)	0.05-6.85
Calcium	Hofmeyr GJ, 2014	Calcium supplements vs pl (Low calcium diet)	8	515/10678	0.35 (0.20-0.64)	0.92 (0.75-1.13)	5×10^{-4}	0.000	76 (<0.01)	0.06-2.12
Calcium	Hofmeyr GJ, 2014	Routine high-dose calcium supplements	13	889/15730	0.44(0.31-0.64)	0.92 (0.75-1.13)	1×10^{-5}	0.000	70 (<0.01)	0.15-1.28
Calcium	Patrelli TS 2012	Adequate calcium intake vs placebo	6	700/9641	0.78 (0.58-1.06)	0.94 (0.76.1.16)	0.116	0.007	49 (0.08)	0.37-1.65
Calcium	Patrelli TS 2012	Low calcium intake vs placebo	7	474/10154	0.35 (0.18-0.68)	0.92 (0.75-1.13)	2×10^{-3}	0.000	75 (<0.01)	0.05-2.59
Calcium	Patrelli TS 2012	Calcium supplements vs pl (high risk)	3	41/346	0.17 (0.07-0.42)	0.21 (0.07-0.58)	9×10^{-5}	0.095	0 (0.80)	0.00-50.1
Calcium	Patrelli TS 2012	Calcium supplements vs pl (low risk)	7	515/11059	0.51 (0.30-0.86)	0.92 (0.75-1.13)	0.012	0.059	75 (<0.01)	0.10-2.55

Abbreviations: Random effects, summary risk ratio (95% CI) using random effects model; Largest effect, risk ratio (95% CI) of the largest study in the meta-analysis; Egger, p-value from Egger's regression asymmetry test for evaluation of publication bias; P, p-value; NP, not pertinent, because the estimated is larger than the observed, and there is no evidence of excess of statistical significance based on the assumption made for the plausible effect size; PE, Preeclampsia; LMWH, low-molecular-weight heparin; AUAD, abnormal uterine artery Doppler; PA, Physical activity.

* Summary random effects risk ratio (95% CI) of each meta-analysis.
‡ Risk ratio (95% CI) of the largest study in each meta-analysis, except for two meta-analyses.
§ P-value from the Egger regression asymmetry test for evaluation of publication bias
|| I^2 metric of inconsistency and P-value of the Cochran Q test for evaluation of heterogeneity
≠ 95% Prediction Interval

Summary Effect Sizes and Significant Findings

Of the 57 meta-analyses, 24 (42%) had nominally statistically significant findings at $P<0.05$ using the random effects model, and all showed a protective effect of the intervention against developing PE. Out of these, 10 (18%) associations presented statistically significant effect at $P<0.001$ (Table 5.1). The ten interventions that presented a significant effect at $P<10^{-3}$ for PE prevention were; low-dose aspirin ≤16 weeks for preterm PE, antiplatelet agents (heparin or aspirin) vs placebo in high risk women, antiplatelet agents vs. placebo in the general population, L-arginine vs. placebo, diet and nutrition counseling, calcium supplementation vs. placebo, calcium supplementation vs. placebo in women with low calcium diet, routine high-dose calcium supplementation, dietary interventions, and calcium supplementation vs. placebo (high risk women). Additional information on all 57 meta-analyses is available online (Supplementary Table 5.4).

Between-Study Heterogeneity and Prediction Intervals

Thirteen (23%) meta-analyses had large heterogeneity estimates ($50\% \leq I^2 \leq 75\%$) and 3 (5%) had very large heterogeneity estimates ($I^2 > 75\%$) (Table 5.1). The 3 meta-analyses where the I^2 exceeded 75% included meta-analyses of calcium supplementation. Uncertainty around heterogeneity estimates was often large, especially when the number of individual studies was limited, and is reflected by wide 95% confidence intervals of I^2. When we calculated the 95% prediction intervals, the null value was excluded in only 3 (5%) of the included meta-analyses. These were for low-dose aspirin ≤16 weeks for preterm PE, diet and nutrition counseling and dietary interventions (Table 5.1).

Small-Study Effects and Excess Significance Bias

Evidence of statistically significant small-study effect (Egger test $p<0.10$ and the random effects summary estimate was larger compared to the point estimate of the largest study in the meta-analysis) was not identified in any of the included meta-analyses (Supplementary Table 5.4). Fifteen (26%) of the associations had hints of excess statistical significance bias with statistically significant ($P<0.05$) excess of positive studies under any of the three assumptions for the plausible effect size, namely the fixed effects summary, the random effects summary, and the results of the largest study (Supplementary Table 5.4). Eight (14%) pertained to calcium supplementation, five (9%) pertained to the antiplatelets, one (2%) pertained to vitamins, and one (2%) pertained to other drugs. Table 5.2 shows the results of excess of statistical significance bias according to category of intervention.

Table 5.2. Observed and expected number of positive studies by type of intervention*

Area	No. of studies	Observed positive	Expected positive (fixed) †	P‡ (fixed)	Expected positive (random)§	P‡ (random)	Expected positive (largest)‖	P‡ (largest)	Expected positive (composite) ¶	P‡ (composite)
All	456	110	57.27	0.00	99.95	0.26	54.73	0.00	54.28	0.00
Antiplatelets	189	46	24.35	0.00	37.71	0.14	26.05	0.00	24.35	0.00
Other drugs	52	4	4.60	0.77	6.34	0.40	4.67	NP	4.60	0.77
Vitamins	47	8	3.90	0.06	4.88	0.15	5.43	0.25	3.90	0.06
Diet & life-style	86	8	8.45	NP	8.05	NP	10.52	0.51	8.05	NP
Calcium	82	44	15.98	0.00	42.97	0.83	7.61	0.00	7.61	0.00

* NP, not pertinent, because the estimated is larger than the observed, and there is no evidence of excess of statistical significance based on the assumption made for the plausible effect size.

† Expected number of statistically significant studies using the summary fixed effects estimate of each meta-analysis as the plausible effect size.

‡ P value of the excess of statistically significant test. All statistical tests were two-sided.

§ Expected number of statistically significant studies using the summary random effects estimate of each meta-analysis as the plausible effect size.

‖ Expected number of statistically significant studies using the effect of the largest study of each meta-analysis as the plausible effect size.

¶ Expected number of statistically significant studies using the most conservative of the three estimates (fixed effects summary, random effects summary, largest study) of each meta-analysis as the plausible effect size.

Epidemiological credibility of findings

After applying our classification criteria, 3 interventions were classified as Class I level of evidence. These were low dose aspirin ≤16 weeks of gestation for preterm PE, diet and nutrition counseling and dietary interventions. In the original meta-analyses, the included studies were characterized as having low risk of bias by using the Cochrane Handbook Criteria and GRADE tools, therefore the quality of evidence supports the findings. Two associations, L-arginine vs. placebo, calcium supplementation vs. placebo (high risk group) presented Class II evidence for PE prevention. The quality assessment for the intervention of L-arginine vs. placebo was graded as regular to high quality using the Jadad scale. Moreover, there were only two studies included in this meta-analysis. The meta-analysis of calcium supplementation vs. placebo (high risk group) did not perform any quality assessment. Nineteen interventions were supported by Class III evidence. An overall assessment of statistically significant associations for PE prevention is presented in Table 5.3.

Table 5.3. Assessment across the statistically significant associations for preeclampsia prevention

Level of evidence	**Criteria**
Class I	[a] $P<10^{-3}$, not large heterogeneity (I^2 <50%), 95% prediction interval excluding the null value, no evidence for small-study effects [b] and excess significance bias [c]
Interventions supported by Class I evidence	Low-dose aspirin ≤16 weeks for preterm PE, diet and nutrition counseling, dietary interventions
Class II	[a] $P<10^{-3}$ and nominally statistically significant effect present at the largest study
Interventions supported by Class II evidence	L-arginine vs placebo, calcium supplementation vs placebo (high risk women)
Class III	The rest associations with [a] $P < 0.05$
Interventions supported by Class III evidence	Aspirin < 16 weeks (severe PE), LMWH and low-dose aspirin or aspirin alone, Aspirin vs placebo, Aspirin ≤16 weeks (abnormal uterine artery flow), Low-dose aspirin in high risk women, Aspirin vs placebo in women with AUAD, Antiplatelet agents vs pl (moderate risk women), Antiplatelet agents vs pl (high risk women), Antiplatelet agents vs placebo, Vitamin D + calcium vs no treatment/placebo, All type of interventions (diet, PA, lifestyle), Calcium supplementation vs placebo, Low baseline calcium, Calcium supplementation (High risk of PE), Calcium supplementation (Normal risk of PE), Calcium supplements vs pl (Low calcium diet), Routine high-dose calcium supplementation, Low calcium intake vs placebo, Calcium supplementation vs placebo (low risk)

Abbreviations: PE, Preeclampsia; LMWH, low-molecular-weight heparin; AUAD, abnormal uterine artery Doppler; PA, Physical activity.

[a] P indicates the P-values of the meta-analysis random effects model.
[b] Small study effect is based on the P-value from the Egger's regression asymmetry test (P< 0.10).
[c] Based on the P-value (P<0.05) of the excess significance test using the largest study (smallest standard error) in a meta-analysis as the plausible effect size.

5.5 Discussion

In this study, we provided a comprehensive overview of the reported associations between a wide range of interventions for PE prevention by incorporating evidence from systematic reviews and meta-analyses of RCTs. We also evaluated the reported evidence using criteria previously applied in other research fields (265–269). Our study is comprised of 29 meta-analyses of RCTs, which covered 57 interventions.

Main Findings and possible explanations

Overall, meta-analyses on 57 pharmacologic and non-pharmacologic interventions for PE prevention were assessed, including antiplatelets, vitamins, diet and lifestyle interventions, calcium supplementation, and other drugs. Of those, low dose aspirin ≤16 weeks of gestation for preterm PE, diet and nutrition counseling and dietary interventions had strong evidence (Class I) for prevention of PE, as specified by the epidemiological credibility criteria. Another two interventions, namely L-arginine vs. placebo and, calcium supplementation vs placebo (high risk group) presented Class II evidence for PE prevention.

Our results are in agreement with the latest evidence from large multi-centered randomized trials (430) as well as form the recent Individual Patient Data (IPD) meta-analysis (290), which demonstrated a significant benefit of early administration of low dose aspirin in women at high risk of developing PE. However, the available tools for risk stratification in the population are currently an issue of debate as a result of the lack of consensus on exactly what high risk translates to, what the characteristics of women who will benefit from this intervention are and what the magnitude of this benefit is. The study of Rolnik et al. (2017) (430) used a previously developed

algorithm that combined maternal factors, mean arterial pressure, uterine-artery pulsatility index, maternal serum pregnancy-associated plasma protein A and placental growth factor (213). However, given that this algorithm includes sophisticated ultrasound and biochemical markers, we have to acknowledge the fact that not all centers have access to specialists in fetal ultrasound or laboratories that will analyze a broad range of biomarkers.

The American College of Obstetricians and Gynecologists (ACOG) supports the recommendations of the United States Preventive Services Task Force (USPSTF) of daily low-dose aspirin (81 mg) beginning at 12 weeks of gestation in patients who are considered to be at high risk for PE. This recommendation is based on observational data. The ACOG reaffirmed in 2017 that the risk should be assessed on clinical criteria only, including primiparity, personal or family history of PE, chronic hypertension, type 1 or 2 diabetes, renal or autoimmune disease, in vitro fertilization, obesity, systemic lupus erythematosus, maternal age ≥40 years, history of thrombophilia, and carrying a multifetal gestation (431,432), given that previous proposed models were too optimistic with a high false positive rate. The National Institute for Health and Care Excellence (NICE) guidelines recommend 75 mg of aspirin daily from 12 weeks in high-risk women, including hypertensive disorders during previous pregnancy, chronic kidney disease, autoimmune disease such as systemic lupus erythrematosis or antiphospholipid syndrome. This is supported by a recent evaluation of early-pregnancy clinical risk factors (433). Our umbrella review confirms and supports the results from the previous evaluations that aspirin is an effective intervention for women that are destined to develop PE. But, even at the narrow range of clinical risk

factors, the amount of contribution from each and the possible effect modifications are yet to be determined.

The exact mechanism by which aspirin acts to prevent PE remains unclear. Based on the hypothesis of abnormal placentation such an effect would result to a restriction in the platelet aggregation and contraction of arterial smooth muscle that potentially improves the pathophysiological implications of PE (425,434). This is a possible reason why early administration (<16 weeks of gestation) has proven to be more effective as opposed to late administration.

Dietary and lifestyle interventions in pregnancy may also reduce the risk of PE and they have the advantage of being sustainable and cost-effective, albeit the most difficult to implement with success. Typical dietary interventions included a balanced diet consisting of carbohydrates, proteins and fat, calorie-controlled or low-fat diet, and keeping of a food diary. It is possible that dietary interventions are effective in modifying metabolic factors such as lipid levels, blood pressure, and glycose or reducing gestational weight gain with a potential contribution to a lower risk for PE (418). This is highlighting the importance of patient education on nutrition and general lifestyle in preventing not only PE, but other important co-morbidities, such as gestational diabetes. Keeping a normal weight gain in pregnancy has been proven to have a beneficial preventive effect in minimizing adverse pregnancy and neonatal outcomes in general (435). Given the fact that one adverse outcome is increasing the risk of other adverse outcomes, the actions taken on preventing one can have a beneficial effect in developing another. For example, raised triglyceride concentrations in pregnancy are associated with the risk of PE (436). It has been shown that an

increase in dietary total fiber intake reduces the levels of triglycerides which consequently reduce the risk in PE (437). Also, diets based on low fat meat and dairy products, whole grains, fruit, vegetables, and fish from the second trimester until delivery is effective in reducing maternal total and low-density lipoprotein cholesterol (438).

Over the last three decades, epidemiological evidence has suggested an association between low calcium intake and PE (439,440,219). This relationship is supported by the fact that the incidence of PE is low in populations with elevated mean calcium intake (e.g. South America and Ethiopia) and by the fact that women with PE have blood and urine calcium levels lower than normotensive pregnant women (420,441). Additional intake of calcium during pregnancy could also reduce the incidence of PE, especially in populations at high risk of PE due to ethnicity, gender, age and high Body Mass Index (BMI) (408,442). It has been proposed that low-calcium intake increases blood pressure by stimulating either the parathyroid hormone or renin release, thus increasing intracellular calcium in vascular smooth muscle and thus leading to vasoconstriction (439). Calcium supplementation could possibly reduce parathyroid release, smooth muscle contractility or increase serum magnesium levels and thus prevent preterm labour and delivery (443,444). The specific characteristics of the populations that could benefit from calcium supplementation are still not clear.

Circulating L-arginine, an essential amino acid, is the substrate of nitric oxide (NO), a potent vasodilator, which has an important role in regulating blood pressure, maintaining the stabilization of homeostasis, cardiovascular activity, and immune responses (406,445). Administration of L-arginine seems to improve uterine-placental

circulation and reduce maternal blood pressure (446–448), and thus aid to reduce oxidative stress, a key factor in PE pathogenesis (449–451). Hence, L-arginine could be a potential therapeutic option for pregnant women with hypertension. However, further large-scale RCTs are needed to draw a definitive conclusion as the enrolled trials in the meta-analysis in our assessment were only two with quite small sample sizes.

To claim discovery of novel findings, researchers widely use a p-value threshold at the level of $P<0.05$. However, findings based on this threshold can only constitute weak evidence in many cases, as suggested by ongoing discussions to redefine the level of statistical significance using more stringent criteria (452). As shown in this paper, even though 42% of the examined associations claim a statistically significant finding at $P<0.05$, only 9% of the eligible associations provided convincing or highly suggestive evidence. Recently, prominent scientists have proposed changing the threshold of statistical significance to 0.005 for studies that examine the null hypothesis aiming to increase statistical standards of evidence for claiming new discoveries and improve reproducibility in many fields of science (452).

We acknowledge some limitations of our work. Umbrella reviews focus on existing meta-analyses and therefore interventions that were not assessed in a previous meta-analysis were not included in our review. Also, it is possible that for some types of interventions, only meta-analyses of observational data exist with no respective randomized evidence and these would not have been captured by our search. Moreover, although our analysis identified diet and nutrition counseling and dietary interventions to had strong epidemiological credibility for prevention of PE, yet, some

of the included studies had a large proportion of obese pregnant women, hence, results should be interpreted with caution. In addition, due to the heterogeneity of both the pathophysiological pathways and clinical presentations of PE, it is possible that our results to be modified based on the presence of the other risk factors such as diabetes and obesity which are associated with cardiovascular disease.

Furthermore, we did not appraise the quality of the individual studies directly, since this was beyond the scope of the current umbrella review. Jadad scale for quality assessment that was used for quality assessment is outdated and this needs to be considered in the overall evaluation of the evidence. Furthermore, both Egger and excess of significance test that we used offer hints of bias, not a definitive proof thereof. The Egger test is difficult to interpret when the between-study heterogeneity is large. The interpretation of the excess of statistical significance test for the results of a single meta-analysis, especially in those with few studies, should be cautious because a negative test does not exclude potential bias (68). Lastly, we cannot exclude the possibility of selective reporting in several trials as typically some interventions are more likely to be reported, if they had statistically significant results.

This umbrella review supports the administration of low dose aspirin in early PE (less than 34 weeks), to women at high-risk for preterm PE. We must underline the fact that PE is not a single disease entity and early versus late PE has different risk profiles, recurrence risks and responses to therapy. We did not find robust epidemiologic evidence for aspirin use in the entire spectrum of PE. On the other hand, diet and lifestyle interventions are measures that can be used for the benefit of the overall cardiovascular health of women. Given the obvious similarity of most clinical PE risk

factors with cardiovascular disease risk factors, pregnancy might be a crucial opportunity of reducing women's risk not only for PE, but the life-time risk for cardiovascular events.

5.6 Conclusion

Early administration of low dose aspirin ≤16 weeks of gestation for prevention of early PE and patient education on a balanced diet and nutrition during pregnancy seem to be effective preventive measures for PE in high risk women. Future research should focus on developing useful and effective screening tools, to have a uniformity of risk stratification in multiple populations, test the proposed prevention measures in large scale studies and evaluate the best cost-effective options in every day clinical practice. Policymakers and clinical experts should be aware of possible biases in published meta-analyses and they should scrutinize all the available evidence to increase the validity of their recommendations.

Contribution to authorship: KG and SP were involved in formulating the hypothesis and the design of the study protocol. KG and SP performed the literature search, the selection of eligible articles and the data extraction. KG analyzed the data. All authors (KG, EE, SP, CC, NM and PY) were involved in data interpretation. KG and SP wrote the first draft of the manuscript and EE, CC, NM and PY were involved in the revision of the manuscript. All authors (KG, EE, SP, CC, NM, PY) approved the final version of the submitted manuscript. KG and SP are guarantors.

Chapter 5: Supplemental material

Supplemental Table 5.4. Analytical description of the 57 selected meta-analyses with observed and expected number of "positive" study datasets

Area	Author, year	Comparison	N‡	Events /participants	Random effects*	Fixed effects†	Largest effect‡	Egger§	I^2 (95% CI) (P)‖	95% PI ≠	O¶	E #	P** (fixed)	E ¥	P** (random)	E ϑ	P** (largest)	P** (largest)
Antiplatelets	Roberge S 2017	Aspirin < 16 weeks (severe PE)	9	231/4194	0.50 (0.29-0.86)	0.76 (0.58-0.99)	0.96 (0.67-1.37)	0.001	18 (0-77) (0.03)	0.12-2.09	2	1.06	0.29	3.22	0.50	0.46	0.07	0.07
Antiplatelets	Roberge S 2017	Aspirin >16 weeks	21	1103/15571	0.83 (0.68-1.01)	0.91 (0.81-1.02)	1.23 (0.90-1.68)	0.006	42 (0-64) (0.02)	0.48-1.43	4	1.36	0.04	2.34	0.28	2.57	0.32	0.32
Antiplatelets	Roberge S 2016	Low dose aspirin ≤ 16 weeks (60mg)	3	281/3293	0.93 (0.75-1.15)	0.93 (0.75-1.15)	1.05 (0.69-1.60)	0.606	0 (0-73) (0.79)	0.23-3.77	0	0.19	NP	0.19	NP	0.17	NP	NP
Antiplatelets	Roberge S 2016	Low dose aspirin > 16 weeks (60 mg)	3	601/8483	0.93 (0.70-1.23)	0.94 (0.81-1.09)	1.23 (0.90-1.68)	0.895	66 (0-88) (0.05)	0.04-22.40	1	0.22	0.20	0.25	0.98	0.98	NP	NP
Antiplatelets	Roberge S 2016	Low dose aspirin vs placebo	6	882/11776	0.94 (0.81-1.09)	0.94 (0.83-1.06)	1.23 (0.90-1.68)	0.980	22 (0-69) (0.27)	0.68-1.29	1	0.42	0.35	0.42	0.35	1.5	NP	NP
Antiplatelets	Roberge S 2016	LMWH and low-dose aspirin or aspirin alone	5	54/590	0.54 (0.32-0.92)	0.54 (0.32-0.92)	0.35 (0.14-0.86)	0.649	0 (0-67) (0.68)	0.23-1.28	1	0.88	NP	0.88	NP	1.84	0.66	0.66
Antiplatelets	Henderson JT 2014	Aspirin vs placebo	13	1977/21865	0.78 (0.64-0.95)	0.86 (0.77-0.95)	0.88 (0.75-1.03)	0.002	36 (0-66) (0.09)	0.50-1.21	4	1.96	0.12	2.81	0.50	1.65	0.07	0.07
Antiplatelets	Villa PM 2013	Aspirin ≤16 weeks (abnormal uterine flow)	3	97/346	0.55 (0.36-0.83)	0.56 (0.41-0.77)	0.57 (0.40-0.82)	0.631	16 (0-77) (0.31)	0.02-17.67	2	0.89	0.21	0.93	0.23	0.84	0.19	0.19
Antiplatelets	Dodd JM 2013	Heparin (alone or with other)	7	91/761	0.47 (0.210.1.01)	0.45 (0.28-0.71)	0.35 (0.14-0.86)	0.957	58 (0-80) (0.03)	0.05-4.27	4	2.16	0.21	1.98	0.11	3.18	0.71	0.71
Antiplatelets	Roberge S 2012	Low-dose aspirin ≤16 wks preterm PE	5	45/556	0.11 (0.03-0.33)	0.11 (0.03-0.33)	0.11 (0.01-0.86)	0.850	0 (0-64) (0.72)	0.02-0.68	1	3.75	0.02	3.75	0.02	3.72	0.02	0.02
Antiplatelets	Trivedi NA 2011	Low-dose aspirin in low risk women	5	729/16550	0.87 (0.64-1.17)	0.97 (0.84-1.12)	1.14 (0.94-1.38)	0.170	67 (0-85) (0.02)	0.33-2.29	2	0.28	0.03	0.78	0.17	0.68	0.14	0.14
Antiplatelets	Trivedi NA 2011	Low-dose aspirin in high risk women	14	1365/11687	0.79 (0.65-0.97)	0.85 (0.77-0.94)	0.88 (0.75-1.04)	0.059	50 (0-72) (0.02)	0.47-1.33	3	1.73	0.40	2.62	0.73	1.35	0.15	0.15
Antiplatelets	Bujold E 2009	Aspirin vs placebo in women with AUAD	9	245/1317	0.67 (0.47-0.94)	0.70 (0.56-0.87)	0.95 (0.67-1.35)	0.450	36 (0-69) (0.13)	0.30-1.47	3	1.28	0.12	1.52	0.18	0.47	0.01	0.01
Antiplatelets	Duley L 2007	Antiplatelet agents (moderate risk)	25	1625/28509	0.77 (0.64-0.92)	0.89 (0.81-0.98)	0.88 (0.74-1.03)	0.000	43 (0-64) (0.01)	0.46-1.30	4	2.02	0.14	4.36	NP	2.14	0.16	0.16
Antiplatelets	Duley L 2007	Antiplatelet agents (high risk)	18	748/4121	0.60 (0.45-0.81)	0.78 (0.69-0.89)	0.91 (0.77-1.06)	0.002	42 (0-66) (0.03)	0.29-1.27	5	2.02	0.04	3.65	0.39	1.09	0.00	0.00
Antiplatelets	Duley L 2007	Antiplatelet agents vs placebo	43	2373/32590	0.72 (0.62-0.83)	0.85 (0.79-0.92)	0.88 (0.75-1.03)	0.000	43 (12-60) (<0.01)	0.43-1.19	9	4.13	0.02	8.02	0.70	3.39	0.01	0.01
Other drugs	Chen B 2015	Fish oil vs control (low risk)	7	155/3720	0.82 (0.53-1.26)	0.84 (0.62-1.15)	0.87 (0.60-1.25)	0.617	13 (0-64) (0.33)	0.36-1.85	0	0.50	NP	0.56	NP	0.45	NP	NP
Other drugs	Chen B 2015	Fish oil vs control (high risk)	5	129/1965	1.04 (0.72-1.50)	1.02 (0.73-1.43)	0.96 (0.53-1.76)	0.226	15(0-69) (0.32)	0.47-2.30	0	0.25	NP	0.26	NP	0.26	NP	NP
Other drugs	Chen B 2015	Fish oil vs control	12	413/7650	0.93 (0.72-1.21)	0.92 (0.73-1.16)	0.87 (0.60-1.25)	0.979	10 (0-55) (0.34)	0.60-1.45	0	0.66	NP	0.65	NP	0.78	NP	NP
Other drugs	Makrides M 2014	Magnesium supplementation	3	78/1042	0.88 (0.58-1.34)	0.88 (0.58-1.34)	1.04 (0.15-7.35)	0.613	0 (0-73) (0.60)	0.06-12.83	0	0.19	NP	0.19	NP	0.15	NP	NP
Other drugs	Gui S 2014	L-arginine vs placebo	2	125/524	0.38 (0.25-0.60)	0.38 (0.25-0.58)	0.34 (0.21-0.55)	NA	4 (NA) (0.31)	NA	1	1.59	0.36	1.59	0.37	1.69	0.28	0.28
Other drugs	Rumbold A 2008	Antioxidants versus control	9	586/5446	0.72 (0.49-1.04)	0.91 (0.78-1.06)	0.97 (0.80-1.17)	0.057	58 (0-78) (0.02)	0.27-1.90	3	0.62	0.02	2.08	0.44	0.47	0.01	0.01
Other drugs	Meher S 2007	Nitric oxide vs placebo/control	4	42/170	0.78 (0.37-1.66)	0.87 (0.50-1.51)	1.35 (0.61-3.01)	0.341	37 (0-79) (0.19)	0.06-10.48	0	0.22	NP	0.26	NP	0.30	NP	NP
Other drugs	Imhoff-Kunsch 2012	n-3 LCPUFA supplementation	4	93/1683	0.80 (0.44-1.46)	0.91 (0.61-1.36)	1.15 (0.66-1.99)	0.027	38 (0-79) (0.18)	0.10-6.29	0	0.23	NP	0.35	NP	0.26	NP	NP
Other drugs	Allen R 2014	Essential fatty acids supplementation	6	226/4579	0.88 (0.63-1.24)	0.94 (0.73-1.21)	1.04 (0.73-1.48)	0.007	24 (0-70) (0.25)	0.42-1.86	0	0.33	NP	0.42	NP	0.31	NP	NP
Vitamins	Rumbold A 2015	Any vitamin E supplementation vs pl	14	1965/20878	0.91 (0.79-1.06)	0.98 (0.90-1.06)	1.07 (0.93-1.24)	0.004	47 (0-70) (0.02)	0.62-1.35	3	0.73	0.03	1.25	0.12	0.98	0.07	0.07
Vitamins	Rumbold A 2015	Vitamin C alone or with other suppl vs pl	16	2003/21956	0.92 (0.80-1.05)	0.98 (0.90-1.06)	1.07 (0.93-1.24)	0.010	41 (0-66) (0.04)	0.64-1.31	3	0.83	0.05	1.29	0.13	1.09	0.09	0.09
Vitamins	De-Regil LM 2015	Vitamin D alone versus no treatment	2	25/219	0.52 (0.25-1.07)	0.52 (0.25-1.07)	0.53 (0.25-1.10)	NA	0 (NA) (0.79)	NA	0	0.44	NP	0.44	NP	0.42	NP	NP
Vitamins	De-Regil LM 2015	Vitamin D + calcium vs no treatment	3	78/1114	0.50 (0.32-0.80)	0.50 (0.32-0.80)	0.39 (0.21-0.73)	0.658	0 (0-73) (0.47)	0.03-9.95	1	1.28	NP	1.28	NP	1.81	0.57	0.57
Vitamins	Pérez-López FR 2015	Vitamin D intervention vs pl	3	47/654	0.92 (0.45-1.87)	0.89 (0.51-1.56)	0.67 (0.33-1.35)	0.943	22 (0-78) (0.28)	0.00-400.2	0	0.17	NP	0.16	NP	0.40	NP	NP
Vitamins	Conde-Agudelo A 2011	Supplementation & vitamins C and E	9	1903/19810	0.99 (0.90-1.09)	1.00 (0.92-1.09)	1.07 (0.93-1.24)	0.113	13 (0-60) (0.32)	0.83-1.18	1	0.45	0.37	0.45	0.37	0.72	0.53	0.53
Lifestyle	Zheng J 2017	Exercise	2	35/1009	1.05 (0.53-2.08)	1.05 (0.53-2.08)	1.00 (0.49-2.03)	NA	0 (NA) (0.62)	NA	0	0.10	NP	0.10	NP	0.10	NP	NP
Lifestyle	Muktabhant B 2015	Diet and exercise counselling	8	177/3139	0.99 (0.74-1.31)	0.99 (0.74-1.31)	1.05 (0.73-1.51)	0.018	0 (0-56) (0.95)	0.69-1.41	0	0.40	NP	0.40	NP	0.41	NP	NP
Lifestyle	Muktabhant B 2015	Supervised exercise	3	47/1024	0.91 (0.52-1.60)	0.91 (0.52-1.60)	1.00 (0.51-1.97)	0.463	0 (0-73) (0.76)	0.02-34.2	0	0.16	NP	0.16	NP	0.15	NP	NP
Lifestyle	Muktabhant B 2015	Unsupervised exercise	2	8/229	1.60 (0.38-6.70)	1.60 (0.38-6.70)	1.34 (0.27-6.72)	NA	0 (NA) (0.63)	NA	0	0.16	NP	0.16	NP	0.12	NP	NP
Lifestyle	Muktabhant B 2015	Diet counselling/other	4	54/634	0.90 (0.54-1.47)	0.90 (0.54-1.47)	2.69 (0.55-13.0)	0.878	0 (0-68) (0.44)	0.30-2.67	0	0.22	NP	0.22	NP	1.71	0.14	0.14
Lifestyle	Muktabhant B 2015	All diet and/or exercise vs standard	18	336/5280	0.95 (0.77-1.16)	0.95 (0.77-1.16)	1.05 (0.73-1.51)	0.337	0 (0-44) (0.99)	0.76-1.18	0	0.93	NP	0.93	NP	0.93	NP	NP
Lifestyle	Allen R, 2014	Diet and nutrition counseling	6	249/2695	0.68 (0.54-0.86)	0.68 (0.54-0.86)	0.65 (0.48-0.88)	0.699	0 (0-61) (0.61)	0.49-0.95	2	1.40	0.63	1.40	0.63	1.62	0.66	0.66
Lifestyle	Allen R, 2014	Mixed interventions	6	113/1438	0.92 (0.64-1.31)	0.92 (0.64-1.31)	1.00 (0.55-1.79)	0.691	0 (0-61) (0.59)	0.55-1.51	0	0.33	NP	0.42	NP	0.30	NP	NP
Lifestyle	Allen R, 2014	All type of interventions	18	588/8712	0.81 (0.69-0.96)	0.81 (0.69-0.95)	1.04 (0.73-1.48)	0.583	4 (0-46) (0.41)	0.64-1.03	2	1.73	0.69	1.71	0.69	0.93	0.24	0.24
Lifestyle	Thangaratinam S 2012	Dietary interventions	6	249/2624	0.68 (0.54-0.86)	0.68 (0.54-0.86)	0.65 (0.48-0.88)	0.788	0 (0-61) (0.55)	0.48-0.95	2	1.41	0.63	1.41	0.63	1.62	0.66	0.66
Lifestyle	Thangaratinam S 2012	Mixed approach	3	16/369	1.40 (0.49-3.95)	1.40 (0.49-3.95)	2.69 (0.55-13.03)	0.948	0 (0-73) (0.37)	0.00-1179	0	0.21	NP	0.21	NP	0.68	NP	NP
Lifestyle	Thangaratinam S 2012	All interventions (diet, mixed, PA)	10	272/3072	0.78 (0.56-1.09)	0.72 (0.57-0.91)	0.65 (0.48-0.88)	0.125	22 (0-62) (0.24)	0.40-1.52	2	1.40	0.64	1.02	0.27	1.96	NP	NP

Calcium	An LB 2015	Calcium supplementation vs placebo	4	754/1452	0.86 (0.69-1.05)	0.89 (0.77-1.02)	0.92 (0.75-1.13)	0.045	37 (0-79) (0.19)	0.42-1.74	1	0.56	0.45	0.87	NP	0.38	0.33	0.33
Calcium	Tang R, 2015	Calcium supplementation vs placebo	10	1513/2478	0.62 (0.48-0.81)	0.86 (0.78-0.95)	0.94 (0.81-1.09)	0.000	72 (39-84) (<0.01)	0.29-1.32	6	1.64	0.00	4.66	0.53	0.69	0.00	0.00
Calcium	Tang R, 2015	Low baseline calcium	6	494/1053	0.42 (0.23-0.76)	0.76 (0.63-0.90)	0.92 (0.75-1.12)	0.002	77 (35-88) (<0.01)	0.06-2.75	4	1.42	0.03	4.07	NP	0.42	0.00	0.00
Calcium	Tang R, 2015	High baseline calcium	2	359/5045	0.70 (0.34-1.44)	0.89 (0.72-1.09)	0.94 (0.76-1.16)	NA	74 (NA) (0.05)	NA	1	0.28	0.26	1.13	NP	0.15	0.14	0.14
Calcium	Tang R, 2015	Unknown baseline calcium	2	660/9208	0.47 (0.08-2.84)	0.93 (0.80-1.08)	0.94 (0.81-1.09)	NA	72 (NA) (0.06)	NA	1	0.22	0.21	1.16	NP	0.18	0.18	0.18
Calcium	Tang R, 2015	Calcium supplements vs pl (High risk of PE)	4	410/8665	0.36 (0.14-0.98)	0.82 (0.68-1.00)	0.92 (0.75-1.12)	0.021	79 (12-90) (<0.01)	0.01-25.6	3	0.71	0.02	2.53	NP	0.30	0.00	0.00
Calcium	Tang R, 2015	Calcium supplements vs pl (Normal risk of PE)	6	1103/16122	0.67 (0.48-0.92)	0.87 (0.78-0.98)	0.94 (0.81-1.09)	0.004	72 (8-86) (<0.01)	0.27-1.67	3	0.97	0.06	2.89	NP	0.44	0.01	0.01
Calcium	Hofmeyr GJ, 2014	Calcium supplements vs pl (Adequate calcium diet)	4	366/5022	0.61 (0.32-1.19)	0.87 (0.72-1.06)	0.94 (0.77-1.16)	0.097	51 (0-82) (0.10)	0.05-6.85	1	0.42	0.36	1.47	NP	0.25	0.22	0.22
Calcium	Hofmeyr GJ, 2014	Calcium supplements vs pl (Low calcium diet)	8	515/10678	0.35 (0.20-0.64)	0.72 (0.61-0.87)	0.92 (0.75-1.13)	0.000	76 (44-87) (<0.01)	0.06-2.12	5	1.79	0.02	5.54	0.71	0.52	0.00	0.00
Calcium	Hofmeyr GJ, 2014	Routine high-dose calcium supplements	13	889/15730	0.44(0.31-0.64)	0.78 (0.69-0.89)	0.92 (0.75-1.13)	0.000	70 (41-82) (<0.01)	0.15-1.28	7	2.30	0.00	6.54	NP	0.86	0.00	0.00
Calcium	Patrelli TS 2012	Adequate calcium intake vs placebo	6	700/9641	0.78 (0.58-1.06)	0.90 (0.78-1.03)	0.94 (0.76.1.16)	0.007	49 (0-78) (0.08)	0.37-1.65	2	0.59	0.11	1.58	0.66	0.39	0.05	0.05
Calcium	Patrelli TS 2012	Low calcium intake vs placebo	7	474/10154	0.35 (0.18-0.68)	0.77 (0.64-0.92)	0.92 (0.75-1.13)	0.000	75 (34-87) (<0.01)	0.05-2.59	4	1.33	0.03	4.63	0.70	0.46	0.00	0.00
Calcium	Patrelli TS 2012	Calcium supplements vs pl (high risk)	3	41/346	0.17 (0.07-0.42)	0.17 (0.07-0.42)	0.21 (0.07-0.58)	0.095	0 (0-73) (0.80)	0.00-50.1	2	2.31	0.54	2.31	0.54	2.11	NP	NP
Calcium	Patrelli TS 2012	Calcium supplements vs pl (low risk)	7	515/11059	0.51 (0.30-0.86)	0.77 (0.64-0.91)	0.92 (0.75-1.13)	0.059	75 (32-86) (<0.01)	0.10-2.55	4	1.45	0.04	3.60	NP	0.47	0.00	0.00

Abbreviations: Random effects, summary risk ratio (95% CI) using random effects model; Fixed effects, summary risk ratio (95% CI) using fixed effects model; Largest effect, risk ratio (95% CI) of the largest study in the meta-analysis; Egger, p-value from Egger's regression asymmetry test for evaluation of publication bias; P, p-value; O, observed number of "positive" studies; E, expected number of "positive" studies; NP, not pertinent, because the estimated E is larger than the O, thus there is no evidence of excess statistical significance based on the assumption made for the plausible effect size; PA, physical activity; pl, placebo; NP, not pertinent, because the estimated is larger than the observed, and there is no evidence of excess of statistical significance based on the assumption made for the plausible effect size; PE, Preeclampsia; LMWH, low-molecular-weight heparin; AUAD, abnormal uterine artery Doppler.

± Number of studies.
* Summary random effects risk ratio (95% CI) of each meta-analysis
† Summary fixed effects risk ratio (95% CI) of each meta-analysis
‡ Risk ratio (95% CI) of the largest study in each meta-analysis
§ P-value from the Egger regression asymmetry test for evaluation of publication bias
|| I^2 metric of inconsistency (95% confidence intervals of I^2) and P-value of the Cochran Q test for evaluation of heterogeneity
≠ 95% Prediction Interval
¶ Observed number of statistically significant studies
# Expected number of statistically significant studies using the summary fixed effects estimate of each meta-analysis as the plausible effect size
** P-value of the excess statistical significance test
¥ Expected number of statistically significant studies using the summary random effects estimate of each meta-analysis as the plausible effect size
ϑ Expected number of statistically significant studies using the effect of the largest study of each meta-analysis as the plausible effect size

Chapter 6 – Risk factors for gestational diabetes: An umbrella review of meta-analyses of observational studies

6.1 Abstract

Background: Gestational diabetes mellitus (GDM) is a common pregnancy complication, defined as glucose intolerance with onset or first recognition during pregnancy, in women without diabetes history during pregnancy. The etiology of GDM is multifactorial and has not completely been established yet. GDM is a major cause for prenatal morbidity and affects approximately 15% of all pregnancies, depending on population characteristics and diagnostic criteria used. GDM is also considered to be a risk factor for long-term complications such as type 2 diabetes mellitus and cardiovascular disease. Early detection of the risk of developing GDM would be vital for its prevention and the long-term consequences.

Objectives: An umbrella review was performed to summarize evidence on the risk factors associated with GDM, evaluate whether there are hints of biases in this literature and how they manifest and finally identify which of the previously studied associations include convincing evidence to support their results.

Methods: We searched PubMed and ISI Web of Science from inception to July 2017, to identify meta-analyses of observational studies examining associations between risk factors for GDM. For each meta-analysis we estimated the summary effect size by random-effects and fixed-effects models, the 95% confidence interval, the 95% prediction interval, the between-study heterogeneity expressed by I^2 (considering above 75% as very large), evidence of small-study effects and evidence of excess significance bias.

Results: Twenty-one eligible meta-analyses were identified, providing data on 43 associations based on 480 primary studies covering a very wide range of risk factors: diet

and lifestyle factors, diseases and disorders, infections and a range of biomarkers. Thirty-eight (88%) associations had nominally statistically significant findings at $P<0.05$, while only 14 (32%) were significant at $P<10^{-6}$ under the random-effects model. Eighteen (42%) associations had large or very large heterogeneity. Evidence for small-study effects and excess significance bias was found in three (7%) and four (9%) associations, respectively. Only five risk factors presented convincing evidence for an association with GDM: vitamin D deficiency, low vs. normal BMI (cohort studies), BMI ~30-35 kg/m^2 vs. normal BMI, BMI >35 kg/m^2 vs. normal BMI, and hypothyroidism.

Conclusions: Vitamin D deficiency, low vs. normal BMI, moderately and severely obese vs. normal weight, and hypothyroidism show the strongest consistent evidence. Our findings highlight the importance of patient education on diet and lifestyle modifications as candidate interventions to reduce the risk of GDM.

6.2 Introduction

Gestational diabetes mellitus (GDM) is a common pregnancy complication, defined as glucose intolerance with onset or first recognition during pregnancy, in women without prior diabetes history prior to pregnancy (232,453). During the last 20 years the prevalence of GDM has increased worldwide and it is expected to continue to rise along with the increase in pre-conception obesity and obese pregnant women (454). GDM affects approximately 15% of all pregnancies worldwide, depending on population characteristics, and this prevalence may in fact be higher under the new diagnostic criteria (455,456). GDM is associated with an increased risk of maternal and infant morbidity, including macrosomia, large for gestational age (LGA), cesarean section delivery and preterm birth, but it is also considered to be a risk factor for long-term complications, such as type 2 diabetes mellitus and cardiovascular disease in the mother and the offspring (457–460). The etiology of GDM is multifactorial and has not completely been established yet. Several risk factors may contribute to its onset. Age, being overweight or obese, ethnicity, family history of diabetes, and history of GDM are some of the proposed risk factors for GDM (461–464).

To further expand the identification of risk factors for GDM, in the current study we aimed to conduct an umbrella review of meta-analyses of risk factors for GDM. We applied the methodology of umbrella review, as outlined below, to map all the risk factors that have been associated with GDM. Using a standardized approach, we aimed to assess the credibility of the findings in order to identify which associations are supported by robust epidemiological evidence.

6.3 Methods

This study was performed according to the guidelines for systematic reviews under the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (102).

We conducted an umbrella review, which is a systematic collection and evaluation of multiple systematic reviews and meta-analyses performed on a specific research topic (22). An umbrella review brings together comparisons of a large number of existing systematic reviews and meta-analyses on risk factors into one accessible and usable document (22,25). The methods of performing an umbrella review are standardized and, in this work, we follow the same principles used in previously published umbrella reviews across various fields of research (265–269). We used a ranking system to grade the evidence from meta-analyses of observational studies in terms of the significance of the summary effect, 95% prediction interval, and presence of large heterogeneity, small study effects, and excess significance bias.

Two researchers (KG and SP) independently searched PubMed and ISI Web of Science from inception to July 2017 to identify meta-analyses of observational studies examining associations regarding risk factors for GDM. The search strategy used the keywords (“gestational diabetes” OR “pregnancy diabetes” OR “pregnancy hyperglycemia” OR “3 h abnormal gtt test” OR “insulin during pregnancy” OR “antidiabetics during pregnancy” OR “metformin in pregnancy”) AND (“systematic review” OR “meta-analysis”). All identified publications went through a three-step parallel review of title, abstract, and full text, performed by KG and SP, based on predefined inclusion and exclusion criteria. We also

screened the references of the retrieved articles for possible eligible papers. Any disagreement was resolved with discussion.

We included meta-analyses of observational studies (i.e., cross-sectional, case-control and cohort studies), which investigated risk factors for GDM. Meta-analyses were retained if they included at least three studies in which information was provided per included study on a measure of association, its standard error, the number of cases and the total population. We did not apply any language restrictions in the selection of eligible studies. We included only meta-analyses of epidemiological studies in humans. We excluded studies in which risk factors were used for screening, diagnostic, or prognostic purposes, or meta-analyses that examined GDM as a risk factor for other medical conditions. We also excluded studies on women with pre-existing type II diabetes. We excluded systematic reviews and meta-analyses of genetic risk factors, narrative reviews, letters to the editor, meta-analyses of Randomised Control Trials (RCTs), and systematic reviews without a quantitative synthesis of data. If an article presented separated meta-analyses on other medical diseases including GDM, we only extracted information on the latter. When more than one meta-analysis on the same research question was eligible, the meta-analysis with the largest number of component studies with data on individual studies' effect sizes was retained for the main analysis.

Data extraction

Data extraction was performed independently by two investigators (KG, SP), and in case of discrepancies, the final decision was reached by consensus, involving a third investigator, when necessary (EE). From each eligible meta-analysis, we extracted information on the first

author, year of publication, the examined risk factors, the number of studies included, the study-specific relative risk estimates (risk ratio, odds ratio, or standardized mean differences) along with the corresponding confidence intervals (CI). Also, we recorded the reported summary meta-analytic estimates using both fixed and random effect methods along with the corresponding confidence intervals, the total population, and number of cases for each study. We also recorded whether the selected meta-analyses applied any criteria to evaluate the quality of the included studies.

Statistical analysis

For each unique meta-analysis, we estimated the summary effect and its 95% CI by using both fixed and random effect models (94,270). We also calculated the 95% prediction intervals (PI) for the summary random effects estimates, which further accounts for between-study heterogeneity and indicates the uncertainty for the effect that would be expected in a new study examining the same association (271,294). For the largest study of each meta-analysis, we calculated the standard error (SE) of the effect size and we examined whether the standard error was less than 0.10, indicating that the difference between the effect estimate and the upper or lower 95% confidence interval is less than 0.20 (i.e. this uncertainty is less than what is considered a small effect size), and whether the largest study presented a statistically significant effect.

We assessed heterogeneity among studies and we reported the P value of the χ^2-based Cochran Q test and the I^2 metric for inconsistency, which could reflect either diversity or bias. I^2 ranges between 0% and 100% and quantifies the variability in effect estimates that is

due to heterogeneity rather than sampling error (273). Values exceeding 50% or 75% are usually considered to represent large or very large heterogeneity, respectively. Confidence intervals were calculated as per Ioannidis et al. (274).

Moreover, we assessed whether there is evidence for small study effect (i.e. whether smaller studies tend to give substantially larger estimates of effect size compared with larger studies). Small study effects can indicate publication and other selective reporting biases, but they can also reflect genuine heterogeneity, chance, or other reasons for differences between small and large studies (275). We used the regression asymmetry test proposed by Egger for this assessment (276). A P value <0.10 accompanied by a more conservative effect in larger studies was considered evidence of small-study effects.

We further applied the excess significant test to evaluate whether there is a relative excess of formally significant findings in published literature due to any reason (e.g. publication bias, selective reporting of outcomes or analyses). This is a chi-squared-based test, in which the number of expected positive studies is estimated and compared against the number of observed number of studies with statistically significant results (P<0.05) (68). A binomial test was then used to evaluate whether the number of positive studies in a meta-analysis is too large according to the power that these studies have to detect plausible effects at α=0.05. Briefly, a comparison between observed vs. expected is performed separately for each meta-analysis and it is also extended to groups of many meta-analyses after summing the observed and expected from each meta-analysis. The expected number of significant studies for each meta-analysis is calculated by the sum of the statistical power estimates for each component

study (68). The power of each component study was estimated using the fixed effects summary, the random effects summary, or the effect size of the largest study (smallest SE) as the plausible effect size (72). The power of each study was calculated with an algorithm using a non-central t distribution (278). Excess statistical significance for single meta-analyses was claimed at $P<0.10$ (one-sided $P<0.05$, with observed > expected as previously proposed) (68).

We classified risk factors into categories based on biological pathways or types of exposures involved: biomarkers, diet and lifestyle factors, diseases and disorders, infections, and other factors. We examined excess of statistical significance separately in each of these categories as selective reporting bias may arise in different categories of research. The excess of statistical significance test was also conducted separately for meta-analyses with I^2 values less than or equal to 50% and those with I^2 values greater than 50%, because values above 50% typically reflect evidence of large heterogeneity beyond chance (295).

Assessment of epidemiologic credibility

We characterized as convincing the associations fulfilling the following criteria: they had a significant effect under the random-effects model at $P<10^{-6}$, they were based on evidence from more than 1000 cases, the between-study heterogeneity was not large ($I^2<50\%$), the 95% PI excludes the null value, and there was no evidence of small-study effects or excess of significance bias. Additionally, associations with more than 1000 cases, a significant effect at $P<10^{-6}$, and a nominally statistically significant effect present at the largest study were characterized as highly suggestive. We considered as suggestive the associations with

significant effect at $P<10^{-3}$ and more than 1000 cases. The remaining statistically significant associations at P<0.05 under random-effects model were graded as weak associations. All authors had full access to all the data in the study. Statistical analysis and the power calculations were performed in STATA version 14 (STATA Corp, College Station, TX).

6.4 Results

Description of Eligible Meta-analyses

Overall, the literature search identified 673 publications of which 607 were excluded after the title and abstract review. Of the 66 articles screened in full text, 15 articles did not report the appropriate information for the calculation of excess of statistical significance (either because the total sample size was missing or the study-specific relative risk estimates were missing), 10 articles were excluded because the outcome of interest was not gestational diabetes, 8 because were editorials or narrative reviews, 5 because were meta-analyses of RCTs, 5 articles excluded because a larger systematic review or meta-analysis investigating the same risk factor was available, and 2 articles were excluded because included only 2 component studies (Figure 6.1). The 21 eligible papers (339,343,465–483) included data on 43 different meta-analyses (comparisons) in five broad areas (biomarkers [n=20 comparisons], diet and lifestyle [n=13 comparisons], diseases and disorders [n=5 comparisons], infections [n=2 comparisons], and other factors [n=3 comparisons]). There were 3 to 40 studies per meta-analysis, with a median of 7 studies. The publication date of the eligible articles ranged between 2009 and 2017. The median number of case and control participants in each study was 95 and 106, respectively. The median number of case and

control subjects in each meta-analysis was 1596 and 5574, respectively. The number of cases was greater than 1000 in 23 meta-analyses (Table 6.1).

Six articles (29%) used the Newcastle Ottawa Scale (NOS) to qualitatively assess the included primary studies. Three articles (14%) used the Cochrane Collaboration's risk of bias tool, two articles (10%) used assessment criteria for non-randomized observational studies adapted from Duckitt & Harrington, two (10%) articles used the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) Statement, and three (14%) articles used other assessment tools. Five papers (25%) did not perform any quality assessment. Supplementary Table 6.4 summarizes these 43 meta-analyses, which included 480 individual study estimates.

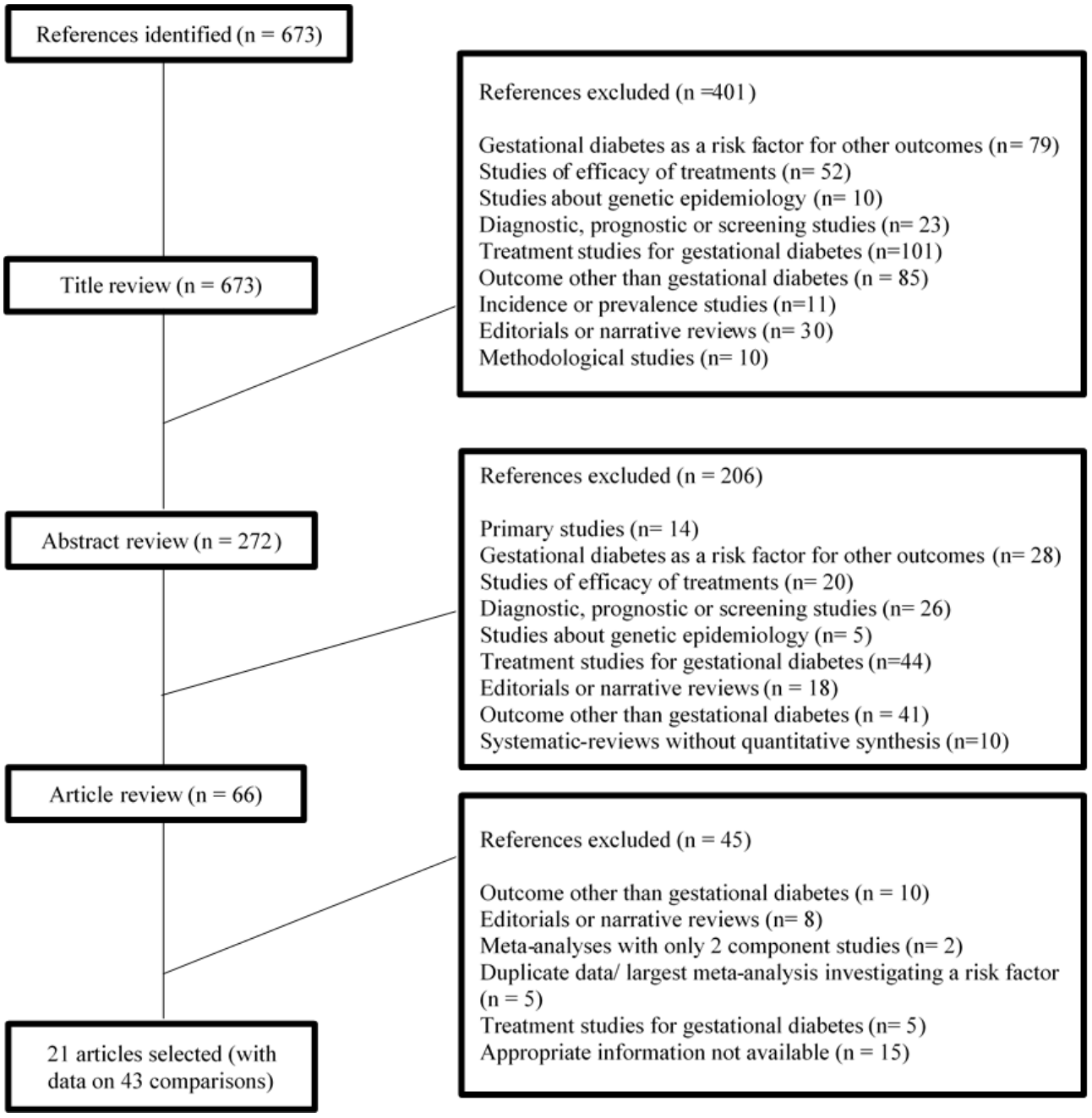

Figure 6.1. Flowchart of the included studies

Table 6.1. Quantitative synthesis and assessment of bias across the 43 associations of risk factors for gestational diabetes

Area	Author, year	Comparison	Studies	Cases/controls	Random effects*	Largest effect‡	P Random	Egger§	I^2 (P)‖	95% PI ≠
Biomarkers	Kong FJ 2017	Betatrophin levels	8	401/421	6.65 (2.12-20.9)	16.5 (9.18-29.8)	1.17×10^{-3}	0.191	94 (<0.001)	0.11-411.7
Biomarkers	Fu S 2016	Ferritin (highest vs lowest ferritin levels) (cohorts)	4	214/1662	3.22 (1.73-6.00)	4.98 (1.46-17.03)	2.37×10^{-4}	0.953	0 (0.815)	0.82-12.65
Biomarkers	Fu S 2016	Dietary total iron intake	3	1007/13850	1.01 (1.00-1.01)	1.12 (0.87-1.45)	2.78×10^{-8}	NA	0 (0.73)	0.99-1.03
Biomarkers	Fu S 2016	Serum ferritin (GMD-women vs non-GMD)	6	403/498	4.89 (2.06-11.58)	6.45 (4.07-10.24)	3.10×10^{-4}	0.756	91 (<0.001)	0.22-106.6
Biomarkers	Fernández-Cao JC 2016	Hemoglobin levels	9	792/4393	1.54 (1.18-2.03)	0.81 (0.36-1.82)	1.80×10^{-3}	0.752	33 (0.157)	0.81-2.93
Biomarkers	Fernández-Cao JC 2016	Ferritin (highest vs lowest ferritin levels) (mixed)	7	330/5574	2.09 (1.48-2.96)	2.27 (1.20-4.30)	3.27×10^{-5}	0.600	1 (0.42)	1.31-3.34
Biomarkers	Kong FJ 2016	Selenium level	7	178/391	0.12 (0.03-0.53)	0.12 (0.06-0.26)	5.00×10^{-3}	0.499	93 (<0.001)	0.00-19.81
Biomarkers	Hu S 2016	Serum retinol-binding protein-4	17	647/620	4.38 (2.10-9.14)	1.27 (0.70-2.30)	8.47×10^{-5}	0.025	91 (<0.001)	0.18-106.7
Biomarkers	Iliodromiti S 2016	Adiponectin	11	794/2071	6.35 (4.08-9.88)	5.05 (3.55-7.18)	2.44×10^{-16}	0.770	71 (<0.001)	1.56-25.9
Biomarkers	Guo CC 2016	DQ2	12	2333/2687	1.36 (1.10-1.66)	0.96 (0.79-1.16)	3.65×10^{-3}	0.008	43 (0.06)	0.80-2.30
Biomarkers	Guo CC 2016	DQ6	11	2270/2576	0.81 (0.69-0.94)	0.75 (0.55-1.02)	7.56×10^{-3}	0.551	0 (0.743)	0.67-0.97
Biomarkers	Guo CC 2016	DR 13	4	209/225	2.46 (1.02-5.90)	0.73 (0.29-1.87)	.04437	0.982	67 (0.03)	0.07-88.5
Biomarkers	Guo CC 2016	DR17	5	329/335	3.16 (1.31-7.64)	3.13 (1.11-8.81)	.01054	0.116	69 (0.01)	0.16-62.9
Biomarkers	Yang Y 2015	Thyroid antibodies (cohort)	11	1596/30012	1.07 (0.97-1.19)	1.18 (0.77-1.81)	.19124	0.546	0 (0.44)	0.95-1.21
Biomarkers	Yang Y 2015	Thyroid antibodies (case-control)	10	856/2062	1.21 (1.05-1.41)	1.33 (1.09-1.63)	.01042	0.402	0 (0.73)	1.02-1.44
Biomarkers	Yang Y 2015	Thyroid antibodies (All studies)	21	2452/32074	1.12 (1.03-1.22)	1.18 (0.77-1.81)	.01065	0.485	0 (0.60)	1.02-1.23
Biomarkers	Zhang MX 2015	Vitamin D deficiency	20	1737/7472	1.55 (1.32-1.82)	1.38 (1.05-1.82)	1.04×10^{-7}	0.110	16 (0.25)	1.10-2.19
Biomarkers	Aghajafari F 2013	25(OH)D concentration	10	687/3425	1.49 (1.18-1.88)	1.35 (0.77-2.35)	6.74×10^{-4}	0.580	0 (0.58)	1.14-1.96
Biomarkers	Wei SQ 2013	25(OH)D5<50 nmol/l	10	623/3503	1.37 (1.11-1.70)	1.20 (0.72-2.00)	3.18×10^{-3}	0.147	0 (0.51)	1.07-1.76
Biomarkers	Wei SQ 2013	25(OH)D<75 nmol/l	8	542/3298	1.52 (1.17-1.98)	1.63 (0.79-3.33)	1.64×10^{-3}	0.954	7 (0.37)	1.01-2.30
Diet and lifestyle	Aune D 2016	Leisure-time physical activity before pregnancy	8	2401/30191	0.78 (0.61-1.00)	0.81 (0.68-1.01)	.05027	0.869	47 (0.07)	0.41-1.47
Diet and lifestyle	Aune D 2016	Leisure-time physical activity during pregnancy	5	580/5140	0.97 (0.73-1.28)	0.91 (0.37-2.21)	.81601	0.430	0 (0.80)	0.61-1.52
Diet and lifestyle	Torloni MR 2009	Low vs. Normal BMI (cohort)	16	75669/280734	0.75 (0.69-0.83)	0.80 (0.69-0.92)	1.55×10^{-9}	0.022	16 (0.27)	0.63-0.90
Diet and lifestyle	Torloni MR 2009	Low vs. Normal BMI (case-control)	3	5957/11651	0.65 (0.51-0.83)	0.61 (0.47-0.81)	4.47×10^{-4}	0.572	0 (0.83)	0.13-3.16
Diet and lifestyle	Torloni MR 2009	Overweight vs. Normal BMI (cohort)	17	112880/282458	1.97 (1.76-2.19)	2.29 (2.12-2.47)	8.01×10^{-35}	0.521	56 (0.003)	1.44-2.68
Diet and lifestyle	Torloni MR 2009	Overweight vs. Normal BMI (case-control)	3	287/501	2.68 (1.78-4.04)	3.85 (2.30-6.47)	2.33×10^{-6}	0.889	40 (0.19)	0.05-138
Diet and lifestyle	Torloni MR 2009	Obese (BMI >30) vs. normal weight	31	56333/308335	3.76 (3.31-4.28)	4.80 (4.43-5.21)	0	0.661	73 (<0.001)	2.23-6.34
Diet and lifestyle	Torloni MR 2009	Obese 1 (BMI ~30–35) vs. Normal weight	6	3087/20901	3.01 (2.34-3.86)	3.21 (2.68-3.85)	8.88×10^{-18}	0.612	27 (0.23)	1.71-5.28
Diet and lifestyle	Torloni MR 2009	Obese 2 (BMI >35) vs. Normal weight	7	1747/21001	5.52 (4.28-7.11)	5.10 (3.18-8.19)	0	0.157	7 (0.37)	3.62-8.42
Diet and lifestyle	Torloni MR 2009	Overweight vs. Non-overweight (cohort)	34	174233/391991	2.95 (2.68-3.24)	3.10 (2.91-3.31)	0	0.132	72 (<0.001)	1.97-4.41
Diet and lifestyle	Torloni MR 2009	Overweight vs. Non-overweight (case-control)	10	6214/19567	3.78 (2.49-5.76)	3.06 (2.51-3.73)	5.18×10^{-10}	0.248	90 (<0.001)	0.83-17.2
Diet and lifestyle	Torloni MR 2009	Obese vs. non-obese women (cohort)	40	68013/520879	3.36 (3.01-3.74)	3.44 (3.20-3.70)	0	0.724	77 (<0.001)	1.97-5.72
Diet and lifestyle	Torloni MR 2009	Obese vs. non-obese women (case-control)	3	238/922	3.24 (1.28-8.19)	7.49 (4.58-12.3)	.01289	0.938	88 (0.001)	0-285401
Diseases/disorders	Gong LL 2016	Overt hypothyroidism	3	3444/222161	2.44 (1.08-5.52)	1.88 (1.67-2.12)	.03262	0.688	57 (0.10)	0-15039
Diseases/disorders	Gong LL 2016	Subclinical hypothyroidism	6	1859/61708	1.59 (1.32-1.92)	1.49 (1.04-2.13)	1.29×10^{-6}	0.208	0 (0.50)	1.22-2.07
Diseases/disorders	Gong LL 2016	Hypothyroidism (all)	7	5770/278609	1.72 (1.51-1.95)	1.88 (1.67-2.12)	4.21×10^{-17}	0.137	14 (0.32)	1.35-2.18
Diseases/disorders	Luque-Fernandez 2013	Sleep-disordered breathing	9	673/9122	2.18 (1.59-2.98)	1.44 (0.99-2.10)	1.22×10^{-6}	0.011	52 (0.03)	0.95-4.97
Diseases/disorders	Kjerulff LE 2011	Polycystic ovary syndrome	18	2385/89669	2.83 (1.95-4.10)	2.69 (2.33-3.11)	4.63×10^{-8}	0.653	52 (0.005)	0.94-8.46
Infections	Abariga SA 2016	Periodontitis	10	624/5100	1.66 (1.16-2.36)	1.73 (0.91-3.30)	5.18×10^{-3}	0.008	51 (0.03)	0.61-4.49
Infections	Soepnel LM 2016	HIV infection	4	593/1070	0.83 (0.48-1.42)	1.00 (0.37-2.71)	.49148	0.472	0 (0.61)	0.25-2.71

Other	Moosazadeh M 2016	Family history of diabetes	33	2697/29134	3.46 (2.80-4.27)	4.36 (2.89-6.58)	5.41×10^{-31}	0.861	76 (<0.001)	1.17-10.2
Other	Xu Y 2016	Isolated Single Umbilical Artery	7	1880/490712	1.38 (1.06-1.80)	2.08 (1.47-2.96)	.01842	0.569	35 (0.16)	0.73-2.61
Other	Pandey S 2012	IVF/ICSI versus spontaneous conception	6	13399/574391	1.31 (0.98-1.75)	1.55 (1.37-1.75)	.07039	0.169	42 (0.13)	0.63-2.72

Abbreviations: Random effects, summary odds ratio (95% CI) using random effects model; Largest effect, odds ratio (95% CI) of the largest study in the meta-analysis; Egger, p-value from Egger's regression asymmetry test for evaluation of publication bias; P, p-value; NP, not pertinent, because the estimated is larger than the observed, and there is no evidence of excess of statistical significance based on the assumption made for the plausible effect size; BMI, Body Mass Index; GDM, gestational diabetes mellitus; PA, physical activity.

* Summary random effects odds ratio (95% CI) of each meta-analysis, except for three meta-analyses (Fu S 2016, Aune D 2016 and Pandey S 2012) where the RR was used.
‡ Odds ratio (95% CI) of the largest study in each meta-analysis, except for three meta-analyses (Fu S 2016, Aune D 2016 and Pandey S 2012) where the RR was used.
§ P-value from the Egger regression asymmetry test for evaluation of publication bias
|| I^2 metric of inconsistency and P-value of the Cochran Q test for evaluation of heterogeneity
≠ 95% Prediction Interval

Summary Effect Sizes and Significant Findings

Of the 43 meta-analyses, 38 (88%) had nominally statistically significant findings at P<0.05 using the random effects model. Out of these, a total of 23 (53%) associations presented statistically significant effects at P<0.001, while only 14 (33%) remained significant after the application of the more stringent p-value threshold of $P<10^{-6}$ (Table 6.1). The fourteen risk factors that presented a significant effect for an association with GDM at $P<10^{-6}$ were the following: dietary total iron intake, adiponectin, vitamin D deficiency, low vs. normal BMI (cohort studies), overweight vs. normal BMI (cohort studies), BMI >30 vs. normal weight, BMI ~30–35 vs. normal weight, BMI >35 vs. normal weight, overweight vs. non-overweight women (cohort studies), overweight vs. non-overweight (case-control), obese vs. non-obese women (cohort studies), hypothyroidism, polycystic ovary syndrome, and family history of diabetes. Additional information on all 43 meta-analyses is available online (Supplementary Table 6.4).

Across the five areas of risk factors there were differences in the proportion of associations that had nominally statistically significant summary effects. Based on the random effects calculations at P<0.05, the proportion of studies with nominally statistically significant summary effects was: 100% for diseases and disorders, 95% for biomarkers, and 85% for diet and lifestyle. On the contrary, this was seen only in 66% and 50% of the meta-analyses on other risk factors and infections, respectively.

Between-Study Heterogeneity and Prediction Intervals

Ten (23%) meta-analyses had large heterogeneity estimates ($I^2 \geq 50\%$ and $I^2 \leq 75\%$) and 8 (19%) meta-analyses had very large heterogeneity estimates ($I^2 > 75\%$) (Table

6.1). When we calculated the 95% prediction intervals, in 19 (44%) meta-analyses the null value was excluded. This included nine biomarkers [ferritin levels, adiponectin, DQ6, thyroid antibodies (case-control studies), thyroid antibodies (all studies), vitamin D deficiency, 25(OH)D concentration, 25(OH)D5 <50 nmol/l, 25(OH)D <75 nmol/l], seven diet and lifestyle factors [low vs. normal BMI (cohort studies), overweight vs. normal BMI (cohort studies), BMI >30 vs. normal weight, BMI ~30–35 vs. normal weight, BMI >35 vs. normal weight, overweight vs. non-overweight women (cohort studies), obese vs. non-obese women (cohort studies)], two diseases and disorders (subclinical hypothyroidism, hypothyroidism), and one other risk factor (family history of diabetes) (Table 6.1).

Small-Study Effects and Excess Significance Bias

Evidence for statistically significant small-study effects (Egger test P<0.10 and random effects summary estimate larger compared to the point estimate of the largest study in the meta-analysis) was identified in 3 of 43 (7%) meta-analyses (Supplementary Table 6.4). These included two meta-analyses on biomarkers (serum retinol-binding protein-4, DQ2), and one on diseases and disorders (sleep-disordered breathing). Four (9%) associations had hints of excess statistical significance bias with statistically significant (P<0.05) excess of positive studies under any of the three assumptions for the plausible effect size - the fixed effects summary, the random effects summary or the results of the largest study (Supplementary Table 6.4). Two (3%) of them pertained to biomarkers, one (1%) pertained to diseases and disorders, and one (1%) pertained to other risk factors. Table 6.2 shows the results of excess of statistical significance bias according to category of risk factor.

Table 6.2. Observed and expected number of positive studies by type of risk factor*

Area	No. of studies	Observed positive	Expected positive (fixed) †	P‡ (fixed)	Expected positive (random)§	P‡ (random)	Expected positive (largest)‖	P‡ (largest)	Expected positive (composite) ¶	P‡ (composite)
All	480	268	328	0.00	338	0.00	301	0.00	301	0.00
Biomarkers	194	73	83.88	0.13	91.75	0.01	71.7	0.88	71.7	0.88
Diet and lifestyle	183	140	165.7	0.00	166	0.00	165.2	0.00	165.2	0.00
Diseases & disorders	42	25	32.85	0.01	34.8	0.00	31	0.05	31	0.05
Infections	15	3	4.16	0.77	5.50	0.28	5.50	0.28	4.16	0.77
Other	46	27	41.41	0.00	39.9	0.00	27.47	0.88	27.47	0.88

* NP, not pertinent, because the estimated is larger than the observed, and there is no evidence of excess of statistical significance based on the assumption made for the plausible effect size.

† Expected number of statistically significant studies using the summary fixed effects estimate of each meta-analysis as the plausible effect size.

‡ P value of the excess of statistically significant test. All statistical tests were two-sided.

§ Expected number of statistically significant studies using the summary random effects estimate of each meta-analysis as the plausible effect size.

‖ Expected number of statistically significant studies using the effect of the largest study of each meta-analysis as the plausible effect size.

¶ Expected number of statistically significant studies using the most conservative of the three estimates (fixed effects summary, random effects summary, largest study) of each meta-analysis as the plausible effect size.

Risk factors with Strong Evidence of Association

After applying our credibility criteria, five risk factors, vitamin D deficiency, low vs. normal BMI (cohort studies), BMI ~30–35 vs. normal weight, BMI >35 vs. normal weight, and hypothyroidism (all types) presented convincing evidence for an association with GDM, supported by more than 1000 cases, $P<10^{-6}$ under the random effect model, no hints for small-study effects and for excess statistical significance, not large heterogeneity ($I^2<50\%$), and a 95% PI excluding the null value. Seven risk factors [overweight vs. normal BMI (cohort), BMI >30 vs. normal weight, overweight vs. non-overweight women (cohort), overweight vs. non-overweight (case-control), obese vs. non-obese women (cohort), polycystic ovary syndrome, family history of diabetes] presented highly suggestive evidence for GDM. Three risk factors were supported by suggestive evidence and twenty-three associations presented weak evidence ($P<0.05$). An overall assessment of statistically significant associations for GDM is presented in Table 6.3.

Table 6.3. Assessment across the statistically significant associations for gestational diabetes

Level of evidence	Criteria used	Decreased risk	Increased risk
Convincing	>1000 cases, [a] $P<10^{-6}$, not large heterogeneity (I^2 <50%), 95% prediction interval excluding the null value, no evidence for small-study effects [b] and excess significance bias [c]	Low vs. Normal BMI (cohort)	Vitamin D deficiency, BMI ~30–35 vs. Normal weight, BMI >35 vs. Normal weight, Hypothyroidism (all)
Highly suggestive	>1000 cases, [a] $P<10^{-6}$ and nominally statistically significant effect present at the largest study		Overweight vs. Normal BMI (cohort), BMI >30 vs. normal weight, Overweight vs. Non-overweight women (cohort), Overweight vs. Non-overweight (case-control), Obese vs. non-obese women (cohort), Polycystic ovary syndrome, Family history of diabetes
Suggestive	>1000 cases, [a] $P<10^{-3}$	Low vs. Normal BMI (case-control)	Dietary total iron intake, Subclinical hypothyroidism
Weak	The rest associations with [a] P < 0.05	Selenium level, DQ6	Betatrophin levels, Ferritin (highest vs lowest ferritin levels) (cohorts), Serum ferritin (GMT-women vs non-GMD), Hemoglobin levels, Ferritin (highest vs lowest ferritin levels) (mixed), Serum retinol-binding protein-4, Adiponectin, DQ2, DR13, DR17, Thyroid antibodies (case-control), Thyroid antibodies (All studies), 25(OH)D concentration, 25(OH)D5 <50 nmol/l, 25(OH)D <75 nmol/l, Overweight vs. Normal BMI (case-control), Obese vs. non-obese women (case-control), Overt hypothyroidism, Sleep-disordered breathing, Periodontitis, Isolated Single Umbilical Artery

Abbreviations: BMI, Body Mass Index; GDM, gestational diabetes mellitus.

[a] P indicates the P-values of the meta-analysis random effects model.

[b] Small study effect is based on the P-value from the Egger's regression asymmetry test (P< 0.10).

[c] Based on the P-value (P<0.05) of the excess significance test using the largest study (smallest standard error) in a meta-analysis as the plausible effect size.

6.5 Discussion

Main Findings

In this umbrella review we evaluated the current evidence, derived from meta-analyses on the association between various risk factors for GDM. Overall 43 associations have been examined, including biomarkers, diet and lifestyle factors, diseases and disorders, infections, and other risk factors. However, only a minority of these associations, had strongly significant results with no suggestion of bias, as can be inferred by substantial heterogeneity between studies, small study effects, and excess significance bias. Five risk factors were supported by convincing evidence, including vitamin D deficiency, low vs. normal BMI (cohort studies), BMI ~30–35 vs. normal weight, BMI >35 vs. normal weight, and hypothyroidism. Another seven non-genetic risk factors from various fields [overweight vs. normal BMI (cohort), BMI >30 vs. normal weight, overweight vs. non-overweight women (cohort), overweight vs. non-overweight (case-control), obese vs. non-obese women (cohort), polycystic ovary syndrome, family history of diabetes], achieved highly suggestive evidence for an association with GDM.

Interpretation in light of evidence

It is well-known that maternal weight, as determined from pre-conception BMI, is critical on the development of insulin resistance and type II diabetes as well as GDM. Our findings show that the more robust associations were related to overweight and obesity, as three out of five associations that met the criteria for convincing evidence and five out of seven highly suggestive associations were concentrated on maternal pre-pregnancy BMI and the risk of GDM. We found that the association between BMI ~30–35 vs. normal weight, and BMI >35 vs. normal weight was supported by

convincing evidence for the risk of GMD, while another five risk factors that examined the link between overweight or obesity and GDM were supported by highly suggestive evidence. The association of low BMI vs. normal BMI was the only protective factor, which it was supported by convincing evidence for protection against GDM (Table 6.3).

Our findings further support the current guidelines regarding pregnancy weight, nutrition and activity, issued from the National Institute for Health and Clinical Excellence (NICE), the Institute of Medicine (IOM) and the American College of Obstetricians and Gynecologists (ACOG), which they accepted lifestyle change as an essential component of prevention and management of GDM (484–486). In the IOM report, pre-conception BMI was recognized as an independent predictor factor for many adverse pregnancy outcomes for both mother and child, while a BMI within the normal range between 18.5–24.9 prior to conception was recommended (484). ACOG endorses IOM weight guidelines and encourages weight loss before considering pregnancy in overweight and obese women, but has also developed separate recommendations for physical activity, recommending exercise for 30 minutes daily for all pregnant women (486). NICE recommendations include specific guidelines for healthy eating, low-fat diet and moderate physical activity before, during, and after pregnancy (485). Our findings are also in agreement with the latest evidence from a Cochrane systematic review of RCTs in which a possible reduction in GDM was found in women who received diet and exercise interventions during pregnancy compared with women who received standard care. Nevertheless, authors concluded that due to the variability of the diet and exercise components tested in the included studies, the evidence was insufficient to inform practice (487). However, issues of consistency and

clarity between reporting definitions and outcomes, could lead to incorrect inferences, which in turn may culminate in uninformed and inappropriate treatment choices. Likewise, subsequent clinical trials may waste limited resources and fail to confirm the previous published results. Large, well-designed, RCTs are needed to confirm the effectiveness of pre-conception weight and gestational weight gain reduction and the effects of dietary interventions in pregnancy for preventing GDM.

The observed association between obesity and GDM is biologically plausible. Normal pregnancy is characterized by a state of insulin resistance defined as an impaired response to insulin. This physiological insulin resistance also occurs in women with GDM on a background of chronic insulin resistance due to obesity to which the insulin resistance of pregnancy is partially additive. Obesity can cause major changes in maternal intermediary metabolism, where co-existing conditions associated with increased insulin resistance, higher serum lipids, and lower plasma levels of adiponectin, appear to play a central role to the development of GDM (488–490). Chronic inflammation is another possible explanation for the link between obesity and GDM. The exact inflammatory mechanisms involved in the development of GDM are not completely understood (491). However, several studies have shown a strong association between obesity and inflammatory markers, leading to the recognition of obesity as a state of chronic low-grade inflammation (492–497).

The association between vitamin D deficiency and risk of GDM was supported by convincing evidence. Even though vitamin D supplementation during pregnancy has been shown to have beneficial effects on glycaemia, insulin sensitivity, insulin resistance and metabolic profiles (498–500), it remains unclear to date whether routine

measurement of vitamin D levels during pregnancy should be recommended and/or whether supplementation should be recommended in all pregnant women or only in populations with insufficiency (501). Our findings are not compatible with the findings from two recent systematic reviews and meta-analyses of RCTs, in which authors report no reduction on the risk of gestational diabetes among those taking vitamin D supplements versus placebo (402,502). Routine vitamin D supplementation is not recommended for pregnant women to improve maternal and perinatal outcomes by the World Health Organization (WHO) due to the limited evidence currently available to directly assess the benefits and harms of the use of vitamin D supplementation alone in pregnancy (503). Although currently there is no specific national evidence-based guideline for vitamin D intake or supplementation in pregnancy, yet, existing guidelines agree that particular attention should be taken in high-risk groups, including vegetarians, women with limited sun exposure (e.g. those who live in cold climates, reside in northern latitudes, or wear sun and winter protective clothing), ethnic minorities, especially those with darker skin, and obese women (504–508). Proposed mechanisms that describe the link between vitamin D deficiency in relation to glucose metabolism include; the direct or indirect action of vitamin D on the pancreatic β-cell function and modulating insulin secretion by binding its circulating active form, 1,25-hydroxy vitamin D (1,25(OH)2D3), to a β-cell vitamin D receptor (509–511), as well as the influence of vitamin D on insulin resistance through regulation of extracellular and intracellular β-cell calcium pools, which is essential for insulin-mediated intracellular processes in insulin-responsive tissues (510–513).

The association between hypothyroidism, which includes both subclinical and overt hypothyroidism, and risk of GDM, was supported by convincing evidence. Increased

levels of human chorionic gonadotropin (hCG) in the first trimester of pregnancy directly stimulate the thyroid gland to increase production of thyroid hormone, which leads in decreased secretion of thyroid stimulating hormone (TSH) (514). Thyroid hormones exert profound effects in the regulation of glucose homeostasis, and hypothyroidism can have profound effects on glucose metabolism and insulin secretion. Proposed mechanisms that describe the relationship between hypothyroidism and gestational diabetes are supported from studies that show that both overt and subclinical hypothyroidism can lead to significantly increased insulin resistance (515–518). It is possible, therefore, that pregnant women with hypothyroidism have further amplified insulin resistance, and consequently an increased risk of gestational diabetes (477). Although, these findings would suggest that routine screening of thyroid hormones during pregnancy could be essential, nevertheless, universal thyroid screening in pregnancy is controversial (519). For example, the 2002 practice guidelines from ACOG recommend thyroid testing only in symptomatic high-risk pregnant women who have a personal history of thyroid disorders, type 1 diabetes or other autoimmune disorders, and do not recommend testing in asymptomatic women or women with small goiters (520). The Endocrine society recommends screening of pregnant women or those who wish to become pregnant and are at "high risk for thyroid illness" (e.g. women over 30 years old, with a family history or autoimmune thyroid disease or hypothyroidism, with a goiter, with symptoms or clinical signs suggestive of thyroid hypofunction etc.) on the basis of their medical history, physical exam, or prior biochemical data (521). On the contrary, the American Association of Clinical Endocrinologists (AACE) recommends routine thyroid function screening before pregnancy for all patients intended to be pregnant or during their first trimester (522).

In the current umbrella review, we applied a transparent and replicable set of criteria and statistical tests to evaluate and categorize the level of existing observational evidence. Although, an impressive 88% of the included meta-analyses report a nominally ($P<0.05$) statistically significant random-effects summary estimate, when stringent P value was considered ($P<10^{-6}$), the proportion of significant associations decreased to 32%. Eighteen (42%) associations had large or very large heterogeneity, while when we calculated the 95% prediction intervals, which further account for heterogeneity, we found that the null value was excluded in about half of the associations. Only five of the assessed risk factors found to provide convincing evidence, indicating that several published meta-analyses in the field could be susceptible to biases and the reported associations in the existing studies are often exaggerated.

The ability to modify those factors, mainly those related to overweight and obesity, through clinical interventions or public health policy measures remains to be established. Furthermore, there is no guarantee that even a convincing observational association for a modifiable risk factor would necessarily translate into large preventive benefits for GDM if these risk factors were to be modified (523). With obesity becoming a global epidemic, the assessment of the strength of the evidence supporting the impact of overweight and obesity in GDM could allow the identification of women at high risk for adverse outcomes and allow better prevention. Obesity is generating an unfavorable metabolic environment from early gestation; therefore, initiation of interventions for weight loss during pregnancy might be belated to prevent or reverse adverse effects, which highlights the need of weight management strategies before conception (188). GDM does not only increase the risk for maternal and fetal

complication in pregnancy, but also significantly increases a woman's risk of type 2 diabetes, metabolic syndrome (characterized by glucose intolerance, central obesity, dyslipidemia, and insulin resistance), and cardiovascular disease (CVD) after pregnancy (524–527).

Limitations

Umbrella reviews focus on existing systematic reviews and meta-analyses and therefore some studies may have not been included either because the original systematic reviews did not identify them, or they were too recent to be included. Statistical tests of bias in the body of evidence (small study effect and excess significance tests) offer hints of bias, not definitive proof thereof, while the Egger test is difficult to interpret when the between-study heterogeneity is large. These tests have low power if the meta-analyses include less than 10 studies and they may not identify the exact source of bias (82,275,528). Furthermore, we did not appraise the quality of the individual studies on our own, since this should be the responsibility of the authors of the original meta-analysis and it was beyond the scope of the current umbrella review. However, we recorded whether and how they performed a quality assessment of the synthesized studies. Lastly, we cannot exclude the possibility of selective reporting for some associations in several studies. For example, perhaps some risk factors were more likely to be reported, if they had statistically significant results.

6.6 Conclusion

The present umbrella review of meta-analyses identified 43 unique risk factors for GDM. Our analysis identified five risk factors with convincing evidence and strong epidemiological credibility pertaining to vitamin D deficiency, hypothyroidism and

BMI (specifically, low vs. normal BMI (cohort studies), BMI ~30–35 vs. normal weight, BMI >35 vs. normal weight). Most of these associations have an apparent biological plausibility but the exact mechanisms are not fully understood. As previously suggested, the use of standardized definitions and protocols for exposures, outcomes, and statistical analyses may diminish the threat of biases, allow for the computation of more precise estimates and will promote the development and training of prediction models that could promote public health.

Authors' contributions: KG and SP were involved in formulating the hypothesis and the design of the study protocol. KG and SP performed the literature search, the selection of eligible articles and the data extraction. KG analyzed the data. All authors (KG, EE, SP, CC, NM, EP and PY) were involved in data interpretation. KG and SP wrote the first draft of the manuscript and EE, CC, NM, EP and PY were involved in the revision of the manuscript. All authors (KG, EE, SP, CC, NM, EP, PY) approved the final version of the submitted manuscript. KG and SP are guarantors.

Chapter 6: Supplemental material

Supplemental Table 6.4. Analytical description of the 43 selected meta-analyses with observed and expected number of "positive" study datasets

Area	Author, year	Comparison	Studies	Cases/controls	Random effects*	Fixed effects†	Largest effect‡	Egger§	I^2 (95% CI) (P)‖	95% PI ≠	O¶	E # (fixed)	P** (fixed)	E ¥ (random)	P** (random)	E ∂ (largest)	P** (largest)
Biomarkers	Kong FJ 2017	Betatrophin levels	8	401/421	6.65 (2.12-20.9)	10.5 (7.92-13.9)	16.5 (9.18-29.8)	0.191	94 () (<0.001)	0.11-411.7	8	7.97	1.00	7.82	1.00	8.00	1.00
Biomarkers	Fu S 2016	Ferritin (highest vs lowest) (cohorts)	4	214/1662	3.22 (1.73-6.00)	3.22 (1.73-6.00)	4.98 (1.46-17.03)	0.953	0 () (0.815)	0.82-12.65	2	2.76	0.59	2.76	0.59	3.40	0.11
Biomarkers	Fu S 2016	Dietary total iron intake	3	1007/13850	1.01 (1.00-1.01)	1.01 (1.00-1.01)	1.12 (0.87-1.45)	NA	0 () (0.73)	0.99-1.03	2	0.15	0.01	0.15	0.01	0.59	0.10
Biomarkers	Fu S 2016	Serum ferritin (GMD-women vs non-GMD)	6	403/498	4.89 (2.06-11.58)	4.51 (3.50-5.82)	6.45 (4.07-10.24)	0.756	91 () (<0.001)	0.22-106.6	5	5.90	0.10	5.93	0.07	5.99	0.01
Biomarkers	Fernández-Cao2016	Hemoglobin levels	9	792/4393	1.54 (1.18-2.03)	1.53 (1.24-1.89)	0.81 (0.36-1.82)	0.752	33 () (0.157)	0.81-2.93	3	3.58	1.00	3.72	0.74	1.24	0.12
Biomarkers	Fernández-Cao2016	Ferritin (highest vs lowest) (mixed)	7	330/5574	2.09 (1.48-2.96)	2.09 (1.48-2.96)	2.27 (1.20-4.30)	0.600	1 () (0.42)	1.31-3.34	3	5.40	0.05	5.40	0.05	5.79	0.02
Biomarkers	Kong FJ 2016	Selenium level	7	178/391	0.12 (0.03-0.53)	0.21 (0.14-0.30)	0.12 (0.06-0.26)	0.499	93 () (<0.001)	0.00-19.81	6	5.73	1.00	6.52	0.39	6.51	0.40
Biomarkers	Hu S 2016	Serum retinol-binding protein-4	17	647/620	4.38 (2.10-9.14)	2.64 (2.12-3.27)	1.27 (0.70-2.30)	0.025	91 () (<0.001)	0.18-106.7	9	9.77	0.81	14.2	0.00	1.45	0.00
Biomarkers	Iliodromiti S 2016	Adiponectin	11	794/2071	6.35 (4.08-9.88)	6.27 (5.09-7.72)	5.05 (3.55-7.18)	0.770	71 () (<0.001)	1.56-25.9	10	10.6	0.37	10.6	0.36	10.3	0.52
Biomarkers	Guo CC 2016	DQ2	12	2333/2687	1.36 (1.10-1.66)	1.20 (1.06-1.36)	0.96 (0.79-1.16)	0.008	43 () (0.06)	0.80-2.30	4	2.03	0.13	3.91	1.00	0.68	0.00
Biomarkers	Guo CC 2016	DQ6	11	2270/2576	0.81 (0.69-0.94)	0.81 (0.69-0.94)	0.75 (0.55-1.02)	0.551	0 () (0.743)	0.67-0.97	1	2.40	0.48	2.40	0.48	3.52	0.19
Biomarkers	Guo CC 2016	DR 13	4	209/225	2.46 (1.02-5.90)	2.54 (1.62-3.99)	0.73 (0.29-1.87)	0.982	67 () (0.03)	0.07-88.5	2	2.82	0.59	2.71	0.60	0.57	0.10
Biomarkers	Guo CC 2016	DR17	5	329/335	3.16 (1.31-7.64)	2.59 (1.61-4.18)	3.13 (1.11-8.81)	0.116	69 () (0.01)	0.16-62.9	3	4.01	0.26	4.54	0.07	4.51	0.08
Biomarkers	Yang Y 2015	Thyroid antibodies (cohort)	11	1596/30012	1.07 (0.97-1.19)	1.07 (0.97-1.19)	1.18 (0.77-1.81)	0.546	0 () (0.44)	0.95-1.21	1	0.78	0.56	0.78	0.56	1.87	0.70
Biomarkers	Yang Y 2015	Thyroid antibodies (case-control)	10	856/2062	1.21 (1.05-1.41)	1.21 (1.05-1.41)	1.33 (1.09-1.63)	0.402	0 () (0.73)	1.02-1.44	1	1.25	1.00	1.25	1.00	1.90	0.70
Biomarkers	Yang Y 2015	Thyroid antibodies (All studies)	21	2452/32074	1.12 (1.03-1.22)	1.12 (1.03-1.22)	1.18 (0.77-1.81)	0.485	0 () (0.60)	1.02-1.23	2	1.91	1.00	1.91	1.00	2.93	0.76
Biomarkers	Zhang MX 2015	Vitamin D deficiency	20	1737/7472	1.55 (1.32-1.82)	1.53 (1.33-1.75)	1.38 (1.05-1.82)	0.110	16 () (0.25)	1.10-2.19	6	8.04	0.49	8.38	0.37	5.38	0.80
Biomarkers	Aghajafari F 2013	25(OH)D concentration	10	687/3425	1.49 (1.18-1.88)	1.49 (1.18-1.88)	1.35 (0.77-2.35)	0.580	0 () (0.58)	1.14-1.96	2	3.49	0.51	3.49	0.51	2.20	1.00
Biomarkers	Wei SQ 2013	25(OH)D5<50 nmol/l	10	623/3503	1.37 (1.11-1.70)	1.37 (1.11-1.70)	1.20 (0.72-2.00)	0.147	0 () (0.51)	1.07-1.76	1	2.24	0.70	2.24	0.70	1.06	1.00
Biomarkers	Wei SQ 2013	25(OH)D<75 nmol/l	8	542/3298	1.52 (1.17-1.98)	1.53 (1.19-1.96)	1.63 (0.79-3.33)	0.954	7 () (0.37)	1.01-2.30	2	3.09	0.72	3.06	0.72	3.85	0.29
Diet and lifestyle	Aune D 2016	Leisure-time PA before pregnancy	8	2401/30191	0.78 (0.61-1.00)	0.78 (0.67-0.90)	0.81 (0.68-1.01)	0.869	47 () (0.07)	0.41-1.47	4	3.28	0.72	3.25	0.72	2.70	0.46
Diet and lifestyle	Aune D 2016	Leisure-time PA during pregnancy	5	580/5140	0.97 (0.73-1.28)	0.97 (0.73-1.28)	0.91 (0.37-2.21)	0.430	0 () (0.80)	0.61-1.52	0	0.27	1.00	0.27	1.00	0.40	1.00
Diet and lifestyle	Torloni MR 2009	Low vs. Normal BMI (cohort)	16	75669/280734	0.75 (0.69-0.83)	0.77 (0.71-0.82)	0.80 (0.69-0.92)	0.022	16 () (0.27)	0.63-0.90	6	11.2	0.01	11.5	0.00	10.3	0.03
Diet and lifestyle	Torloni MR 2009	Low vs. Normal BMI (case-control)	3	5957/11651	0.65 (0.51-0.83)	0.65 (0.51-0.82)	0.61 (0.47-0.81)	0.572	0 () (0.83)	0.13-3.16	1	1.78	0.57	1.78	0.57	1.96	0.28
Diet and lifestyle	Torloni MR 2009	Overweight vs. Normal BMI (cohort)	17	112880/282458	1.97 (1.76-2.19)	2.05 (1.94-2.15)	2.29 (2.12-2.47)	0.521	56 () (0.003)	1.44-2.68	11	16.2	0.00	16	0.00	16.5	0.00
Diet and lifestyle	Torloni MR 2009	Overweight vs. Normal BMI (case-control)	3	287/501	2.68 (1.78-4.04)	2.66 (1.95-3.63)	3.85 (2.30-6.47)	0.889	40 () (0.19)	0.05-138	3	2.86	1.00	2.86	1.00	2.99	1.00
Diet and lifestyle	Torloni MR 2009	Obese (BMI >30) vs. normal weight	31	56333/308335	3.76 (3.31-4.28)	3.94 (3.75-4.13)	4.80 (4.43-5.21)	0.661	73 () (<0.001)	2.23-6.34	26	31	0.00	31	0.00	31	0.00
Diet and lifestyle	Torloni MR 2009	Obese 1 (BMI ~30–35) vs. Normal weight	6	3087/20901	3.01 (2.34-3.86)	3.06 (2.62-3.56)	3.21 (2.68-3.85)	0.612	27 () (0.23)	1.71-5.28	5	5.79	0.19	5.78	0.20	5.83	0.16
Diet and lifestyle	Torloni MR 2009	Obese 2 (BMI >35) vs. Normal weight	7	1747/21001	5.52 (4.28-7.11)	5.42 (4.36-6.73)	5.10 (3.18-8.19)	0.157	7 () (0.37)	3.62-8.42	6	6.98	0.02	6.98	0.02	6.96	0.03
Diet and lifestyle	Torloni MR 2009	Overweight vs. Non-overweight (cohort)	34	174233/391991	2.95 (2.68-3.24)	2.81 (2.71-2.91)	3.10 (2.91-3.31)	0.132	72 () <0.001)	1.97-4.41	31	33.9	0.00	33.9	0.00	34	0.00
Diet and lifestyle	Torloni MR 2009	Overweight vs. Non-overweight (case-control)	10	6214/19567	3.78 (2.49-5.76)	3.03 (2.68-3.42)	3.06 (2.51-3.73)	0.248	90 () (<0.001)	0.83-17.2	9	9.65	0.30	9.87	0.13	9.67	0.29
Diet and lifestyle	Torloni MR 2009	Obese vs. non-obese women (cohort)	40	68013/520879	3.36 (3.01-3.74)	3.27 (3.14-3.41)	3.44 (3.20-3.70)	0.724	77 () (<0.001)	1.97-5.72	36	40	0.00	40	0.00	40	0.00
Diet and lifestyle	Torloni MR 2009	Obese vs. non-obese women (case-control)	3	238/922	3.24 (1.28-8.19)	3.36 (2.48-4.55)	7.49 (4.58-12.27)	0.938	88 () (0.001)	0-285401	2	2.92	0.08	2.9	0.10	3.00	0.00
Diseases/disorders	Gong LL 2016	Overt hypothyroidism	3	3444/222161	2.44 (1.08-5.52)	1.90 (1.69-2.14)	1.88 (1.67-2.12)	0.688	57 () (0.10)	0-15039	2	2.20	1.00	2.68	0.29	2.17	1.00
Diseases/disorders	Gong LL 2016	Subclinical hypothyroidism	6	1859/61708	1.59 (1.32-1.92)	1.59 (1.32-1.92)	1.49 (1.04-2.13)	0.208	0 () (0.50)	1.22-2.07	3	5.26	0.03	5.26	0.03	5.00	0.06
Diseases/disorders	Gong LL 2016	Hypothyroidism (all)	7	5770/278609	1.72 (1.51-1.95)	1.76 (1.60-1.94)	1.88 (1.67-2.12)	0.137	14 () (0.32)	1.35-2.18	5	6.69	0.04	6.64	0.05	6.78	0.02
Diseases/disorders	Luque-Fernandez 2013	Sleep-disordered breathing	9	673/9122	2.18 (1.59-2.98)	1.80 (1.51-2.14)	1.44 (0.99-2.10)	0.011	52 (0.03)	0.95-4.97	7	4.58	0.18	5.64	0.50	2.91	0.01
Diseases/disorders	Kjerulff LE 2011	Polycystic ovary syndrome	18	2385/89669	2.83 (1.95-4.10)	2.68 (2.36-3.05)	2.69 (2.33-3.11)	0.653	52 () (0.005)	0.94-8.46	8	14.9	0.00	15.4	0.00	14.9	0.00
Infections	Abariga SA 2016	Periodontitis	10	624/5100	1.66 (1.16-2.36)	1.48 (1.17-1.87)	1.73 (0.91-3.30)	0.008	51 () (0.03)	0.61-4.49	3	2.72	0.74	4.03	0.75	4.54	0.33
Infections	Soepnel LM 2016	HIV infection	4	593/1070	0.83 (0.48-1.42)	0.83 (0.48-1.42)	1.00 (0.37-2.71)	0.472	0 () (0.61)	0.25-2.71	0	0.67	1.00	0.67	1.00	0.20	1.00
Other	Moosazadeh 2016	Family history of diabetes	33	2697/29134	3.46 (2.80-4.27)	3.50 (3.17-3.86)	4.36 (2.89-6.58)	0.861	76 () (<0.001)	1.17-10.2	25	32.2	0.00	32.2	0.00	16.1	0.00
Other	Xu Y 2016	Isolated Single Umbilical Artery	7	1880/490712	1.38 (1.06-1.80)	1.43 (1.17-1.75)	2.08 (1.47-2.96)	0.569	35 () (0.16)	0.73-2.61	1	4.96	0.00	4.54	0.01	6.86	0.00
Other	Pandey S 2012	IVF/ICSI versus spontaneous conception	6	13399/574391	1.31 (0.98-1.75)	1.50 (1.34-1.68)	1.55 (1.37-1.75)	0.169	42 () (0.13)	0.63-2.72	1	4.26	0.01	3.22	0.10	4.48	0.00

Abbreviations: Random effects, summary odds ratio (95% CI) using random effects model; Largest effect, odds ratio (95% CI) of the largest study in the meta-analysis; Egger, p-value from Egger's regression asymmetry test for evaluation of publication bias; P, p-value; NP, not pertinent, because the estimated is larger than the observed, and there is no evidence of excess of statistical significance based on the assumption made for the plausible effect size; BMI, Body Mass Index; GDM, gestational diabetes mellitus; PA, physical activity.

* Summary random effects odds ratio (95% CI) of each meta-analysis, except for three meta-analyses (Fu S 2016, Aune D, 2014 and Pandey S 2012) where the RR was used.
† Summary fixed effects odds ratio (95% CI) of each meta-analysis, except for three meta-analyses (Fu S 2016, Aune D, 2014 and Pandey S 2012) where the RR was used.
‡ Odds ratio (95% CI) of the largest study in each meta-analysis, except for three meta-analyses (Fu S 2016, Aune D, 2014 and Pandey S 2012) where the RR was used.
§ P-value from the Egger regression asymmetry test for evaluation of publication bias
|| I^2 metric of inconsistency (95% confidence intervals of I^2) and P-value of the Cochran Q test for evaluation of heterogeneity
≠ 95% Prediction Interval
¶ Observed number of statistically significant studies
# Expected number of statistically significant studies using the summary fixed effects estimate of each meta-analysis as the plausible effect size
** P-value of the excess statistical significance test
¥ Expected number of statistically significant studies using the summary random effects estimate of each meta-analysis as the plausible effect size
ϑ Expected number of statistically significant studies using the effect of the largest study of each meta-analysis as the plausible effect size

Chapter 7 – Summary and Future Directions

7.1 Summary of Major Findings

Currently, biomedical and public health research is conducted on a massive scale, where nearly one million articles on humans are published each year (9). Driven by the rapid increase in the number of published studies, scientists turn into systematic reviews and meta-analyses to summarize the evidence on a research question, using multiple related studies in a rigorous and replicable way (10). Although systematic reviews and meta-analyses are considered the highest level of evidence and may accelerate evidence uptake, their credibility is under threat as most of them appear to be either not useful or of uncertain utility (9,529). The problem is that the majority are unnecessary (duplicative), inaccurate or misleading due to biases in the methodology and selective reporting of results, or they address questions that have no clinical value (9,26). The increase in the number of systematic reviews, along with escalating demand from policy makers for rapid reviews of research, has driven an increase in a newer form of synthesis, umbrella reviews (11,530). An umbrella review can provide an overall assessment of the body of evidence that is available on a given topic using the data from multiple systematic reviews and meta-analyses (11,25). This comprehensive assessment of epidemiological evidence with the goal of providing an objective summary of the available data is central not only for understanding the reliability of an evidence-base to effective decision making but also as the basis for clinical and public health recommendations.

This thesis focuses on the systematic assessment of current evidence across the published systematic reviews and meta-analyses of clinical entities with a large impact on perinatal epidemiology, specifically preeclampsia and gestational diabetes, through the performance of umbrella review approach. As previously explained, an umbrella review is a new method to summarise and synthesise the evidence from multiple systematic reviews and meta-analyses into one accessible publication. We believe this evaluation of the quality of research evidence that includes a robust hierarchical classification of the published evidence and its interpretation will help to inform decision making of clinicians, policy-makers and regulatory bodies as well as to researchers looking to contribute to the evidence base through targeted evidence synthesis. Additionally, it could also serve for the optimization of preeclampsia and gestational diabetes prediction models and identification of the women at high risk.

In Chapter 4, an umbrella review of systematic reviews and meta-analyses of observational studies was carried out to summarize evidence on the factors that have been associated with preeclampsia, evaluate whether there are hints of biases in this literature and how they manifest and finally identify which of the previously studied associations include convincing evidence to support their results. Fifty-eight eligible papers were identified providing data on 130 associations including 1466 primary studies, covering a very wide range of risk factors. Sixty-five (50%) associations had nominally statistically significant findings at $P<0.05$, while sixteen (12%) were significant at $P<10^{-6}$. Sixty-four (49%) associations had large or very large heterogeneity. Evidence for small-study effects and excess significance bias was found in ten (8%) and twenty-six (20%) associations, respectively. Oocyte donation vs spontaneous conception (OR 4.33, 95% CI: 3.11-6.03) had >1000 cases, 95%

prediction intervals excluding the null, not suggestive of large heterogeneity (I^2<50%), small-study effects (P for Egger's test>0.10), or excess of significance (P>0.05). Across the statistically significant genetic risk factors (P<0.05), only PAI-1 4G/5G (recessive model) polymorphism was supported with strong evidence for a contribution to the pathogenesis of preeclampsia. In addition, another eleven risk factors, presented highly suggestive evidence for preeclampsia. The results indicate that a large proportion of systematic reviews and meta-analyses of genetic and non-genetic risk factors for preeclampsia have caveats, which threaten their validity. Only oocyte donation vs spontaneous conception and PAI-1 4G/5G polymorphism (recessive model) show the strongest consistent evidence for a contribution to the pathogenesis of preeclampsia.

In Chapter 5, the evidence from published systematic reviews and meta-analyses of randomized controlled trials for preeclampsia prevention was collectively summarized and evaluated using the umbrella review approach. Twenty-nine eligible meta-analyses were identified that included 456 primary studies, providing data on 57 associations. Twenty-four (42%) associations had nominally statistically significant findings at p<0.05, while only 10 (18%) were significant at $p<10^{-3}$ under the random-effects model. Sixteen (28%) associations had large or very large heterogeneity. Evidence of excess significance bias was found in 15 (26%) associations. After applying our classification criteria, the following three interventions were classified as Class I level of evidence including low dose aspirin ≤16 weeks of gestation for preterm preeclampsia, diet and nutrition counselling and dietary interventions. This analysis demonstrated that from the available pharmacologic and non-pharmacologic interventions, early administration of low dose aspirin ≤16 weeks of gestation for

prevention of preterm preeclampsia, diet and nutrition counselling and dietary interventions had the strongest epidemiologic evidence suggesting their effectiveness. The findings also highlight the importance of patient education on diet and lifestyle modifications in reducing the risk of preeclampsia, as well as the recommendation for early administration of aspirin in women at high risk pregnancies.

In Chapter 6, an umbrella review of meta-analyses of observational studies was performed to summarize evidence on the protective or risk factors associated with gestational diabetes mellitus, evaluate whether there are indications of biases in this literature, and identify which of the previously reported associations are supported by convincing evidence. Twenty-one eligible meta-analyses were identified, providing data on 43 associations based on 480 primary studies covering a very wide range of risk factors. Thirty-eight (88%) associations had nominally statistically significant findings at $P<0.05$, while only 14 (32%) were significant at $P<10^{-6}$ under the random-effects model. Only five risk factors presented convincing evidence for an association with GDM: vitamin D deficiency, low vs. normal BMI (cohort studies), BMI ~30-35 kg/m^2 vs. normal BMI, BMI >35 kg/m^2 vs. normal BMI, and hypothyroidism. The results highlight the importance of patient education on diet and lifestyle modifications as candidate interventions to reduce the risk of GDM.

7.2 Limitations

In the present study, we applied the umbrella review approach for summarizing data from already published systematic reviews and meta-analyses. This approach takes full advantage of building on existing systematic reviews and meta-analyses to perform a standardized methodological assessment of the epidemiological credibility

of the findings using a wide range of tests and criteria. Although, the present study adds an additional level of different and relatively objective critical appraisal and evidence grading criteria, some limitations exist that should be considered when interpreting the findings.

As with other forms of evidence synthesis, the utility of umbrella reviews will be largely dependent on the availability of published systematic reviews and meta-analyses. Hence, this approach may favour the selection of more commonly and readily studied risk factors or interventions, since they are more likely to be included in a meta-analysis. Consequently, we cannot eliminate the possibility that some promising factors or interventions were excluded, either because were too recent to be fulfil the eligibility criteria, or despite having sufficient data, do not have a corresponding eligible meta-analysis. However, this possibility is becoming less likely in the current era, with systematic reviews and meta-analyses being conducted on a massive scale, to the point that for several topics multiple meta-analyses are available (9,12). In addition, for most putative risk factors or interventions that are difficult or uncommon to study, the current evaluation of evidence is unlikely to be remarkable, given the limited data.

Furthermore, we did not appraise the quality of the individual studies directly, since this should be the responsibility of the authors of the original systematic reviews and meta-analyses, and it was beyond the scope of the umbrella review. Thus, because we depended on the original meta-analyses quality assessment, and ultimately the studies that they include, we cannot exclude the possibility that deficiencies in the methodological quality at each level can compromise the results and conclusions of an

umbrella review. Nevertheless, we examine whether the original systematic reviews and meta-analyses applied any criteria to assess the quality of the synthesized studies. For instance, in Chapter 4, we found that the quality assessment of the primary studies was very heterogeneous, and this reflects the lack of standardized quality assessment methodologies. In the same Chapter, we have also assessed the quality of the included studies of the meta-analysis of the risk factors that presented convincing evidence for an association with preeclampsia, using the Newcastle Ottawa Scale and the Q-Genie tool. Likewise, in Chapter 5, we noted the conclusions from any evidence classification applied by the authors of the original meta-analyses for the interventions presented Class I and Class II evidence for preeclampsia prevention. The quality assessment for one of the two interventions that presented Class II evidence, namely L-arginine vs. placebo, was graded as regular to high quality using the Jadad scale, however, this scale is outdated and this needs to be considered in the overall evaluation of the evidence.

Statistical tests of bias in the body of evidence, namely Egger and excess of significance test, offer hints of bias, not a definitive proof thereof. The Egger test is difficult to interpret when the between-study heterogeneity is large, whereas the interpretation of the excess of statistical significance test for the results of a single meta-analysis, especially in those with few studies, should be cautious because a negative test does not exclude potential bias (82). In addition, it is possible that the results of studies included in a meta-analysis to have previously been standardized (e.g. cleaned or made to follow consistent definitions and adjustments) compared with the results presented in each study's original paper. Such standardization efforts are

likely to reduce, if anything, inconsistency and selective reporting bias, whereas the last, may be more prominent in the primary study reports.

In Chapter 4, most of the included studies for non-genetic associations were retrospective which is indicative of a higher potential for bias inherent in the included studies. We address this limitation by performing a standardized methodological process for the assessment of the epidemiological credibility of the findings using a variety of test, to accomplish the incorporation of all these biases together and provide a complete picture of the totality of evidence as it stands today. In Chapter 5, it is probable that for some types of interventions, only meta-analyses of observational data exist with no respective randomized evidence and these would not have been captured by our search. Likewise, even though our analysis in Chapter 5 identified diet and nutrition counselling and dietary interventions to had strong epidemiological credibility for prevention of preeclampsia, yet, some of the included studies had a large proportion of obese pregnant women, therefore our results should be interpreted with caution. In addition, due to the heterogeneity of both the pathophysiological pathways and clinical presentations of preeclampsia, it is possible that our results to be modified based on the presence of the other risk factors such as diabetes and obesity which are associated with cardiovascular disease.

7.3 Clinical Implications

Since the first successful use of donated oocytes in 1984, oocyte donation (OD) has become an increasingly more accepted method of assisted reproduction, leading to a dramatically increased of OD cycles in Europe and USA (531–533). In nowadays, it is considered to be an integral part of infertility treatment, especially to overcome

infertility due to advanced age (533). Nevertheless, OD has its own associated risks, and this should call for clinician's special awareness given that OD is becoming increasingly prevalent in line with modern living, not only for mothers who are older (aged over 45), but also in younger women (534). It is now well documented, that OD pregnancies are associated with increased risk of hypertensive diseases in pregnancy. In fact, our results showed that OD vs. spontaneous conception have the strongest consistent evidence for a contribution to the pathogenesis of preeclampsia. The etiology of preeclampsia in donor oocyte pregnancies is yet to be clarified. As previously discussed, the most likely hypothesis to explain preeclampsia in OD pregnancies has been postulated based on the lack of immunological tolerance of a fetus whose entire genome is allogeneic to that of the mother's (533,535). In light of these results and regardless of the preeclampsia etiology, a certain number of factors should be considered by clinicians.

Before authorizing a reproductive assistance with OD, it seems critical to careful select patients. Women should be screened accurately and assess for a certain number of preexisting risk factors for preeclampsia, such as hypertension, diabetes, obesity, renal disease and chronic infections (535). In the presence of risk factors, counseling by fertility experts prior to treatment should be mandatory to advice on the increased risks of OD and possible treatment options of any modifiable risk factors (533). Furthermore, as multiple pregnancy increases the risk of adverse maternal and fetal outcomes, the option of transferring a single high-quality embryo ought to be favored (536). Pregnant woman who have conceived using OD need to be categorized as high-risk patients, and close clinical, ultrasound, and biological monitoring throughout the pregnancy and repeated measurement of blood pressure in both arms to identify

hypertension is recommended (535,537). In addition to close monitoring of the pregnancy, women conceived using OD should, if possible, be under the care of obstetricians specializing in maternofetal medicine who will be prepared appropriately antenatally for delivery and the puerperium (537).

The recognition of risks associated with OD pregnancies should lead clinicians to consider tailored clinical surveillance and possibly preventive strategies such as early aspirin therapy (before the 16th week of gestation), which improves deep placentation and could prevent or delay the appearance of early preeclampsia (290,409,538). In line with previous recommendations from ACOG and NICE, our umbrella review confirms and supports the results from the earlier evaluations that early administration of low dose aspirin ≤16 weeks of gestation is an effective intervention for prevention of preterm preeclampsia. In addition, we demonstrated that from the available non-pharmacologic interventions, diet and nutrition counselling and dietary interventions had the strongest consistent evidence suggesting their effectiveness. It is possible that dietary interventions, such as a balanced diet consisting of carbohydrates, proteins and fat, calorie-controlled or low-fat diet are effective in modifying metabolic factors such as lipid levels, blood pressure, and glycose or reducing gestational weight gain with a potential contribution to a lower risk for preeclampsia (418). This highlights the importance of patient education on nutrition and lifestyle modifications in preventing not only preeclampsia, but other important co-morbidities, such as gestational diabetes.

Lifestyle change is an essential component of prevention and management of GDM too. Our findings demonstrated that among the non-genetic risk factors for GDM the

most epidemiological credible factors were concentrated on maternal pre-pregnancy BMI, overweight and obesity. Preconception counseling on the factors associated with GDM, the short and long-term risks of GDM, and the importance of healthy lifestyle should be incorporated into routine medical care for all women of childbearing potential (539,540). With type 2 diabetes becoming a global epidemic, which in part related to the epidemic of overweight and obesity in the population, and due to the possible complications of undiagnosed gestational diabetes, universal screening for this entity is widely practiced and is recommended (541,542). It is important that all women to be screened early in their pregnancy for preexisting risk factors associated with gestational diabetes (e.g. maternal age, previous GDM, and obesity) as well as for other independently predictor factors (e.g. pre-pregnancy BMI) to identify high risk women and consequently allow better prevention. Clinicians and other care providers should focus together, on how to support pregnant women to make positive lifestyle changes from the time of the initial comprehensive medical evaluation. For the pregnant categorized as high-risk patients, a close clinical monitoring throughout the pregnancy is recommended to maintain a high index of suspicion for associated conditions and complications (543,544).

7.4 Future Directions

With the ever-increasing number of systematic reviews published every day, reviews of systematic reviews, aka umbrella reviews can provide a comprehensive assessment of the body of information using data from all systematic reviews and meta-analyses on a given research topic (11,257). Such reviews emerged only recently, and their number is increasing since their content is an attractive way to summarize, evaluate and translate large amounts of evidence into one accessible document that can be used

by scientists, clinicians, and policy makers to support evidence-based decision-making (22,25). This higher-level synthesis of evidence permits an understanding of the spread of summary effects, heterogeneity, hints of bias and quality features that affect the credibility of the results in different systematic reviews in a whole research field (11). They can also bring efficiencies that could lessen research waste and provide suggestions on how to improve the design, quality and rigour of future primary studies (11,545). Such evaluations can also form the basis for higher level integrative documents such as risk assessments, practice guidelines, and decision tools (546).

Notwithstanding their weaknesses, systematic reviews and meta-analyses will continue to be extremely influential and have a major value. Their credibility and utility are probably better than almost any other type of biomedical article published, excluding large randomized trials (19). While the development of methods of higher-level synthesis such as umbrella reviews is essential to improve evidence-based decision-making, this effort needs to happen in tandem with improvements in the conduct and reporting of systematic reviews. To achieve this, a coordination amongst review teams examining different parts of a broad evidence synthesis is essential. The international prospective register of systematic reviews (PROSPERO) is a promising new initiative that could play an important role in this coordinated effort through the linking of review teams (545). The purpose of this international initiative is not only to stimulate collaboration, but also reduce unplanned duplication of research efforts and to provide transparency in the review process with the aim of minimizing reporting bias (547). As previously described, because most of systematic reviews and meta-analyses published today are retrospective, explicitly objective methods for the conduct of the review, focusing on the control of error, both from bias and random

error, are essential to be defined a *priori* in a protocol (548). That protocol should be published or at least registered online (e.g. PROSPERO) so that it can be accessed and compared to what is finally published in the completed review. In addition, the protocol is important in distinguishing a systematic review from a narrative review which can so easily drift by being influenced by what is found in the searches rather than remaining focused on the defined question (548).

As previously clarified, the overall validity and quality of a systematic review is inextricably linked with the use of accurate synthesis methods and good reporting of individual studies (10,21). Although there has been an implosion of guidelines and tools to ensure proper reporting and adoption of rigorous methods of systematic reviews and meta-analyses, still, many of those being published are poorly conducted and reported. Poor conduct can lead to systematic reviews with misleading results, while poor reporting prevents users from being able to determine the validity of the methods used. Consequently, apart from research waste, the validity of these systematic reviews and meta-analyses is diminished which limit their value to guide policy-decisions and clinical practice. Strategies are needed to increase the adoption of reporting guidelines that may help improve the reliability of this important literature in the future. The endorsement of guidelines by journals that would not only encouraging their use, but rather implementing systems to monitor adherence is also very important (41). In addition, it is vital that clinicians and other healthcare specialists, researchers, and editors to be trained with critical appraisal skills to be able to distinguish high-and low quality systematic reviews. It is also significant to underline the importance of encourage efforts to stop the growing number of

“predatory journals” that publishing anything quickly, with little or no peer review or quality control (41).

The pervasive documentation of bias suggests that more should be done to improve the quality of the primary evidence that forms the backbone of an evidence base, rather than expect from systematic reviews to correct deficiencies after the fact. The problem of having so much unreliable and non-useful published medical research could be eliminated if we attack at its root, that is by funding, conducting, publishing and disseminating more true and useful primary research. There are quite a few ways to improve primary evidence in the future. Foremost, if studies were more completely and transparently reported according to published guidelines, such as the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE), Consolidated Standards of Reporting Trials (CONSORT), and the Animal Research Reporting In Vivo Experiments (ARRIVE) guidelines, it is possible that many of the biases in the scientific literature to be substantially lessened. Although, reporting guidelines are often onerous, and authors and editors present challenges to managing journal page constraints, nevertheless, they have increased the standardization of reporting study results, which help to ensure that critical information is available for systematic reviews or other evidence-synthesis studies (530). Lack of standardization on reporting of data in individual studies can make quality appraisal difficult when conducting a systematic review and has a potential to contribute to missing data (530). In addition, despite that many journals require the adherence to specific reporting guidelines for a research manuscript to be considered for publication, yet, endorsement of reporting guidelines by journals is highly variable, leaving areas for improvement.

At the same time, it cannot be stressed enough the need for a system for registering animal experiments, analogous to that for clinical trials, which would help to reduce publication bias and provide a more informed view before proceeding to clinical trials. In addition, there must be a firmer attitude toward insisting on complete systematic reviews and meta-analyses of animal studies before embarking in clinical trials. Many clinical trials would probably not have gone ahead if all the data had been subjected to meta-analysis. Such reviews would also provide robust estimates of effect size and variance for adequately powering randomized trials. As previously suggested, the use of standardized definitions and protocols for exposures, outcomes, and statistical analyses may diminish the threat of biases, allow for the computation of more precise estimates which also help improve the evidence in the future. Advancements in evidence synthesis such as umbrella reviews and reporting guidelines will ultimately improve the quality, scope, and applicability of results and consequently future health care research, clinical practice, and public health policy.

REFERENCES

1. Ioannidis JPA, Boyack KW, Klavans R. Estimates of the continuously publishing core in the scientific workforce. *PLoS One*. 2014;9(7):e101698.
2. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: How will we ever keep up? *PLoS Med*. 2010;7(9):e1000326.
3. Créquit P, Trinquart L, Yavchitz A, Ravaud P. Wasted research when systematic reviews fail to provide a complete and up-to-date evidence synthesis: The example of lung cancer. *BMC Med*. 2016;14(1):8.
4. Ioannidis JPA. Why Most Clinical Research Is Not Useful. *PLoS Med*. 2016;13(6):e1002049.
5. Ioannidis JPA. How to Make More Published Research True. *PLoS Med*. 2014;11(10):e1001747.
6. Liberati A. Need to realign patient-oriented and commercial and academic research. *The Lancet*. 2011;378:1777–8.
7. Fox DM. Evidence and health policy: Using and regulating systematic reviews. *American Journal of Public Health*. 2017;107(1): 88–92.
8. Starr M, Chalmers I, Clarke M, Oxman AD. The origins, evolution, and future of the Cochrane Database of Systematic Reviews. *International Journal of Technology Assessment in Health Care*. 2009;25(S1):182–95.
9. Ioannidis JPA. The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. *Milbank Quarterly*. 2016;94(3): 485–514.
10. Page MJ, Shamseer L, Altman DG, Tetzlaff J, Sampson M, Tricco AC, et al. Epidemiology and Reporting Characteristics of Systematic Reviews of Biomedical Research: A Cross-Sectional Study. PLoS Med. 2016;13(5).

11. Ioannidis J. Next-generation systematic reviews: Prospective meta-analysis, individual-level data, networks and umbrella reviews. *British Journal of Sports Medicine*. 2017;51(20):1456–8.

12. Siontis KC, Hernandez-Boussard T, Ioannidis JPA. Overlapping meta-analyses on the same topic: survey of published studies. *BMJ*. 2013;347:f4501.

13. Bolland MJ, Grey A. A case study of discordant overlapping meta-analyses: vitamin d supplements and fracture. *PLoS One*. 2014;9(12):e115934.

14. Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, et al. Research: increasing value, reducing waste 5 Reducing waste from incomplete or unusable reports of biomedical research. *Lancet*. 2014;383:267–76.

15. Ioannidis JPA, Tarone R, McLaughlin JK. The false-positive to false-negative ratio in epidemiologic studies. *Epidemiology*. 2011;22(4):450–6.

16. Ioannidis JPA, Chang CQ, Lam TK, Schully SD, Khoury MJ. The Geometric Increase in Meta-Analyses from China in the Genomic Era. *PLoS One*. 2013;8(6):e65602.

17. Schuit E, Ioannidis JPA. Network meta-analyses performed by contracting companies and commissioned by industry. *Syst Rev*. 2016;5(1):198.

18. Ioannidis JPA. Why most discovered true associations are inflated. *Epidemiology*. 2008;19(5):640–8.

19. Ioannidis JPA. Why most published research findings are false. *PLoS Medicine*.2005; 2(8):0696–701.

20. Rifai N, Altman DG, Bossuyt PM. Reporting bias in diagnostic and prognostic studies: Time for action. *Clinical Chemistry*. 2008;54(7):1101–3.

21. Klimo P, Thompson CJ, Ragel BT, Boop FA. Methodology and reporting of

meta-analyses in the neurosurgical literature. *J Neurosurg*. 2014;120(4):796–810.

22. Ioannidis JPA. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ*. 2009;181(8):488–93.
23. Serghiou S, Patel CJ, Tan YY, Koay P, Ioannidis JPA. Field-wide meta-analyses of observational associations can map selective availability of risk factors and the impact of model specifications. *J Clin Epidemiol*. 2016;71:58–67.
24. Ioannidis JPA, Fanelli D, Dunne DD, Goodman SN. Meta-research: Evaluation and Improvement of Research Methods and Practices. *PLoS Biol*. 2015;13(10):1–7.
25. Tsagris M, Fragkos KC. Umbrella reviews, overviews of reviews, meta-epidemiologic studies: similarities and differences. In Umbrella Reviews. Springer International Publishing: Switzeland, 2016;43-54.
26. Ioannidis JPA, Stuart ME, Brownlee S, Strite SA. How to survive the medical misinformation mess. *Eur J Clin Invest*. 2017;47(11):795–802.
27. Khan KS, Kunz R, Kleijnen J, Antes G, Royal Society of Medicine (Great Britain). Systematic reviews to support evidence-based medicine : how to review and apply findings of healthcare research. Crc Press; 2011.
28. Lipp A. A Guide to Developing a Systematic Review. *AORN J*. 2003;78(1):90–107.
29. Pölkki T, Kanste O, Kääriäinen M, Elo S, Kyngäs H. The methodological quality of systematic reviews published in high-impact nursing journals: A review of the literature. *Journal of Clinical Nursing.* 2014;23:315–32.

30. Murad MH, Montori VM. Synthesizing evidence: Shifting the focus from individual studies to the body of evidence. *JAMA*. 2013; 309(21):2217–8.

31. Murad MH, Montori VM, Ioannidis JPA, Jaeschke R, Devereaux PJ, Prasad K, et al. How to read a systematic review and meta-analysis and apply the results to patient care: Users' guides to the medical literature. *JAMA*.2014;312(2): 171–9.

32. Greenhalgh T. How to read a paper: Assessing the methodological quality of published papers. *BMJ*. 1997;315(7103):305.

33. Mulrow CD. Rationale for systematic reviews. *Br Med J*. 1994;309(6954):597–9.

34. Oxman AD, Guyatt GH. Guidelines for reading literature reviews. *Can Med Assoc J*. 1988;138(8):697–703.

35. Higgins JP, Green S. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series. Wiley-Blackwell; 2011.

36. Egger M, Smith GD. Meta-analysis: Potentials and promise. *BMJ*. 1997;315(7119):1371–4.

37. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. *PLoS Medicine*. 2007;4:447–55.

38. Chalmers I. The Cochrane collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Ann N Y Acad Sci.* 1993;703(1):156-65.

39. Levin A. The Cochrane Collaboration. *Annals of Internal Medicine*. 2001;135(4):309-312.

40. Mallett S, Clarke M. How many Cochrane reviews are needed to cover

existing evidence on the effects of health care interventions? *ACP J Club*. 2003;139(1):A11–A11.

41. Page MJ, Moher D. Mass Production of Systematic Reviews and Meta-analyses: An Exercise in Mega-silliness? *Milbank Quarterly*. 2016;94:515–9.
42. Owens DK, Lohr KN, Atkins D, Treadwell JR, Reston JT, Bass EB, et al. AHRQ Series Paper 5: Grading the strength of a body of evidence when comparing medical interventions-Agency for Healthcare Research and Quality and the Effective Health-Care Program. *J Clin Epidemiol*. 2010;63(5):513–23.
43. Patsopoulos NA, Analatos AA, Ioannidis JPA. Relative citation impact of various study designs in the health sciences. *J Am Med Assoc*. 2005;293(19):2362–6.
44. Garritty C, Tsertsvadze A, Tricco AC, Sampson M, Moher D. Updating systematic reviews: An international survey. *PLoS One*. 2010;5(4):e9914.
45. Cook DJ, Reeve BK, Guyatt GH, Heyland DK, Griffith LE, Buckingham L, et al. Stress Ulcer Prophylaxis in Critically III Patients. *JAMA*. 1996;275(4):308.
46. Campbell J, Bellamy N, Gee T. Differences between systematic reviews/meta-analyses of hyaluronic acid/hyaluronan/hylan in osteoarthritis of the knee. *Osteoarthr Cartil*. 2007;15(12):1424–36.
47. Poolman RW, Abouali JAK, Conter HJ, Bhandari M. Overlapping systematic reviews of anterior cruciate ligament reconstruction comparing hamstring autograft with bone-patellar tendon-bone autograft: Why are they different? *Journal of Bone and Joint Surgery - Series A*. 2007;89:1542–52.
48. Hernandez A V, Walker E, Ioannidis JPA, Kattan MW. Challenges in meta-analysis of randomized clinical trials for rare harmful cardiovascular events: the case of rosiglitazone. *Am Heart J*. 2008;156(1):23–30.

49. Thorlund K, Druyts E, Antonio Aviña-Zubieta J, Wu P, Mills EJ. Why the findings of published multiple treatment comparison meta-analyses of biologic treatments for rheumatoid arthritis are different: An overview of recurrent methodological shortcomings. *Ann Rheum Dis*. 2013;72(9):1524–35.

50. Minozzi S, Davoli M, Bargagli AM, Amato L, Vecchi S, Perucci CA. An overview of systematic reviews on cannabis and psychosis: Discussing apparently conflicting results. *Drug and Alcohol Review*. 2010;29: 304–17.

51. Thompson R, Bandera E, Burley V, Cade J, Forman D, Freudenheim J, et al. Reproducibility of systematic literature reviews on food, nutrition, physical activity and endometrial cancer. *Public Health Nutr*. 2008;11(10):1006–14.

52. Chang SM, Carey TS, Kato EU, Guise J-M, Sanders GD. Identifying Research Needs for Improving Health Care. *Ann Intern Med*. 2012;157(6):439.

53. Ebrahim S, Bance S, Athale A, Malachowski C, Ioannidis JPA. Meta-analyses with industry involvement are massively published and report no caveats for antidepressants. *J Clin Epidemiol*. 2016;70:155–63.

54. Vickers A, Goyal N, Harland R, Rees R. Do Certain Countries Produce Only Positive Results? A Systematic Review of Controlled Trials. *Control Clin Trials*. 1998;19:159–66.

55. Wu T, Li Y, Bian Z, Liu G, Moher D. Randomized trials published in some Chinese journals: How many are randomized? *Trials*. 2009;10(1):46.

56. Pan Z, Trikalinos TA, Kavvoura FK, Lau J, Ioannidis JPA. Local literature bias in genetic epidemiology: An empirical evaluation of the Chinese literature. *PLoS Med*. 2005;2(12):1309–17.

57. Salanti G, Giovane C Del, Chaimani A, Caldwell DM, Higgins JPT, Tu Y-K. Evaluating the Quality of Evidence from a Network Meta- Analysis. *PLoS*

One. 2014;9(7):e99682.

58. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol*. 1997;50(6):683–91.

59. Higgins JPT, Whitehead A. Borrowing strength from external trials in a meta-analysis. In: Statistics in Medicine. 1996; 2733–49.

60. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med*. 2004;23(20):3105–24.

61. Bafeta A, Trinquart L, Seror R, Ravaud P. Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review. *BMJ*. 2013;347:f3675.

62. Bafeta A, Trinquart L, Seror R, Ravaud P. Reporting of results from network meta-analyses: Methodological systematic review. *BMJ (Online)*. 2014;348:g1741

63. Chen Y long, Yang K hu. Avoidable waste in the production and reporting of evidence.*The Lancet*. 2009;374:786.

64. Song F, Parekh S, Hooper L, Loke YK, Ryder J, Sutton AJ, et al. Dissemination and publication of research findings: An updated review of related biases. *Health Technol Assess (Rockv)*. 2010;14(8):1–220.

65. Scherer RW, Langenberg P, Von Elm E. Full publication of results initially presented in abstracts. Cochrane Database of Systematic Reviews. John Wiley & Sons, Ltd; 2007.

66. Dwan K, Gamble C, Kolamunnage-Dona R, Mohammed S, Powell C, Williamson PR. Assessing the potential for outcome reporting bias in a review: A tutorial. *Trials*. 2010;11(1):52.

67. Dickersin K, Min Y-I, Meinert CL. Factors Influencing Publication of Research Results. *JAMA*. 1992;267(3):374.

68. Ioannidis JP, Trikalinos TA. An exploratory test for an excess of significant findings. *Clin Trials*. 2007;4(3):245–53.

69. Ioannidis JPA. Meta-research: The art of getting it wrong. *Res Synth Methods*. 2010;1(3–4):169–84.

70. Ioannidis JPA, Munafò MR, Fusar-Poli P, Nosek BA, David SP. Publication and other reporting biases in cognitive sciences: Detection, prevalence, and prevention. *Trends in Cognitive Sciences*. 2014;18:235–41.

71. Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan AW, Cronin E, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS ONE*. 2008;3(8):e3081.

72. Ioannidis JPA. Clarifications on the application and interpretation of the test for excess significance and its extensions. *Journal of Mathematical Psychology*. 2013;57184–7.

73. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy. *N Engl J Med*. 2008;358(3):252–60.

74. Rothstein HR, Sutton AJ, Borenstein M. Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments. John Wiley & Sons; 2006.

75. Song F, Hooper L, Loke YK. Publication bias: what is it? How do we measure it? How do we avoid it? *Open Access J Clin Trials*. 2013;5:71–81.

76. Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ*. 2010;340(7747):c365–c365.

77. Hahn S, Williamson PR, Hutton JL, Garner P, Victor Flynn E. Assessing the potential for bias in meta-analysis due to selective reporting of subgroup analyses within studies. In: Statistics in Medicine 2000;3325–36.

78. Moreno SG, Sutton AJ, Turner EH, Abrams KR, Cooper NJ, Palmer TM, et al. Novel methods to deal with publication biases: Secondary analysis of antidepressant trials in the FDA trial registry database and related journal publications. *BMJ*. 2009;339(7719):494–7.

79. Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ*. 1997;315(7109):640–5.

80. Loannidis JPA. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *J Am Med Assoc*. 1998;279(4):281–6.

81. Reyes MM, Panza KE, Martin A, Bloch MH. Time-lag bias in trials of pediatric antidepressants: A systematic review and meta-analysis. *J Am Acad Child Adolesc Psychiatry*. 2011;50(1):63–72.

82. Ioannidis JP, Trikalinos TA. An exploratory test for an excess of significant findings. *Clin Trials*. 2007;4(3):245–53.

83. Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: Review of publications and survey of authors. *British Medical Journal*. 2005;330:753–6.

84. Chan AW, Krleža-Jerić K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ*. 2004;171(7):735–40.

85. Chan AW, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: Comparison

of protocols to published articles. *Journal of the American Medical Association*. 2004;291:2457–65.

86. Contopoulos-Ioannidis DG, Alexiou GA, Gouvias TC, Ioannidis JPA. An empirical evaluation of multifarious outcomes in pharmacogenetics: Beta-2 adrenoceptor gene polymorphisms in asthma treatment. *Pharmacogenet Genomics*. 2006;16(10):705–11.

87. Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of Registered and Published Primary Outcomes in Randomized Controlled Trials. *JAMA*. 2009;302(9):977.

88. ter Riet G, Korevaar DA, Leenaars M, Sterk PJ, van Noorden CJF, Bouter LM, et al. Publication Bias in Laboratory Animal Research: A Survey on Magnitude, Drivers, Consequences and Potential Solutions. *PLoS One.* 2012;7(9):e43404.

89. Perel P, Roberts I, Sena E, Wheble P, Briscoe C, Sandercock P, et al. Comparison of treatment effects between animal experiments and clinical trials: systematic review. *BMJ*. 2007;334(7586):197.

90. Tsilidis KK, Panagiotou OA, Sena ES, Aretouli E, Evangelou E, Howells DW, et al. Evaluation of Excess Significance Bias in Animal Studies of Neurological Diseases. *PLoS Biol*. 2013;11(7): e1001609.

91. Howells DW, Sena ES, Macleod MR. Bringing rigour to translational medicine. *Nature Reviews Neurology*. 2014;10: 37–43.

92. Ioannidis JPA, Trikalinos TA. Early extreme contradictory estimates may appear in published research: The Proteus phenomenon in molecular genetics research and randomized trials. *J Clin Epidemiol*. 2005;58(6):543–9.

93. Panagiotou OA, Willer CJ, Hirschhorn JN, Ioannidis JPA. The Power of

Meta-Analysis in Genome Wide Association Studies NIH Public Access. *Annu Rev Genomics Hum Genet.* 2013;14:441–65.

94. Lau J, Ioannidis JP, Schmid CH. Quantitative synthesis in systematic reviews. *Ann Intern Med.* 1997;127(9):820–6.

95. Rudmik LR, Walen SG, Dixon E, Dort J. Evaluation of meta-analyses in the otolaryngological literature. *Otolaryngology - Head and Neck Surgery*. 2008;139(2):187–94.

96. Munn Z, MClinSc SM, Lisy K, Riitano D, Tufanaru C. Methodological guidance for systematic reviews of observational epidemiological studies reporting prevalence and cumulative incidence data. *Int J Evid Based Healthc*. 2015;13(3):147–53.

97. Cook DJ, Sackett DL, Spitzer WO. Methodologic guidelines for systematic reviews of randomized control trials in health care from the potsdam consultation on meta-analysis. *J Clin Epidemiol*. 1995;48(1):167–71.

98. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Br J Surg*. 2000;87(11):1448–54.

99. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the Quality of Reporting of Randomized Controlled Trials. *JAMA*. 1996;276(8):637.

100. Moher D, Schulz KF, Altman D. The consort statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *Altern Ther Health Med*. 2002;8(3):96–100.

101. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group

randomized trials. *BMC Med Res Methodol*. 2001;1(1):2.

102. Moher D, Liberati A, Tetzlaff J, Altman DG, Group TP. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement (Reprinted from Annals of Internal Medicine). *Phys Ther*. 2009 Jul;89(9):873–80.

103. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA*. 2000;283(15):2008–12.

104. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol*. 1991;44(11):1271–8.

105. Oxman AD, Guyatt GH, Singer J, Goldsmith CH, Hutchison BG, Milner RA, et al. Agreement among reviewers of review articles. *J Clin Epidemiol*. 1991;44(1):91–8.

106. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-Analyses of Randomized Controlled Trials. *N Engl J Med*. 1987;316(8):450–5.

107. Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). *PLoS ONE*. 2007;2(12):e1350.

108. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*. 2007;7(1):10.

109. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the

methodological quality of systematic reviews. *J Clin Epidemiol*. 2009;62(10):1013–20.

110. Dixon E, Hameed M, Sutherland F, Cook DJ, Doig C. Evaluating meta-analyses in the general surgical literature: A critical appraisal. *Annals of Surgery*. 2005;241:450–9.

111. Delaney A, Bagshaw SM, Ferland A, Manns B, Laupland KB, Doig CJ. A systematic evaluation of the quality of meta-analyses in the critical care literature. *Crit care*. 2005;9(5):R575–82.

112. Bhandari M, Morrow F, Kulkarni A V, Tornetta P. Meta-analyses in orthopaedic surgery. A systematic review of their methodologies. *J Bone Jt Surg*. 2001;83–A(1):15–24.

113. Jadad AR, Moher M, Browman GP, Booker L, Sigouin C, Fuentes M, et al. Systematic reviews and meta-analyses on treatment of asthma: critical evaluation. *BMJ*. 2000;320:537–40.

114. Jadad AR, McQuay HJ. Meta-analyses to evaluate analgesic interventions: A systematic qualitative review of their methodology. *J Clin Epidemiol*. 1996;49(2):235–43.

115. Dijkman BG, Abouali JA, Kooistra BW, Conter HJ, Poolman RW, Kulkarni A V, et al. Twenty years of meta-analyses in orthopaedic surgery: Has quality kept up with quantity? *Journal of Bone and Joint Surgery - Series A*. 2010;92:48–57.

116. Wasiak J, Shen AY, Ware R, O'Donohoe TJ, Faggion CM. Methodological quality and reporting of systematic reviews in hand and wrist pathology. *Journal of Hand Surgery: European Volume*. 2017;42(8):852–6.

117. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias.

Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*. 1995;273(5):408–12.

118. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet*. 1998;352(9128):609–13.

119. Wee B, Hadley G, Derry S. How useful are systematic reviews for informing palliative care practice ? Survey of 25 Cochrane systematic reviews. *BMC Palliat Care.* 2008;6:6–11.

120. Ezzo J, Bausell B, Moerman DE, Berman B, Hadhazy V. Reviewing the reviews. How strong is the evidence? How clear are the conclusions? *Int J Technol Assess Health Care*. 2001;17(4):457–66.

121. Bader J, Ismail A. Survey of systematic reviews in dentistry. J*ournal of the American Dental Association*. 2004;135:464–73.

122. Helena A, Versiani V, Cabrera A, Ii M, Stella M, Iii P. Mapping of the evidence from systematic reviews of the Cochrane Collaboration for decision-making within physiotherapy Mapeamento das evidências de revisões sistemáticas da Colaboração Cochrane. *Sao Paulo Med J*. 2013;131(1):39–45.

123. El Dib RP, Atallah ÁN, Andriolo RB. Mapping the Cochrane evidence for decision making in health care. *J Eval Clin Pract*. 2007;13(4):689–92.

124. Shanks N, Greek R, Greek J. Are animal models predictive for humans? *Philosophy, Ethics, and Humanities in Medicine.* 2009; 4(1):2.

125. Botting JH, Morrison AR. Animal research is vital to medicine. *Sci Am*. 1997;276(2):83–5.

126. Lemon R, Dunnett SB. Surveying the literature from animal experiments. Critical reviews may be helpful - not systematic ones. *BMJ*. 2005;330 (7498):

977–8.

127. Greaves P, Williams A, Eve M. First dose of potential new medicines to humans: How animals help. *Nature Reviews Drug Discovery*. 2004; 3:226–36.

128. Pound P. Where is the evidence that animal research benefits humans? *BMJ*. 2004;328(7438):514–7.

129. Hartung T. Opinion versus evidence for the need to move away from animal testing. *ALTEX*. 2017;34(2):193–200.

130. Garattini S, Grignaschi G. Animal testing is still the best way to find new treatments for patients. *Eur J Intern Med*. 2017;39:32–5.

131. Hackam DG. Translating animal research into clinical benefit. *British Medical Journal*. 2007;334(7586):163–4.

132. Akhtar AZ, Pippin JJ, Sandusky CB. Animal models in spinal cord injury: A review. *Reviews in the Neurosciences*. 2008;19:47–60.

133. Hackam DG, Redelmeier DA. Translation of Research Evidence From Animals to Humans. *JAMA*. 2006;296(14):1727.

134. Garber K. Realistic rodents? Debate grows over new mouse models of cancer. *J Natl Cancer Inst*. 2006;98(17):1176–8.

135. Sausville EA, Burger AM. Contributions of human tumor xenografts to anticancer drug development. *American Association for Cancer Research*. 2006;66:3351–4.

136. Voskoglou-Nomikos T, Pater JL, Seymour L. Clinical predictive value of the in vitro cell line, human xenograft, and mouse allograft preclinical cancer models. *Clinical Cancer Research*. 2003: 9(11):4227-39.

137. Terszowski G, Müller SM, Bleul CC, Blum C, Schirmbeck R, Reimann J, et al. Evidence for a functional second thymus in mice. *Science*.

2006;312(5771):284–7.

138. Cabrera O, Berman DM, Kenyon NS, Ricordi C, Berggren P-O, Caicedo A. The unique cytoarchitecture of human pancreatic islets has implications for islet cell function. *Proc Natl Acad Sci*. 2006;103(7):2334–9.

139. Beauchamp K, Mutlak H, Smith WR, Shohami E, Stahel PF. Pharmacology of Traumatic Brain Injury: Where Is the "Golden Bullet"? *Mol Med*. 2008;14(11–12):731–40.

140. Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? *Nat Rev Drug Discov*. 2004;3(8):711–6.

141. Khanna I. Drug discovery in pharmaceutical industry: Productivity challenges and trends. Drug Discovery Today. *Elsevier Current Trends*. 2012;17: 17(19-20):1088-102.

142. Roberts I, Kwan I, Evans P, Haig S. Does animal experimentation inform human healthcare? Observations from a systematic review of international animal experiments on fluid resuscitation. *BMJ*. 2002;324(7335):474–6.

143. Hawkes N. Poor quality animal studies cause clinical trials to follow false leads. *BMJ*. 2015;351:h5453.

144. Bebarta V, Luyten D, Heard K. Emergency medicine animal research: does use of randomization and blinding affect the results? *Acad Emerg Med*. 2003;10(6):684–7.

145. Kilkenny C, Parsons N, Kadyszewski E, Festing MFW, Cuthill IC, Fry D, et al. Survey of the Quality of Experimental Design, Statistical Analysis and Reporting of Research Using Animals. McLeod M, editor. *PLoS One*. 2009;4(11):e7824.

146. Macleod MR, O'Collins T, Horky LL, Howells DW, Donnan GA. Systematic

review and meta-analysis of the efficacy of melatonin in experimental stroke. *J Pineal Res*. 2005;38(1):35–41.

147. Macleod MR, O'Collins T, Howells DW, Donnan GA. Pooling of animal experimental data reveals influence of study design and publication bias. *Stroke*. 2004;35(5):1203-8.

148. O'Collins VE, Macleod MR, Donnan GA, Horky LL, Van Der Worp BH, Howells DW. 1,026 Experimental treatments in acute stroke. *Ann Neurol*. 2006;59(3):467–77.

149. Sena ES, Bart van der Worp H, Bath PMW, Howells DW, Macleod MR. Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. *PLoS Biol*. 2010;8(3):e1000344.

150. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. The ARRIVE Guidelines Checklist Animal Research : Reporting In Vivo Experiments. *Br J Pharmacol*. 2010;8:8–9.

151. Chen Z, Cheng K, Walton Z, Wang Y, Ebi H, Shimamura T, et al. A murine lung cancer co-clinical trial identifies genetic modifiers of therapeutic response. *Nature*. 2012;483(7391):613–7.

152. Petitti DB. Perinatal epidemiology: Studying the effects of illness and medications during pregnancy. *Immunol Allergy Clin North Am*. 2000;20(4):673–85.

153. Kudlová E. Life cycle approach to child and adolescent health. *Cent Eur J Public Health*. 2004;12(3):166–70.

154. Barker DJ. Fetal origins of coronary heart disease. *Br Med J*. 1995;311(3):171–4.

155. Barker DJP, Godfrey KM, Gluckman PD, Harding JE, Owens JA, Robinson

JS. Fetal nutrition and cardiovascular disease in adult life. *Lancet*. 1993;341(8850):938–41.

156. Roseboom TJ, van der Meulen JH, Osmond C, Barker DJ, Ravelli AC, Schroeder-Tanka JM, et al. Coronary heart disease after prenatal exposure to the Dutch famine, 1944-45. *Heart*. 2000;84(6):595–8.

157. Goldenberg RL, Culhane JF. Low birth weight in the United States. *American Journal of Clinical Nutrition.* 2007;85(2):584S-90S

158. Fall CH. Fetal malnutrition and long-term outcomes. *Nestle Nutrition Institute Workshop Series*. 2013;74: 11–25.

159. Bracken MB. Perinatal Epidemiology. New York: Oxford University Press; 1984.

160. Jackson DJ, Lang JM, Ganiats TG. Epidemiological issues in perinatal outcomes research. *Paediatric and Perinatal Epidemiology*. 1999;13: 392–404.

161. Thompson C. Fortuitous phenomena: On complexity, pragmatic randomised controlled trials, and knowledge for evidence-based practice. *Worldviews Evidence-Based Nurs.* 2004;1(1):9–17.

162. Walker W. The strengths and weaknesses of research designs involving quantitative measures. *J Res Nurs*. 2005;10(5):571–82.

163. Schneider, Z, Whitehead D. Nursing and Midwifery Research: Methods and Appraisal for Evidence-Based Practice -- Searching and reviewing the research literature. *Elsevier Australia*. 2013.

164. Zlowodzki M, Jönsson A, Bhandari M. Common pitfalls in the conduct of clinical research. *Medical Principles and Practice*. 2005;15: 1–8.

165. Miettinen O. Confounding and effect-modification. 1974. *Am J Epidemiol*.

141(12):1113–6.

166. Rothman JK. Modern epidemiology. Boston: Little Brown and Company; 1986.

167. Tilaki KH. Methodological issues of confounding in analytical epidemiologic studies. *Casp J Intern Med*. 2012;3(3):488–95.

168. Lieberman E, Lang JM, Heffner LJ, Cohen A. Assessing the role of case mix in cesarean delivery rates. *Obstet Gynecol*. 1998;92(1):1–7.

169. Souza JP, Gülmezoglu AM, Lumbiganon P, Laopaiboon M, Carroli G, Fawole B, et al. Caesarean section without medical indications is associated with an increased risk of adverse short-term maternal outcomes: The 2004-2008 WHO Global Survey on Maternal and Perinatal Health. *BMC Med*. 2010;8:71.

170. Silver RM, Landon MB, Rouse DJ, Leveno KJ, Spong CY, Thom EA, et al. Maternal morbidity associated with multiple repeat cesarean deliveries. *Obstet Gynecol*. 2006;107(6):1226–32.

171. Bloomberg MA, Jordan HS, Angel KO, Bailit MH, Goonan KJ, Straus J. Development of clinical indicators for performance measurement and improvement: an HMO/purchaser collaborative effort. J*t Comm J Qual Improv*. 1993;19(12):586–95.

172. Krivenko CA, Chodroff C. The analysis of clinical outcomes: getting started in benchmarking. *Jt Comm J Qual Improv*. 1994;20(5):260–6.

173. Romero R, Kuivaniemi H, Tromp G, Olson JM. The design, execution, and interpretation of genetic association studies to decipher complex diseases. *Am J Obstet Gynecol*. 2002;187(5):1299–312.

174. Cnattingius S, Reilly M, Pawitan Y, Lichtenstein P. Maternal and fetal genetic factors account for most of familial aggregation of preeclampsia: A

population-based swedish cohort study. *Am J Med Genet*. 2004;130 A(4):365–71.

175. Williams PJ, Broughton Pipkin F. The genetics of pre-eclampsia and other hypertensive disorders of pregnancy. *Best Pract Res Clin Obstet Gynaecol*. 2011;25(4):405–17.

176. Sutherland A, Cooper DW, Howie PW, Liston WA, MacGillivray I. The incidence of severe pre-eclampsia amongst mothers and mothers-in-law of pre-eclamptics and controls. *BJOG An Int J Obstet Gynaecol*;88(8):785–91.

177. Ros HS, Lichtenstein P, Lipworth L, Cnattingius S. Genetic effects on the liability of developing pre-eclampsia and gestational hypertension. *Am J Med Genet*. 2000;91(4):256–60.

178. Treloar SA, Cooper DW, Brennecke SP, Grehan MM, Martin NG. An Australian twin study of the genetic basis of preeclampsia and eclampsia. *Am J Obstet Gynecol.* 2001;184(3):374–81.

179. Nilsson E, Ros HS, Cnattingius S, Lichtenstein P. The importance of genetic and environmental effects for pre-eclampsia and gestational hypertension: A family study. *BJOG An Int J Obstet Gynaecol*. 2004;111(3):200–6.

180. Boyd HA, Tahir H, Wohlfahrt J, Melbye M. Associations of personal and family preeclampsia historywith the risk of early-, intermediate- and late-onset preeclampsia. *Am J Epidemiol*. 2013;178(11):1611–9.

181. Sugawara J, Oe Y, Wagata M. Genetic Background of Preeclampsia. In Springer, Singapore; 2018; p. 29–43.

182. Staines-Urias E, Paez MC, Doyle P, Dudbridge F, Serrano NC, Ioannidis JP, et al. Genetic association studies in pre-eclampsia: Systematic meta-analyses and field synopsis. *Int J Epidemiol*. 2012;41(6):1764–75.

183. Mütze S, Rudnik-Schöneborn S, Zerres K, Rath W. Genes and the preeclampsia syndrome. *Journal of Perinatal Medicine*. 2008;36: 38–58.

184. Chappell S, Morgan L. Searching for genetic clues to the causes of pre-eclampsia. *Clin Sci*. 2006;110(4):443–58.

185. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009;461(7265):747–53.

186. Uzun A, Triche EW, Schuster J, Dewan AT, Padbury JF. dbPEC: a comprehensive literature-based database for preeclampsia related genes and phenotypes. *Database*. 2016:baw006.

187. Arngrimsson R, Bjornsson S, Geirsson RT, Bjornsson H, Walker JJ, Snaedal G. Genetic and familial predisposition to eclampsia and pre-eclampsia in a defined population. *BJOG An Int J Obstet Gynaecol*. 1990;97(9):762–9.

188. Catalano P, deMouzon SH. Maternal obesity and metabolic risk to the offspring: why lifestyle interventions may have not achieved the desired outcomes. *Int J Obes*. 2015;39(4):642–9.

189. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science*. 1996;273(5281):1516–7.

190. Colhoun HM, McKeigue PM, Smith GD. Problems of reporting genetic associations with complex outcomes. *Lancet*. 2003;361: 865–72.

191. Roberts JM, Cooper DW. Pathogenesis and genetics of pre-eclampsia. *Lancet.* 2001;357(9249):53–6.

192. Maher BS. Polygenic Scores in Epidemiology: Risk Prediction, Etiology, and Clinical Utility. *Curr Epidemiol Reports*. 2015;2(4):239–44.

193. Zhao L, Triche EW, Walsh KM, Bracken MB, Saftlas AF, Hoh J, et al.

Genome-wide association study identifies a maternal copy-number deletion in PSG11 enriched among preeclampsia patients. *BMC Pregnancy Childbirth*. 2012;12(1):61.

194. Johnson MP, Brennecke SP, East CE, Göring HHH, Kent JW, Dyer TD, et al. Genome-wide association scan identifies a risk locus for preeclampsia on 2q14, near the inhibin, beta B gene. *PLoS One*. 2012;7(3):e33666.

195. Zhao L, Bracken MB, DeWan AT. Genome-Wide Association Study of Pre-Eclampsia Detects Novel Maternal Single Nucleotide Polymorphisms and Copy-Number Variants in Subsets of the Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study Cohort. *Ann Hum Genet*. 2013;77(4):277–87.

196. Kusmierska-Urban K, Rytlewski K, Huras H, Wybranska I. Association of single nucleotide polymorphism rs7579169 with hypertension disorders during pregnancy and perinatal outcome. *Neuro Endocrinol Lett*. 2015;36(3):282–7.

197. Guo LF, Wang ZH, Wang YF. Common variant rs7579169 is associated with preeclampsia in Han Chinese women. *Genet Mol Res*. 2016;15(2).

198. Lewis CM, Vassos E. Prospects for using risk scores in polygenic medicine. Genome Medicine. *BioMed Central*. 2017;9(1):96.

199. Smith CJ, Saftlas AF, Spracklen CN, Triche EW, Bjonnes A, Keating B, et al. Genetic risk score for essential hypertension and risk of preeclampsia. *Am J Hypertens*. 2016;29(1):17–24.

200. Spracklen CN, Saftlas AF, Triche EW, Bjonnes A, Keating B, Saxena R, et al. Genetic predisposition to dyslipidemia and risk of preeclampsia. *Am J Hypertens*. 2015;28(7):915–23.

201. Spracklen CN, Smith CJ, Saftlas AF, Triche EW, Bjonnes A, Keating BJ, et

al. Genetic predisposition to elevated levels of C-reactive protein is associated with a decreased risk for preeclampsia. *Hypertens Pregnancy*. 2017;36(1):30–5.

202. Trevethan R. Sensitivity, Specificity, and Predictive Values: Foundations, Pliabilities, and Pitfalls in Research and Practice. *Front public Heal*. 2017;5:307.

203. Conde-Agudelo A, Villar J, Lindheimer M. World Health Organization Systematic Review of Screening Tests for Preeclampsia. *Obstet Gynecol*. 2004;104(6):1367–91.

204. Nicolaides KH. Turning the pyramid of prenatal care. *Fetal Diagn Ther*. 2011;29(3):183–96.

205. Nicolaides KH. A model for a new pyramid of prenatal care based on the 11 to 13 weeks' assessment. *Prenatal Diagnosis*. 2011;31(1):3–6.

206. McDonald SD, Malinowski A, Zhou Q, Yusuf S, Devereaux PJ. Cardiovascular sequelae of preeclampsia/eclampsia: A systematic review and meta-analyses. *Am Heart J*. 2008;156(5):918–30.

207. Duckitt K, Harrington D. Risk factors for pre-eclampsia at antenatal booking: systematic review of controlled studies. *BMJ*. 2005;330(7491):565.

208. North RA, McCowan LM, Dekker GA, Poston L, Chan EH, Stewart AW, et al. Clinical risk prediction for pre-eclampsia in nulliparous women: development of model in international prospective cohort. *Bmj*. 2011;342:d1875.

209. Wallenburg HCS. Prevention of pre-eclampsia: Status and perspectives 2000. *European Journal of Obstetrics Gynecology and Reproductive Biology*. 2001;94(1):13-22.

210. Chaiworapongsa T, Romero R, Korzeniewski SJ, Kusanovic JP, Soto E, Lam J, et al. Maternal plasma concentrations of angiogenic/antiangiogenic factors in the third trimester of pregnancy to identify the patient at risk for stillbirth at or near term and severe late preeclampsia. *Am J Obstet Gynecol*. 2013;208(4):287.e1-287.e15.

211. O'Brien WF. Predicting preeclampsia. *Obstet Gynecol*. 1990;75(3):445–52.

212. Wright D, Akolekar R, Syngelaki A, Poon LCY, Nicolaides KH. A competing risks model in early screening for preeclampsia. *Fetal Diagn Ther*. 2012;32(3):171–8.

213. Akolekar R, Syngelaki A, Poon L, Wright D, Nicolaides KH. Competing Risks Model in Early Screening for Preeclampsia by Biophysical and Biochemical Markers. *Fetal Diagn Ther*. 2013;33(1):8–15.

214. Anderson UD, Gram M, Åkerström B, Hansson SR. First Trimester Prediction of Preeclampsia. *Current Hypertension Reports*. 2015;17: 74.

215. Anderson UD, Olsson MG, Kristensen KH, Åkerström B, Hansson SR. Review: Biochemical markers to predict preeclampsia. *Placenta*. 2012;33:S42-7.2–7.

216. Cetin I, Huppertz B, Burton G, Cuckle H, Gonen R, Lapaire O, et al. Pregenesys pre-eclampsia markers consensus meeting: What do we require from markers, risk assessment and model systems to tailor preventive strategies? *Placenta*. 2011;32:S4–16.

217. Cnossen JS, Riet G Ter, Mol BW, Van Der Post JA, Leeflang MM, Meads CA, et al. Are tests for predicting pre-eclampsia good enough to make screening viable? A review of reviews and critical appraisal. *Acta Obstetricia et Gynecologica Scandinavica*. 2009;88:758–65.

218. Kuc S, Wortelboer EJ, Van Rijn BB, Franx A, Visser GHA, Schielen PCJI. Evaluation of 7 serum biomarkers and uterine artery doppler ultrasound for first-trimester prediction of preeclampsia: A systematic review. *Obstetrical and Gynecological Survey*. 2011;66:225–39.

219. Meads CA, Cnossen JS, Meher S, Juarez-Garcia A, ter Riet G, Duley L, et al. Methods of prediction and prevention of pre-eclampsia: systematic reviews of accuracy and effectiveness literature with economic modelling. *Health Technol Assess.* 2008;12(6): 1-270.

220. Waldo Sepulveda, Amy E. Wong, Alexandra Casasbuenas AS and JLA. Congenital diaphragmatic hernia in a first-trimester ultrasound aneuploidy screening program. *Prenat Diagn*. 2008;28(6):531–4.

221. Anderson UD, Olsson MG, Rutardóttir S, Centlow M, Kristensen KH, Isberg PE, et al. Fetal hemoglobin and α1-microglobulin as first- and early second-trimester predictive biomarkers for preeclampsia. *Am J Obstet Gynecol*. 2011;204(6):520.e1-520.e5.

222. Hui D, Okun N, Murphy K, Kingdom J, Uleryk E, Shah PS. Combinations of Maternal Serum Markers to Predict Preeclampsia, Small for Gestational Age, and Stillbirth: A Systematic Review. *J Obstet Gynaecol Canada*;34(2):142–53.

223. Wu P, Van Den Berg C, Alfirevic Z, O'brien S, Röthlisberger M, Baker PN, et al. Early pregnancy biomarkers in pre-eclampsia: A systematic review and meta-analysis. *International Journal of Molecular Sciences*. 2015;16:23035–56.

224. Giguere Y, Charland M, Bujold E, Bernard N, Grenier S, Rousseau F, et al. Combining Biochemical and Ultrasonographic Markers in Predicting

Preeclampsia: A Systematic Review. *Clin Chem*. 2010;56(3):361–75.

225. HAPO Study Cooperative Research Group, Metzger BE, Lowe LP, Dyer AR, Trimble ER, Chaovarindr U, et al. Hyperglycemia and Adverse Pregnancy Outcomes. *N Engl J Med*. 2008;358(19):1991–2002.

226. Crowther CA, Hiller JE, Moss JR, McPhee AJ, Jeffries WS, Robinson JS. Effect of Treatment of Gestational Diabetes Mellitus on Pregnancy Outcomes. *N Engl J Med*. 2005;352(24):2477–86.

227. Landon MB, Spong CY, Thom E, Carpenter MW, Ramin SM, Casey B, et al. A Multicenter, Randomized Trial of Treatment for Mild Gestational Diabetes. *N Engl J Med*. 2009;361(14):1339–48.

228. Griffin ME, Coffey M, Johnson H, Scanlon P, Foley M, Stronge J, et al. Universal vs. risk factor-based screening for gestational diabetes mellitus: detection rates, gestation at diagnosis and outcome. *Diabet Med*. 2000;17(1):26–32.

229. Teh WT, Teede HJ, Paul E, Harrison CL, Wallace EM, Allan C. Risk factors for gestational diabetes mellitus: Implications for the application of screening guidelines. *Aust New Zeal J Obstet Gynaecol*. 2011;51(1):26–30.

230. Santos-Ayarzagoitia M, Salinas-Martínez AM, Villarreal-Pérez JZ. Gestational diabetes: Validity of ADA and WHO diagnostic criteria using NDDG as the reference test. *Diabetes Res Clin Pract*. 2006;74(3):322–8.

231. American Diabetes Association. Gestational diabetes mellitus. *Diabetes Care*. 2003;26:S103-5.

232. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2014;37(1):S81-90.

233. Alberico S, Strazzanti C, De Santo D, De Seta F, Lenardon P, Bernardon M, et

al. Gestational diabetes: Universal or selective screening? *J Matern Neonatal Med*. 2004;16(6):331–7.

234. Cosson E, Benchimol M, Carbillon L, Pharisien I, Pariès J, Valensi P, et al. Universal rather than selective screening for gestational diabetes mellitus may improve fetal outcomes. *Diabetes Metab*. 2006;32(2):140–6.

235. Dahanayaka N, Agampodi S, Ranasinghe O, Jayaweera P, Wickramasinghe W, Adhikari A, et al. Inadequacy of the risk factor based approach to detect gestational diabetes mellitus. *Ceylon Med J*. 2012;57(1):5.

236. Sattar N, Wannamethee SG, Forouhi NG. Novel biochemical risk factors for type 2 diabetes: Pathogenic insights or prediction possibilities? *Diabetologia*. 2008;51:926–40.

237. Georgiou HM, Lappas M, Georgiou GM, Marita A, Bryant VJ, Hiscock R, et al. Screening for biomarkers predictive of gestational diabetes mellitus. *Acta Diabetol*. 2008;45(3):157–65.

238. Abell SK, De Courten B, Boyle JA, Teede HJ. Inflammatory and other biomarkers: Role in pathophysiology and prediction of gestational diabetes mellitus. *International Journal of Molecular Sciences*. 2015;16:13442–73.

239. Bao W, Baecker A, Song Y, Kiely M, Liu S, Zhang C. Adipokine levels during the first or early second trimester of pregnancy and subsequent risk of gestational diabetes mellitus: A systematic review. *Metabolism*. 2015;64(6):756–64.

240. Xu J, Zhao YH, Chen YP, Yuan XL, Wang J, Zhu H, et al. Maternal circulating concentrations of tumor necrosis factor-alpha, leptin, and adiponectin in gestational diabetes mellitus: a systematic review and meta-analysis. *ScientificWorldJournal*. 2014;2014:926932.

241. Ryckman KK, Spracklen CN, Smith CJ, Robinson JG, Saftlas AF. Maternal lipid levels during pregnancy and gestational diabetes: A systematic review and meta-analysis. *BJOG: An International Journal of Obstetrics and Gynaecology*. 2015;22: 643–51.

242. Sweeting AN, Appelblom H, Ross GP, Wong J, Kouru H, Williams PF, et al. First trimester prediction of gestational diabetes mellitus: A clinical model based on maternal demographic parameters. *Diabetes Res Clin Pract*. 2017;127:44–50.

243. Harrison CL, Lombard CB, East C, Boyle J, Teede HJ. Risk stratification in early pregnancy for women at increased risk of gestational diabetes. *Diabetes Res Clin Pract*. 2015;107(1):61–8.

244. Oostdam N, van Poppel MNM, Wouters MGAJ, van Mechelen W. Interventions for Preventing Gestational Diabetes Mellitus: A Systematic Review and Meta-Analysis. *J Women's Heal*. 2011;20(10):1551–63.

245. Thériault S, Forest JC, Massé J, Giguère Y. Validation of early risk-prediction models for gestational diabetes based on clinical characteristics. *Diabetes Res Clin Pract*. 2014;103(3):419–25.

246. Lamain – de Ruiter M, Kwee A, Naaktgeboren CA, Franx A, Moons KGM, Koster MPH. Prediction models for the risk of gestational diabetes: a systematic review. *Diagnostic Progn Res*. 2017;1(1):3.

247. Teede HJ, Harrison CL, Teh WT, Paul E, Allan CA. Gestational diabetes: Development of an early risk prediction tool to facilitate opportunities for prevention. *Aust New Zeal J Obstet Gynaecol*. 2011;51(6):499–504.

248. Savvidou M, Nelson SM, Makgoba M, Messow CM, Sattar N, Nicolaides K. First-trimester prediction of gestational diabetes mellitus: Examining the

potential of combining maternal characteristics and laboratory measures. *Diabetes*. 2010;59(12):3017–22.

249. Nanda S, Savvidou M, Syngelaki A, Akolekar R, Nicolaides KH. Prediction of gestational diabetes mellitus by maternal factors and biomarkers at 11 to 13 weeks. *Prenat Diagn*. 2011;31(2):135–41.

250. Ferreira AFA, Rezende JC, Vaikousi E, Akolekar R, Nicolaides KH. Maternal serum visfatin at 11-13 weeks of gestation in gestational diabetes mellitus. *Clin Chem*. 2011;57(4):609–13.

251. Ioannidis JPA. Interpretation of tests of heterogeneity and bias in meta-analysis. *Journal of Evaluation in Clinical Practice*. 2008;14: 951–7.

252. Ioannidis JPA, Trikalinos TA. The appropriateness of asymmetry tests for publication bias in meta-analyses: A large survey. *CMAJ*. 2007;176(8):1091–6.

253. Lau J, Ioannidis JPA, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. *BMJ*. 2006;333(7568):597–600.

254. Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic Review of the Empirical Evidence of Study Publication Bias and Outcome Reporting Bias - An Updated Review. Boutron I, editor. *PLoS ONE*. 2013;8:e66844.

255. Hartling L, Chisholm A, Thomson D, Dryden DM. A Descriptive Analysis of Overviews of Reviews Published between 2000 and 2011. *PLoS One*. 2012;7(11):e49667

256. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: How will we ever keep up? *PLoS Med*. 2010;7(9):A11–2.

257. Becker LA, Oxman AD. Overviews of Reviews. In: Cochrane Handbook for Systematic Reviews of Interventions. Chichester, UK: John Wiley & Sons,

Ltd; 2008. 607–31.

258. Zhang W. I-01 Meta-Epidemiology: Building the Bridge From Research Evidence To Clinical Practice. *Osteoarthr Cartil*. 2010;18(2):S1.

259. Roever L B-ZG. Features of Meta-Epidemiology, Meta-Meta- Epidemiology and Network Meta-Epidemiology in Emergency Medicine. *Austin Emerg Med*. 2016;2(8):1044.

260. Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: Meta-epidemiological study. *BMJ*. 2008;336(7644):601–5.

261. Le Lorier J, Grégoire G. Meta-analysis and the meta-epidemiology of clinical research. Comments on paper by author of editorial were unwarranted. *BMJ*. 1998;316(7127):311–2.

262. Bae J-M. Epidemiology and Health Meta-epidemiology. *Epidemiol Health.* 2014; 36:e2014019.

263. Puljak L. If there is only one author or only one database was searched, a study should not be called a systematic review. *J Clin Epidemiol*. 2017;91:4–5.

264. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews. *Int J Evid Based Healthc*. 2015;13(3):132–40.

265. Tsilidis KK, Papatheodorou SI, Evangelou E, Ioannidis JPA. Evaluation of Excess Statistical Significance in Meta-analyses of 98 Biomarker Associations with Cancer Risk. *JNCI J Natl Cancer Inst*. 2012 Dec;104(24):1867–78.

266. Belbasis L, Bellou V, Evangelou E, Ioannidis JPA, Tzoulaki I. Environmental

risk factors and multiple sclerosis: An umbrella review of systematic reviews and meta-analyses. *Lancet Neurol*. 2015;14(3):263–73.

267. Bellou V, Belbasis L, Tzoulaki I, Evangelou E, Ioannidis JPA. Environmental risk factors and Parkinson's disease: An umbrella review of meta-analyses. *Parkinsonism Relat Disord.* 2016;23:1–9.

268. Belbasis L, Bellou V, Evangelou E. Environmental Risk Factors and Amyotrophic Lateral Sclerosis: An Umbrella Review and Critical Assessment of Current Evidence from Systematic Reviews and Meta-Analyses of Observational Studies. *Neuroepidemiology*. 2016;46(2):96–105.

269. Li X, Meng X, Timofeeva M, Tzoulaki I, Tsilidis KK, Ioannidis PA, et al. Serum uric acid levels and multiple health outcomes: umbrella review of evidence from observational studies, randomised controlled trials, and Mendelian randomisation studies. *BMJ*. 2017;357:j2376.

270. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7(3):177–88.

271. Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ*. 2011;342:d549.

272. Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Stat Soc Ser A Stat Soc*. 2009;172(1):137–59.

273. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002;21(11):1539–58.

274. Ioannidis JPA, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. *BMJ*. 2007;335(7626):914–6.

275. Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in

meta-analyses of randomised controlled trials. *BMJ*. 2011;343(7818):d4002.

276. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629–34.

277. Ioannidis JPA. Clarifications on the application and interpretation of the test for excess significance and its extensions. *Journal of Mathematical Psychology*. 2013;57:184–7.

278. Lubin JH, Gail MH. On power and sample size for studying features of the relative odds of disease. *Am J Epidemiol*. 1990;131(3):552–66.

279. Bellou V, Belbasis L, Tzoulaki I, Middleton LT, Ioannidis JPA, Evangelou E. Systematic evaluation of the associations between environmental risk factors and dementia: An umbrella review of systematic reviews and meta-analyses. *Alzheimer's Dement*. 2017;13(4):406–18.

280. Belbasis L, Köhler CA, Stefanis N, Stubbs B, van Os J, Vieta E, et al. Risk factors and peripheral biomarkers for schizophrenia spectrum disorders: an umbrella review of meta-analyses. *Acta Psychiatr Scand*. 2018;137(2):88–97.

281. Ioannidis JPA, Boffetta P, Little J, O'Brien TR, Uitterlinden AG, Vineis P, et al. Assessment of cumulative evidence on genetic associations: Interim guidelines. *Int J Epidemiol*. 2008;37(1):120–32.

282. Cooper H, Koenka AC. Unique challenges and opportunities when research syntheses are the principal elements of new integrative scholarship. *Am Psychol*. 2012;67(6):446–62

283. Smith V, Devane D, Begley CM, Clarke M. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Med Res Methodol*. 2011;11(1):15

284. Duley L. Pre-eclampsia, eclampsia, and hypertension. *BMJ Clin Evid*. 2011;2:

1402

285. Steegers EAP, Von Dadelszen P, Duvekot JJ, Pijnenborg R. Pre-eclampsia. *Lancet*. 2010;376(9741):631–44.

286. Chaiworapongsa T, Chaemsaithong P, Yeo L, Romero R. Pre-eclampsia part 1: current understanding of its pathophysiology. *Nat Rev Nephrol*. 2014;10(8):466–80.

287. Uzan J, Carbonnel M, Piconne O, Asmar R, Ayoubi JM. Pre-eclampsia: Pathophysiology, diagnosis, and management. *Vascular Health and Risk Management*. 2011;7:467–74.

288. Buurma AJ, Turner RJ, Driessen JHM, Mooyaart AL, Schoones JW, Bruijn JA, et al. Genetic variants in pre-eclampsia: A meta-analysis. *Hum Reprod Update.* 2013;19(3):289–303.

289. Sibai B, Dekker G, Kupferminc M. Pre-eclampsia. *Lancet*. 2005; 365(9461): 785–99.

290. Askie LM, Duley L, Henderson-Smart DJ, Stewart LA, PARIS Collaborative Group. Antiplatelet agents for prevention of pre-eclampsia: a meta-analysis of individual patient data. *Lancet*. 2007;369(9575):1791–8.

291. Duley L, Dj H, Meher S, Jf K, Duley L, Henderson-smart DJ, et al. Antiplatelet agents for preventing pre-eclampsia and its complications *Cochrane Database Syst*: 2007:CD004659.

292. Bujold E, Roberge S, Lacasse Y, Bureau M, Audibert F, Marcoux S, et al. Prevention of preeclampsia and intrauterine growth restriction with aspirin started in early pregnancy: A meta-analysis. *Obstetrics and Gynecology*. 2010;116:402–14.

293. Groeneveld E, Lambers MJ, Lambalk CB, Broeze KA, Haapsamo M, De

Sutter P, et al. Preconceptional low-dose aspirin for the prevention of hypertensive pregnancy complications and preterm delivery after IVF: A meta-analysis with individual patient data. *Hum Reprod*. 2013;28(6):1480–8.

294. Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Stat Soc A*. 2008;172(1):137–59.

295. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557–60.

296. Fan Y, Kang Y, Zhang M. A meta-analysis of copper level and risk of preeclampsia: evidence from 12 publications. *Biosci Rep*. 2016;36(4):e00370–e00370.

297. Song Q-Y, Luo W-P, Zhang C-X. High serum iron level is associated with an increased risk of hypertensive disorders during pregnancy: a meta-analysis of observational studies. *Nutr Res*. 2015;35(12):1060–9.

298. Cohen JM, Beddaoui M, Kramer MS, Platt RW, Basso O, Kahn SR. Maternal antioxidant levels in pregnancy and risk of preeclampsia and small for gestational age birth: A systematic review and meta-analysis. *PLoS One*. 2015;10(8):e0135192.

299. Liu HQ, Wang YH, Wang LL, Hao M. Predictive Value of Free β-hCG Multiple of the Median for Women with Preeclampsia. *Gynecol Obstet Invest*. 2016;81(2):137–47.

300. Ma Y, Shen X, Zhang D. The Relationship between Serum Zinc Level and Preeclampsia: A Meta-Analysis. *Nutrients*. 2015;7(9):7806–20.

301. Allen RE, Rogozinska E, Cleverly K, Aquilina J, Thangaratinam S. Abnormal blood biomarkers in early pregnancy are associated with preeclampsia: A meta-analysis. *Eur J Obstet Gynecol Reprod Biol*. 2014;182:194–201.

302. Yang Y, Su X, Xu W, Zhou R. Interleukin-18 and Interferon Gamma Levels in Preeclampsia: A Systematic Review and Meta-analysis. *Am J Reprod Immunol*. 2014;72(5):504–14.

303. Lashley EELO, Meuleman T, Claas FHJ. Beneficial or harmful effect of antipaternal human leukocyte antibodies on pregnancy outcome? A systematic review and meta-analysis. *American Journal of Reproductive Immunology*. 2013;70:87–103.

304. Dai B, Liu T, Zhang B, Zhang X, Wang Z. The polymorphism for endothelial nitric oxide synthase gene, the level of nitric oxide and the risk for pre-eclampsia: A meta-analysis. *Gene*. 2013;519(1):187–93.

305. Kleinrouweler CE, Wiegerinck MMJ, Ris-Stalpers C, Bossuyt PMM, Van Der Post JAM, Von Dadelszen P, et al. Accuracy of circulating placental growth factor, vascular endothelial growth factor, soluble fms-like tyrosine kinase 1 and soluble endoglin in the prediction of pre-eclampsia: A systematic review and meta-analysis. *BJOG: An International Journal of Obstetrics and Gynaecology*. 2012; 119:778–87.

306. Hausvater A, Giannone T, Sandoval YHG, Doonan RJ, Antonopoulos CN, Matsoukis IL, et al. The association between preeclampsia and arterial stiffness. *Journal of Hypertension*. 2012;31:17–33.

307. Do Prado AD, Piovesan DM, Staub HL, Horta BL. Association of anticardiolipin antibodies with preeclampsia: A systematic review and meta-analysis. *Obstetrics and Gynecology*. 2010;116:1433–43.

308. Clark P, Wu O. ABO(H) blood groups and pre-eclampsia. A systematic review and meta-analysis. *Thromb Haemost*. 2008;100(3):469–74.

309. Hu H, Ha S, Roth J, Kearney G, Talbott EO, Xu X. Ambient Air Pollution and

Hypertensive Disorders of Pregnancy: A Systematic Review and Meta-analysis. *Atmos Environ*. 2014;97(1994):336–45.

310. Pedersen M, Stayner L, Slama R, Sørensen M, Figueras F, Nieuwenhuijsen MJ, et al. Ambient air pollution and pregnancy-induced hypertensive disorders: A systematic review and meta-analysis. *Hypertension*. 2014;64(3):494–500.

311. Zeng F, Zhu S, Wong MC-S, Yang Z, Tang J, Li K, et al. Associations between nitric oxide synthase 3 gene polymorphisms and preeclampsia risk: a meta-analysis. *Sci Rep*. 2016;6(1):23407.

312. Zhang G, Zhao J, Yi J, Luan Y, Wang Q. Association Between Gene Polymorphisms on Chromosome 1 and Susceptibility to Pre-Eclampsia: An Updated Meta-Analysis. *Med Sci Monit*. 2016 ;22:2202–14.

313. Li Y, Zhu M, Hu R, Yan W. The effects of gene polymorphisms in angiotensin II receptors on pregnancy-induced hypertension and preeclampsia: A systematic review and meta-analysis. *Hypertens Pregnancy*. 2015;34(2):241–60.

314. Yang W, Zhu Z, Wan J, Ding Y. Evaluation of association of maternal IL-10 polymorphisms with risk of preeclampsia by A meta-analysis Literature selection. *J Cell Mol Med*. 2014;18(12):2466–77.

315. Wang X, Bai T, Liu S, Pan H, Wang B. Association between thrombophilia gene polymorphisms and preeclampsia: A meta-analysis. *PLoS One*. 2014;9(6):e100789

316. Li X, Luo YL, Zhang QH, Mao C, Wang XW, Liu S, et al. Methylenetetrahydrofolate reductase gene C677T, A1298C polymorphisms and pre-eclampsia risk: A meta-analysis. *Mol Biol Rep*. 2014;41(8):5435–48.

317. Gong LL, Liu H, Liu LH. Lack of association between matrix metalloproteinase-9 gene-1562C/T polymorphism and preeclampsia: A meta-analysis. *Hypertens Pregnancy*. 2014;33(4):389–94.

318. Li X, Shen L, Tan H. Polymorphisms and plasma level of transforming growth factor-beta 1 and risk for preeclampsia: A systematic review. *PLoS One*. 2014;9(5).

319. Cheng D, Hao Y, Zhou W, Ma Y. Vascular endothelial growth factor +936C/T, -634G/C, -2578C/A, and -1154G/A polymorphisms with risk of preeclampsia: A meta-analysis. *PLoS One*. 2013;8(11):e78173.

320. Song GG, Kim J-H, Lee YH. Associations between vascular endothelial growth factor gene polymorphisms and pre-eclampsia susceptibility: a meta-analysis. *Immunol Inves*t. 2013;42(8):749–62.

321. Morgan JA, Bombell S, McGuire W. Association of Plasminogen Activator Inhibitor-Type 1 (-675 4G/5G) Polymorphism with Pre-Eclampsia: Systematic Review. *PLoS One*. 2013;8(2):e56907.

322. Zhao L, Bracken MB, DeWan AT, Chen S. Association between the SERPINE1 (PAI-1) 4G/5G insertion/deletion promoter polymorphism (rs1799889) and pre-eclampsia: A systematic review and meta-analysis. *Molecular Human Reproduction*. 2013;19: 136–43.

323. Lin R, Lei Y, Yuan Z, Ju H, Li D. Angiotensinogen Gene M235T and T174M Polymorphisms and Susceptibility of Pre-Eclampsia: A Meta-Analysis. *Ann Hum Genet*. 2012;76(5):377–86.

324. Zhao L, Dewan AT, Bracken MB. Association of maternal AGTR1 polymorphisms and preeclampsia: A systematic review and meta-analysis. *Journal of Maternal-Fetal and Neonatal Medicine*. 2012;25:2676–80.

325. Zhong WG, Wang Y, Zhu H, Zhao X. Meta analysis of angiotensin-converting enzyme I/D polymorphism as a risk factor for preeclampsia in Chinese women. *Genet Mol Res*. 2012;11(3):2268–76.

326. Shaik AP, Sultana A, Bammidi VK, Sampathirao K, Jamil K. A meta-analysis of eNOS and ACE gene polymorphisms and risk of pre-eclampsia in women. *J Obstet Gynaecol*. 2011;31(7):603–7.

327. Xie C, Yao MZ, Liu JB, Xiong LK. A meta-analysis of tumor necrosis factor-alpha, interleukin-6, and interleukin-10 in preeclampsia. *Cytokine*. 2011;56:550–9.

328. Rodger MA, Betancourt MT, Clark P, Lindqvist PG, Dizon-Townson D, Said J, et al. The association of factor V leiden and prothrombin gene mutation and placenta-mediated pregnancy complications: A systematic review and meta-analysis of prospective cohort studies. *PLoS Med*. 2010;7(6):e1000292.

329. Medica I, Kastrin A, Peterlin B. Genetic polymorphisms in vasoactive genes and preeclampsia: A meta-analysis. *European Journal of Obstetrics Gynecology and Reproductive Biology*. 2007;131:115–26.

330. Serrano NC, Díaz LA, Páez MC, Mesa CM, Cifuentes R, Monterrosa A, et al. Angiotensin-converting enzyme I/D polymorphism and preeclampsia risk: Evidence of small-study bias. *PLoS Med*. 2006;3(12):2304–16.

331. Lin J, August P. Genetic thrombophilias and preeclampsia: A meta-analysis. *Obstetrics and Gynecology*. 2005;105:182–92.

332. Saccone G, Berghella V, Sarno L, Maruotti GM, Cetin I, Greco L, et al. Celiac disease and obstetric complications: A systematic review and metaanalysis. *American Journal of Obstetrics and Gynecology*. 2016;214:225–34.

333. Zhang JJ, Ma XX, Hao L, Liu LJ, Lv JC, Zhang H. A systematic review and

meta-analysis of outcomes of pregnancy in CKD and CKD outcomes in pregnancy. *Clin J Am Soc Nephrol*. 2015;10(11):1964–78.

334. Hu R, Li Y, Zhang Z, Yan W. Antenatal depressive symptoms and the risk of preeclampsia or operative deliveries: A meta-analysis. *PLoS One*. 2015;10(3):e0119018.

335. Qin JZ, Pang LH, Li MJ, Fan XJ, Huang RD, Chen HY. Obstetric complications in women with polycystic ovary syndrome: a systematic review and meta-analysis. *Reprod Biol Endocrinol*. 2013;11:56.

336. Zhang S, Ding Z, Liu H, Chen Z, Wu J, Zhang Y, et al. Association between mental stress and gestational hypertension/preeclampsia: a meta-analysis. *Obstet Gynecol Surv*. 2013;68(12):825–34.

337. Grigoriadis S, VonderPorten EH, Mamisashvili L, Tomlinson G, Dennis CL, Koren G, et al. The impact of maternal depression during pregnancy on perinatal outcomes: A systematic review and meta-analysis. J*ournal of Clinical Psychiatry*. 2013; 74:e321–41.

338. Schoenaker DAJM, Soedamah-Muthu SS, Mishra GD. The association between dietary factors and gestational hypertension and pre-eclampsia: A systematic review and meta-analysis of observational studies. *BMC Med*. 2014;12(1):157

339. Wei S-Q, Qi H-P, Luo Z-C, Fraser WD. Maternal vitamin D status and adverse pregnancy outcomes: a systematic review and meta-analysis. *J Matern Neonatal Med*. 2013;26(9):889–99.

340. Huang Q, Chen J, Zhong M, Hang L, Wei S, Yu Y. Chronic Hepatitis B Infection is Associated with Decreased Risk of Preeclampsia: A Meta-Analysis of Observational Studies. *Cell Physiol Biochem*. 2016;38(5):1860–8.

341. Sgolastra F, Petrucci A, Severino M, Gatto R, Monaco A. Relationship between Periodontitis and Pre-Eclampsia: A Meta-Analysis. *PLoS One*. 2013;8(8):e71387.

342. Rustveld LO, Kelsey SF, Sharma R. Association between maternal infections and preeclampsia: A systematic review of epidemiologic studies. *Maternal and Child Health Journal*. 2008;12:223–42.

343. Xu Y, Ren L, Zhai S, Luo X, Hong T, Liu R, et al. Association Between Isolated Single Umbilical Artery and Perinatal Outcomes: A Meta-Analysis. *Med Sci Monit*. 2016;22:1451–9.

344. Wei J, Liu C-X, Gong T-T, Wu Q-J, Wu L. Cigarette smoking during pregnancy and preeclampsia risk: a systematic review and meta-analysis of prospective studies. *Oncotarget.* 2015;6(41):43667–78.

345. Masoudian P, Nasr A, De Nanassy J, Fung-Kee-Fung K, Bainbridge SA, El Demellawy D. Oocyte donation pregnancies and the risk of preeclampsia or gestational hypertension: A systematic review and metaanalysis. *American Journal of Obstetrics and Gynecology*. 2016;214:328–39.

346. Aune D, Saugstad OD, Henriksen T, Tonstad S. Physical activity and the risk of preeclampsia: A systematic review and meta-analysis. *Epidemiology*. 2014;25:331–43.

347. González-Comadran M, Avila JU, Tascón AS, Jimenéz R, Solà I, Brassesco M, et al. The impact of donor insemination on the risk of preeclampsia: A systematic review and meta-analysis. *European Journal of Obstetrics Gynecology and Reproductive Biology*. 2014;182:160–6.

348. Wang Z, Wang P, Liu H, He X, Zhang J, Yan H, et al. Maternal adiposity as an independent risk factor for pre-eclampsia: A meta-analysis of prospective

cohort studies. *Obes Rev*. 2013;14(6):508–21.

349. Kasawara KT, Nascimento SL Do, Costa ML, Surita FG, E Silva JLP. Exercise and physical activity in the prevention of pre-eclampsia: Systematic review. *Acta Obstet Gynecol Scand.* 2012;91(10):1147–57.

350. Basaran A, Basaran M, Topatan B, Martin Jr. JN. Effect of chorionic villus sampling on the occurrence of preeclampsia and gestational hypertension: An updated systematic review and meta-analysis. *J Turkish Ger Gynecol Assoc.* 2016;17(2):65–72.

351. Luo ZC, An N, Xu HR, Larante A, Audibert F, Fraser WD. The effects and mechanisms of primiparity on the risk of pre-eclampsia: A systematic review. *Paediatric and Perinatal Epidemiology*. 2007;21:36–45.

352. Pergialiotis V, Prodromidou A, Frountzas M, Perrea DN, Papantoniou N. Maternal cardiac troponin levels in pre-eclampsia: a systematic review. *Journal of Maternal-Fetal and Neonatal Medicine*. 2016;29:3386–90.

353. Pergialiotis V, Prodromidou A, Pappa E, Vlachos GD, Perrea DN, Papantoniou N. An evaluation of calprotectin as serum marker of preeclampsia: a systematic review of observational studies. *Inflammation Research*. 2016;65:95–102.

354. Martin A, Krishna I, Martina B, Samuel A. Can the quantity of cell-free fetal DNA predict preeclampsia: a systematic review. *Prenatal diagnosis*. 2014;34:685–91.

355. Poursafa P, Keikha M, Kelishadi R. Systematic review on adverse birth outcomes of climate change. *Journal of Research in Medical Sciences.*2015;20: 397–402.

356. Conde-Agudelo A, Rosas-Bermúdez A, Kafury-Goeta AC. Effects of birth

spacing on maternal health: a systematic review. *Am J Obstet Gynecol*. 2007;196(4):297–308.

357. Oyebode F, Rastogi A, Berrisford G, Coccia F. Psychotropics in pregnancy: Safety and other considerations. *Pharmacol Ther*. 2012;135(1):71–7.

358. Smyth A, Oliveira GHM, Lahr BD, Bailey KR, Norby SM, Garovic VD. A Systematic Review and Meta-Analysis of Pregnancy Outcomes in Patients with Systemic Lupus Erythematosus and Lupus Nephritis. *Clin J Am Soc Nephrol*. 2010;5(11):2060–8.

359. Caimari F, Valassi E, Garbayo P, Steffensen C, Santos A, Corcoy R, et al. Cushing's syndrome and pregnancy outcomes: a systematic review of published cases. *Endocrine*. 2017;55(2):555–63.

360. Vrebosch L, Bel S, Vansant G, Guelinckx I, Devlieger R. Maternal and neonatal outcome after laparoscopic adjustable gastric banding: A systematic review. *Obesity Surger*y. 2012;22:1568–79.

361. Maggard MA, Yermilov I, Li Z, Maglione M, Newberry S, Suttorp M, et al. Pregnancy and Fertility Following Bariatric Surgery. *JAMA*. 2008;300(19):2286.

362. Bonzini M, Palmer KT, Coggon D, Carugno M, Cromi A, Ferrario MM. Shift work and pregnancy outcomes: A systematic review with meta-analysis of currently available epidemiological studies. *BJOG: An International Journal of Obstetrics and Gynaecology*. 2011;118:1429–37.

363. Adams JW, Watts DH, Phelps BR. A systematic review of the effect of HIV infection and antiretroviral therapy on the risk of pre-eclampsia. *International Journal of Gynecology and Obstetrics*. 2016;133: 17–21.

364. Kujovich JL. Thrombophilia and pregnancy complications. *Am J Obstet*

Gynecol. 2004;191(2):412–24.

365. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *European Journal of Epidemiology*. 2010;25:603–5.

366. Sohani ZN, Meyre D, de Souza RJ, Joseph PG, Gandhi M, Dennis BB, et al. Assessing the quality of published genetic association studies in meta-analyses: The quality of genetic studies (Q-Genie) tool. *BMC Genet*. 2015;16(1):50.

367. Genest DS, Falcao S, Gutkowska J, Lavoie JL. Impact of exercise training on preeclampsia: Potential preventive mechanisms. *Hypertension*. 2012;60:1104–9.

368. Girardi G, Yarilin D, Thurman JM, Holers VM, Salmon JE. Complement activation induces dysregulation of angiogenic factors and causes fetal rejection and growth restriction. *J Exp Med*. 2006;203(9):2165–75.

369. Van Der Hoorn MLP, Scherjon SA, Claas FHJ. Egg donation pregnancy as an immunological model for solid organ transplantation. *Transplant Immunology*. 2011;25(2–3):89–95.

370. Chernyshov VP, Tumanova LE, Sudoma IA, Bannikov VI. Th1 and Th2 in human IVF pregnancy with allogenic fetus. *Am J Reprod Immunol*. 2008;59(4):352–8.

371. Lashley L, van der Hoorn MLP, Haasnoot GW, Roelen DL, Claas FHJ. Uncomplicated oocyte donation pregnancies are associated with a higher incidence of human leukocyte antigen alloantibodies. *Hum Immunol*. 2014;75(6):555–60.

372. van der Hoorn MLP, Lashley EELO, Bianchi DW, Claas FHJ, Schonkeren

CMC, Scherjon SA. Clinical and immunologic aspects of egg donation pregnancies: A systematic review. *Hum Reprod Updat*e. 2010;16(6):704–12

373. Fenzl V, Flegar-Meštrić Z, Perkov S, Andrišić L, Tatzber F, Žarković N, et al. Trace elements and oxidative stress in hypertensive disorders of pregnancy. *Arch Gynecol Obste*t. 2013;287(1):19–24.

374. Casanueva E, Viteri FE. Iron and oxidative stress in pregnancy. *J Nutr.* 2003;133(5 Suppl 2):1700S–1708S.

375. Rayman MP, Barlis J, Evans RW, Redman CWG, King LJ. Abnormal iron parameters in the pregnancy syndrome preeclampsia. *Am J Obstet Gynecol*. 2002;187(2):412–8.

376. Spencer K, Cowans NJ, Nicolaides KH. Low levels of maternal serum PAPP-A in the first trimester and the risk of pre-eclampsia. *Prenat Diagn*. 2008;28(1):7–10.

377. Imbasciati E, Gregorini G, Cabiddu G, Gammaro L, Ambroso G, Del Giudice A, et al. Pregnancy in CKD Stages 3 to 5: Fetal and Maternal Outcomes. *Am J Kidney Dis*. 2007;49(6):753–62.

378. Jones DC, Hayslett JP. Outcome of Pregnancy in Women with Moderate or Severe Renal Insufficiency. *N Engl J Med*. 1996;335(4):226–32.

379. Williams D, Davison J. Chronic kidney disease in pregnancy. *BMJ*. 2008;336(7637):211–5.

380. Tehrani FR, Behboudi-Gandevani S. Polycystic ovary syndrome. In: Contemporary Gynaecologic Practice. Darwish A (ed). Intech Publishers, Rijeka Croatia. 2015; 79-102.

381. Troisi R, Potischman N, Johnson CN, Roberts JM, Lykins D, Harger G, et al. Estrogen and Androgen Concentrations Are Not Lower in the Umbilical Cord

Serum of Pre-eclamptic Pregnancies. *Cancer Epidemiol Biomarkers Prev*. 2003;12(11 II):1268–70.

382. Choi YK, Kim CK, Lee H, Jeoung D, Ha KS, Kwon YG, et al. Carbon monoxide promotes VEGF expression by increasing HIF-1α protein level via two distinct mechanisms, translational activation and stabilization of HIF-1α protein. *J Biol Chem*. 2010;285(42):32116–25.

383. Levine RJ, Maynard SE, Qian C, Lim K-H, England LJ, Yu KF, et al. Circulating Angiogenic Factors and the Risk of Preeclampsia. *N Engl J Med*. 2004;350(7):672–83.

384. Levine RJ, Lam C, Qian C, Yu KF, Maynard SE, Sachs BP, et al. Soluble Endoglin and Other Circulating Antiangiogenic Factors in Preeclampsia. *N Engl J Med*. 2006;355(10):992–1005.

385. Cudmore M, Ahmad S, Al-Ani B, Fujisawa T, Coxall H, Chudasama K, et al. Negative regulation of soluble Flt-1 and soluble endoglin release by heme oxygenase-1. *Circulation*. 2007;115(13):1789–97.

386. Von Dadelszen P, Magee LA. Could an infectious trigger explain the differential maternal response to the shared placental pathology of preeclampsia and normotensive intrauterine growth restriction? *Acta Obstet Gynecol Scand*. 2002;81(7):642–8.

387. Arechavaleta-Velasco F, Ma Y, Zhang J, McGrath CM, Parry S. Adeno-associated virus-2 (AAV-2) causes trophoblast dysfunction, and placental AAV-2 infection is associated with preeclampsia. *Am J Pathol*. 2006;168(6):1951–9.

388. Tzoulaki I, Siontis KC, Evangelou E, Ioannidis JPA. Bias in associations of emerging biomarkers with cardiovascular disease. *JAMA Intern Med*.

2013;173(8):664–71.

389. Tzoulaki I, Siontis KCM, Ioannidis JPA. Prognostic effect size of cardiovascular biomarkers in datasets from observational studies versus randomised trials: Meta-epidemiology study. *BMJ*. 2011;343(7834):1137.

390. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. 2007;370(9596):1453–7.

391. Little J, Higgins JPT, Ioannidis JPA, Moher D, Gagnon F, Von Elm E, et al. STrengthening the REporting of genetic association studies (STREGA)- An extension of the STROBE statement. *Genet Epidemiol*. 2009;33(7):581–98.

392. Dal-Ré R, Ioannidis JP, Bracken MB, Buffler PA, Chan AW, Franco EL, et al. Making prospective registration of observational research a reality. *Science Translational Medicine*. 2014; 19(6):224cm1-224cm1.

393. Ioannidis JPA. The importance of potential studies that have not existed and registration of observational data sets. *JAMA - Journal of the American Medical Association*. 2012;308:575–6.

394. Staff AC, Sibai BM, Cunningham FG. Prevention of Preeclampsia and Eclampsia. In: Chesley's Hypertensive Disorders in Pregnancy. Elsevier; 2015. 253–67.

395. Chaiworapongsa T, Chaemsaithong P, Korzeniewski SJ, Yeo L, Romero R. Pre-eclampsia part 2: prediction, prevention and management. *Nat Rev Nephrol*. 2014;10(9):531–40.

396. Leslie MS, Briggs LA. Preeclampsia and the Risk of Future Vascular Disease and Mortality: A Review. *J Midwifery Womens Health*. 2016;61(3):315–24.

397. Duley L. The Global Impact of Pre-eclampsia and Eclampsia. *Semin Perinatol.* 2009;33(3):130–7.

398. Dugoff L, Hobbins JC, Malone FD, Vidaver J, Sullivan L, Canick JA, et al. Quad Screen as a Predictor of Adverse Pregnancy Outcome. *Obstet Gynecol.* 2005;106(2):260–7.

399. Poon LCY, Kametas NA, Maiz N, Akolekar R, Nicolaides KH. First-Trimester Prediction of Hypertensive Disorders in Pregnancy. *Hypertension*. 2009;53(5):812–8.

400. Ioannidis JPA, Karassa FB. The need to consider the wider agenda in systematic reviews and meta-analyses: breadth, timing, and depth of the evidence. *BMJ*. 2010;341:c4875.

401. Lubin JH, Gail MH. On power and sample size for studying features of the relative odds of disease. Am J Epidemiol. 1990;131(3):552–566.

402. De-Regil L, Palacios C, Lombardo L, Pena-Rosas J, Peña-Rosas J. Vitamin D supplementation for women during pregnancy. *Cochrane Database Syst Rev*. 2016;1(1):CD008873.

403. Rumbold A, Ota E, Hori H, Miyazaki C, Crowther CA. Vitamin E supplementation in pregnancy. *CochraneDatabase.of Syst*. 2015;(2):1–123.

404. Makrides M, Crosby DD, Bain E, Crowther CA. Magnesium supplementation in pregnancy. *Cochrane Database Syst Rev*. 2014;(4):1–75.

405. Pérez-López FR, Pasupuleti V, Mezones-Holguin E, Benites-Zapata VA, Thota P, Deshpande A, et al. Effect of vitamin D supplementation during pregnancy on maternal and neonatal outcomes: A systematic review and meta-analysis of randomized controlled trials. *Fertil Steril*. 2015;103(5):1278–1288.e4.

406. Gui S, Jia J, Niu X, Bai Y, Zou H, Deng J, et al. Arginine supplementation for improving maternal and neonatal outcomes in hypertensive disorder of pregnancy: A systematic review. *J Renin-Angiotensin-Aldosterone Syst*. 2014 Mar;15(1):88–96.

407. Roberge S, Sibai B, McCaw-Binns A, Bujold E. Low-Dose Aspirin in Early Gestation for Prevention of Preeclampsia and Small-for-Gestational-Age Neonates: Meta-analysis of Large Randomized Trials. *Am J Perinatol*. 2016;33(8):781–5.

408. Patrelli TS, Dall'Asta A, Gizzo S, Pedrazzi G, Piantelli G, Jasonni VM, et al. Calcium supplementation and prevention of preeclampsia: a meta-analysis. *J Matern Neonatal Med*. 2012;25(12):2570–4.

409. Roberge S, Villa P, Nicolaides K, Giguére Y, Vainio M, Bakthi A, et al. Early administration of low-dose aspirin for the prevention of preterm and term preeclampsia: A systematic review and meta-analysis. *Fetal Diagn Ther*. 2012;31(3):141–6.

410. Villa PM, Kajantie E, Räikkönen K, Pesonen AK, Hämäläinen E, Vainio M, et al. Aspirin in the prevention of pre-eclampsia in high-risk women: A randomised placebo-controlled PREDO Trial and a meta-analysis of randomised trials. *BJOG An Int J Obstet Gynaecol*. 2013;120(1):64–74.

411. Rumbold A, Ota E, Nagata C, Shahrook S, Crowther CA. Vitamin C supplementation in pregnancy. *Cochrane Database Syst Rev*. 2015;(9):CD004072.

412. Bujold E, Morency A-M, Roberge S, Lacasse Y, Forest J-C, Giguère Y. Acetylsalicylic Acid for the Prevention of Preeclampsia and Intra-uterine Growth Restriction in Women with Abnormal Uterine Artery Doppler: A

Systematic Review and Meta-analysis. *J Obstet Gynaecol Canada.* 2009;31(9):818–26.

413. Chen B, Ji X, Zhang L, Hou Z, Li C, Tong Y. Fish Oil Supplementation does not Reduce Risks of Gestational Diabetes Mellitus, Pregnancy-Induced Hypertension, or Pre-Eclampsia: A Meta-Analysis of Randomized Controlled Trials. *Med Sci Monit*. 2015;21:2322–30.

414. Roberge S, Demers S, Nicolaides KH, Bureau M, Côté S, Bujold E. Prevention of pre-eclampsia by low-molecular-weight heparin in addition to aspirin: A meta-analysis. *Ultrasound Obstet Gynecol.* 2016;47(5):548–53.

415. Muktabhant B, Ta L, Lumbiganon P, Laopaiboon M. Diet or exercise, or both, for preventing excessive weight gain in pregnancy (Review). *Cochrane Database Syst Rev*. 2015;(6):CD007145.

416. Conde-Agudelo A, Romero R, Kusanovic JP, Hassan SS. Supplementation with vitamins C and E during pregnancy for the prevention of preeclampsia and other adverse maternal and perinatal outcomes: a systematic review and metaanalysis. *American Journal of Obstetrics & Gynecology*. 2011;204(6):503-e1-e12.

417. Imhoff-Kunsch B, Briggs V, Goldenberg T, Ramakrishnan U. Effect of n-3 long-chain polyunsaturated fatty acid intake during pregnancy on maternal, infant, and child health outcomes: A systematic review. *Paediatr Perinat Epidemiol*. 2012;26(1):91–107.

418. Thangaratinam S, Rogozinska E, Jolly K, Glinkowski S, Roseboom T, Tomlinson JW, et al. Effects of interventions in pregnancy on maternal weight and obstetric outcomes: meta-analysis of randomised evidence. *Bmj*. 2012;344:e2088–e2088.

419. Rumbold A, Duley L, Ca C, Rr H. Antioxidants for preventing pre-eclampsia. *The Cochrane Library*. 2008;(4): CD004227.

420. Hofmeyr JG, Lawrie TA, Atallah AN, Duley L, Torloni MR. Calcium supplementation during pregnancy for preventing hypertensive disorders and related problems. *Cochrane Database Syst Rev*. 2014;(6):CD001059.

421. Tang R, Tang IC, Henry A, Welsh A. Limited evidence for calcium supplementation in preeclampsia prevention: a meta-analysis and systematic review. *Hypertens Pregnancy*. 2015;34(2):181–203.

422. Zheng J, Wang H, Ren M. Influence of exercise intervention on gestational diabetes mellitus: a systematic review and meta-analysis. *J Endocrinol Invest.* 2017;40(10):1027–33.

423. Trivedi N. A meta-analysis of low-dose aspirin for prevention of preeclampsia. *J Postgrad Med*. 2011;57(2):91.

424. An L, Li W, Xie T, Peng X, Li B, Xie S, et al. Calcium supplementation reducing the risk of hypertensive disorders of pregnancy and related problems: A meta-analysis of multicentre randomized controlled trials. *Int J Nurs Pract*. 2015;21:19–31..

425. Roberge S, Nicolaides K, Demers S, Hyett J, Chaillet N, Bujold E. The role of aspirin dose on the prevention of preeclampsia and fetal growth restriction: systematic review and meta-analysis. *Am J Obstet Gynecol*. 2017;216(2):110–120.

426. Allen R, Rogozinska E, Sivarajasingam P, Khan KS, Thangaratinam S. Effect of diet- And lifestyle-based metabolic risk-modifying interventions on preeclampsia: A meta-analysis. *Acta Obstet Gynecol Scand*. 2014;93(10):973–85.

427. Meher S, Duley L, Shireen M, Lelia D. Nitric Oxide for Preventing Pre-Eclampsia and Its Complications. *Cochrane Database Syst Rev*. 2014;(2):2–4.

428. Henderson JT, Whitlock EP, O'Connor E, Senger CA, Thompson JH, Rowland MG. Low-dose aspirin for prevention of morbidity and mortality from preeclampsia: A systematic evidence review for the u.s. preventive services task force. *Ann Intern Med*. 2014;160(10):695–703.

429. Dodd JM, McLeod A, Windrim RC, Kingdom J. Antithrombotic therapy for improving maternal or infant health outcomes in women considered at risk of placental dysfunction. *Cochrane Database Syst Rev*. 2013;7(7):Cd006780.

430. Rolnik DL, Wright D, Poon LC, O'Gorman N, Syngelaki A, de Paco Matallana C, et al. Aspirin versus Placebo in Pregnancies at High Risk for Preterm Preeclampsia. *N Engl J Med*. 2017;377(7):613–22.

431. American College of Obstetricians and Gynecologists. Practice Advisory on Low-Dose Aspirin and Prevention of Preeclampsia: Updated Recommendations - ACOG. Available from: https://www.acog.org/About-ACOG/News-Room/Practice-Advisories/Practice-Advisory-Low-Dose-Aspirin-and-Prevention-of-Preeclampsia-Updated-Recommendations [accessed 11 Nov2017].

432. LeFevre ML, U.S. Preventive Services Task Force. Low-Dose Aspirin Use for the Prevention of Morbidity and Mortality From Preeclampsia: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med.* 2014;161(11):819.

433. Bartsch E, Medcalf KE, Park AL, Ray JG, High Risk of Pre-eclampsia Identification Group. Clinical risk factors for pre-eclampsia determined in early pregnancy: systematic review and meta-analysis of large cohort studies.

BMJ. 2016;353:i1753.

434. Greene MF, Solomon CG. Aspirin to Prevent Preeclampsia. *N Engl J Med*. 2017;377(7):690–1.

435. Muktabhant B, Lawrie TA, Lumbiganon P, Laopaiboon M, Ta L, Lumbiganon P, et al. Diet or exercise, or both, for preventing excessive weight gain in pregnancy (Review). Muktabhant B, editor. *Cochrane Database Syst Rev*. 2015;6:CD007145.

436. Ray J, Diamond P, Singh G, Bell C. Brief overview of maternal triglycerides as a risk factor for pre-eclampsia. *BJOG An Int J Obstet Gynaecol*. 2006;113(4):379–86.

437. Qiu C, Coughlin KB, Frederick IO, Sorensen TK, Williams MA. Dietary Fiber Intake in Early Pregnancy and Risk of Subsequent Preeclampsia. *Am J Hypertens*. 2008;21(8):903–9.

438. Khoury J, Henriksen T, Christophersen B, Tonstad S. Effect of a cholesterol-lowering diet on maternal, cord, and neonatal lipids, and pregnancy outcome: A randomized clinical trial. *Am J Obstet Gynecol*. 2005;193(4):1292–301.

439. Belizán JM, Villar J, Repke J. The relationship between calcium intake and pregnancy-induced hypertension: up-to-date evidence. *Am J Obstet Gynecol*. 1988;158(4):898–902.

440. Villar J, Merialdi M, Gülmezoglu AM, Abalos E, Carroli G, Kulier R, et al. Nutritional interventions during pregnancy for the prevention or treatment of maternal morbidity and preterm delivery: an overview of randomized controlled trials. *J Nutr*. 2003;133(5 Suppl 2):1606S–1625S.

441. Roberts JM, Speer P. Antioxidant Therapy to Prevent Preeclampsia. *Semin Nephrol*. 24:557–64.

442. Kumar A, Devi SG, Batra S, Singh C, Shukla DK. Calcium supplementation for the prevention of pre-eclampsia. *Int J Gynecol Obstet*. 2009;104(1):32–6.

443. Repke JT, Villar J. Pregnancy-induced hypertension and low birth weight: the role of calcium. *Am J Clin Nutr*. 1991;54(1 Suppl):237S–241S.

444. Villar J, Repke J, Markush L, Calvert W, Rhoads G. The measuring of blood pressure during pregnancy. *Am J Obs Gynecol*. 1989;161(4):1019–24.

445. Palmer RMJ, Ashton DS, Moncada S. Vascular endothelial cells synthesize nitric oxide from L-arginine. *Nature*.1988;333(6174):664–6.

446. Neri I, Valensise H, Facchinetti F, Menghini S, Romanini C, Volpe A. 24-Hour Ambulatory Blood Pressure Monitoring: A Comparison Between Transdermal Glyceryl-Trinitrate and Oral Nifedipine. *Hypertens Pregnancy.* 1999;18(1):107–13.

447. Molnár M, Sütö T, Tóth T, Hertelendy F. Prolonged blockade of nitric oxide synthesis in gravid rats produces sustained hypertension, proteinuria, thrombocytopenia, and intrauterine growth retardation. *Am J Obstet Gynecol.* 1994;170(5):1458–66.

448. Baylis C, Mitruka B, Deng A. Chronic blockade of nitric oxide synthesis in the rat produces systemic hypertension and glomerular damage. *J Clin Invest.* 1992;90(1):278–81.

449. Biri A, Bozkurt N, Gunaydin G, Korucuoglu U, Durak I, Kavutcu M. Antioxidant enzyme activities and lipid peroxidation in preeclampsia. *Int J Gynecol Obstet*. 2007;96(3):196–7.

450. Chamy VM, Lepe J, Catalán A, Retamal D, Escobar JA, Madrid EM. Oxidative stress is closely related to clinical severity of pre-eclampsia. *Biol Res*. 2006;39(2):229–36.

451. Serdar Z, Gür E, Develioğlu O, Colakoğullari M, Dirican M. Placental and decidual lipid peroxidation and antioxidant defenses in preeclampsia. Lipid peroxidation in preeclampsia. *Pathophysiol Off J Int Soc Pathophysiol*. 2002;9(1):21.

452. Benjamin DJ, Berger JO, Johannesson M, Nosek BA, Wagenmakers E-J, Berk R, et al. Redefine statistical significance. *Nat Hum Behav*. 2017;2(1):6

453. Reece EA, Leguizamón G, Wiznitzer A. Gestational diabetes: the need for a common ground. *Lancet*. 2009;373(9677):1789–97.

454. Ferrara A. Increasing prevalence of gestational diabetes mellitus: a public health perspective. *Diabetes Care*. 2007;30 Suppl 2(Supplement 2):S141-6.

455. American Diabetes Association AD. Standards of medical care in diabetes--2014. *Diabetes Care*. 2014;37 Suppl 1(Supplement 1):S14-80.

456. Jenum AK, Mørkrid K, Sletner L, Vange S, Torper JL, Nakstad B, et al. Impact of ethnicity on gestational diabetes identified with the WHO and the modified International Association of Diabetes and Pregnancy Study Groups criteria: a population-based cohort study. *Eur J Endocrinol*. 166:317–24.

457. Xiong X, Saunders LD, Wang FL, Demianczuk NN. Gestational diabetes mellitus: prevalence, risk factors, maternal and infant outcomes. *Int J Gynecol Obstet.* 2001;75(3):221–8.

458. Catalano PM, McIntyre HD, Cruickshank JK, McCance DR, Dyer AR, Metzger BE, et al. The hyperglycemia and adverse pregnancy outcome study: Associations of GDM and obesity with pregnancy outcomes. *Diabetes Care.* 2012;35(4):780–6.

459. Group THSCR. Hyperglycemia and Adverse Pregnancy Outcomes. *N Engl J Med*. 2008;358(19):1991–2002.

460. Buchanan TA, Xiang AH, Page KA. Gestational diabetes mellitus: Risks and management during and after pregnancy. *Nat Rev Endocrinol.* 2012;8(11):639–49.

461. Ben-Haroush A, Yogev Y, Hod M. Epidemiology of gestational diabetes mellitus and its association with Type 2 diabetes. *Diabet Med.* 2004;21(2):103–13.

462. Kampmann U. Gestational diabetes: A clinical update. *World J Diabetes*. 2015;6(8):1065.

463. Zhang C, Rawal S, Chong YS. Risk factors for gestational diabetes: is prevention possible? *Diabetologia*. 2016;59:1385–90.

464. Zhang C, Ning Y. Effect of dietary and lifestyle factors on the risk of gestational diabetes: review of epidemiologic evidence. *Am J Clin Nutr*. 2011;94(6 Suppl):1975S–1979S.

465. Kong F-J, Ma L-L, Li G, Chen Y-X, Zhou J-Q. Circulating Betatrophin Levels and Gestational Diabetes Mellitus: A Systematic Review and Meta-Analysis. *PLoS One*. 2017;12(1):e0169941.

466. Fu S, Li F, Zhou J, Liu Z. The Relationship Between Body Iron Status, Iron Intake And Gestational Diabetes. *Medicine*. 2016;95(2):e2383.

467. Fernández-Cao JC, Aranda N, Ribot B, Tous M, Arija V. Elevated iron status and risk of gestational diabetes mellitus: A systematic review and meta-analysis. *Matern Child Nutr.* 2017;13(4):e12400.

468. Kong FJ, Ma LL, Chen SP, Li G, Zhou JQ. Serum selenium level and gestational diabetes mellitus: a systematic review and meta-analysis. *Nutr J.* 2016;15(1):94.

469. Hu S, Liu Q, Huang X, Tan H. Serum level and polymorphisms of retinol-

binding protein-4 and risk for gestational diabetes mellitus: A meta-analysis. *BMC Pregnancy Childbirth.* 2016;16(1):1–11.

470. Iliodromiti S, Sassarini J, Kelsey TW, Lindsay RS, Sattar N, Nelson SM. Accuracy of circulating adiponectin for predicting gestational diabetes: a systematic review and meta-analysis. *Diabetologia*. 2016;59:692–9.

471. Guo C, Jin Y, Lee KKH, Yang G, Jing C, Yang X. The relationships between HLA class II alleles and antigens with gestational diabetes mellitus: A meta-analysis. *Sci Rep*. 2016;6(1):35005.

472. Yang Y, Li Q, Wang Q, Ma X. Thyroid antibodies and gestational diabetes mellitus: A meta-analysis. *Fertil Steril*. 2015;104(3):665–671.

473. Zhang MX, Pan GT, Guo JF, Li BY, Qin LQ, Zhang ZL. Vitamin D Deficiency Increases the Risk of Gestational Diabetes Mellitus: A Meta-Analysis of Observational Studies. *Nutrients*. 2015;7(10):8366–75.

474. Aghajafari F, Nagulesapillai T, Ronksley PE, Tough SC, O'Beirne M, Rabi DM. Association between maternal serum 25-hydroxyvitamin D level and pregnancy and neonatal outcomes: systematic review and meta-analysis of observational studies. *Bmj*. 2013;346:f1169–f1169.

475. Aune D, Sen A, Henriksen T, Saugstad OD, Tonstad S. Physical activity and the risk of gestational diabetes mellitus: a systematic review and dose–response meta-analysis of epidemiological studies. *Eur J Epidemiol*. 2016;31(10):967–97.

476. Torloni MR, Betrán AP, Horta BL, Nakamura MU, Atallah AN, Moron AF, et al. Prepregnancy BMI and the risk of gestational diabetes: A systematic review of the literature with meta-analysis: Diagnostic in Obesity and Complications. *Obes Rev*. 2009;10(2):194–203.

477. Gong LL, Liu H, Liu LH. Relationship between hypothyroidism and the incidence of gestational diabetes: A meta-analysis. *Taiwan J Obstet Gynecol*. 2016;55(2):171–5.

478. Luque-Fernandez MA, Bain PA, Gelaye B, Redline S, Williams MA. Sleep-Disordered Breathing and Gestational Diabetes Mellitus: A meta-analysis of 9,795 participants enrolled in epidemiological observational studies. *Diabetes Care.* 2013;36(10):3353–60.

479. Kjerulff LE, Sanchez-Ramos L, Duffy D. Pregnancy outcomes in women with polycystic ovary syndrome: A metaanalysis. *Am J Obstet Gynecol*. 2011;204(6):558.e1-558.e6.

480. Abariga SA, Whitcomb BW. Periodontitis and gestational diabetes mellitus: A systematic review and meta-analysis of observational studies. *BMC Pregnancy Childbirth*. 2016;16(1):344.

481. Soepnel LM, Norris SA, Schrier VJMM, Browne JL, Rijken MJ, Gray G, et al. The association between HIV, antiretroviral therapy, and gestational diabetes mellitus. *Aids*. 2017;31:113-125

482. Moosazadeh M, Asemi Z, Lankarani KB, Tabrizi R, Maharlouei N, Naghibzadeh-Tahami A, et al. Family history of diabetes and the risk of gestational diabetes mellitus in Iran: A systematic review and meta-analysis. *Diabetes Metab Syndr Clin Res Rev*. 2017;11(2):S99–104.

483. Pandey S, Shetty A, Hamilton M, Bhattacharya S, Maheshwari A. Obstetric and perinatal outcomes in singleton pregnancies resulting from ivf/icsi: A systematic review and meta-analysis. *Hum Reprod Update*. 2012;18(5):485–503.

484. Rasmussen KM, Yaktine AL. Committee to Reexamine IOM Pregnancy

Weight Guidelines. Food and Nutrition Board, Board on Children, Youth and Families, Institute of Medicine, National Research Council. Weight gain during pregnancy: reexamining the guidelines. Washington, DC: National Academies Press. 2009..

485. National Institute for Health and Clinical Excellence. Weight management before, during and after pregnancy | Guidance and guidelines | NICE. Available from: https://www.nice.org.uk/guidance/ph27. [Accessed 14 February 2018]

486. American College of Obstetricians and Gynecologists. ACOG Committee opinion. Number 267, January 2002: exercise during pregnancy and the postpartum period. *Obstet Gynecol.* 2002;99(1):171–3.

487. Shepherd E, Gomersall JC, Tieu J, Han S, Crowther CA, Middleton P. Combined diet and exercise interventions for preventing gestational diabetes mellitus. *Cochrane Database Syst Rev*. 2017;11:CD010443.

488. King JC. Maternal Obesity, Metabolism, and Pregnancy Outcomes. *Annu Rev Nutr*. 2006;26(1):271–91.

489. Kahn B, Flier J. Obesity and insulin resistance. *J Clin Invest.* 2000;106(4):473–81.

490. Sathyapalan T, Mellor D, Atkin SL. Obesity and gestational diabetes. *Semin Fetal Neonatal Med.* 2010;15(2):89–93.

491. Pantham P, Aye ILMH, Powell TL. Inflammation in maternal obesity and gestational diabetes mellitus. *Placenta*. 2015;36(7):709–15.

492. Yudkin JS, Stehouwer CDA, Emeis JJ, Coppack SW. C-reactive protein in healthy subjects: Associations with obesity, insulin resistance, and endothelial dysfunction: A potential role for cytokines originating from adipose tissue?

Arterioscler Thromb Vasc Biol. 1999;19(4):972–8.

493. Ford ES. Body mass index, diabetes, and C-reactive protein among U.S. Adults. *Diabetes Care*. 1999;22(12):1971–7.

494. Visser M. Elevated C-Reactive Protein Levels in Overweight and Obese Adults. *JAMA*. 1999;282(22):2131.

495. Wolf M, Sandler L, Hsu K, Vossen-Smirnakis K, Ecker JL, Thadhani R. First-trimester C-reactive protein and subsequent gestational diabetes. *Diabetes Care*. 2003;26(3):819–24.

496. Qiu C, Sorensen TK, Luthy DA, Williams MA. A prospective study of maternal serum C-reactive protein (CRP) concentrations and risk of qestational diabetes mellitus. *Paediatr Perinat Epidemiol*. 2004;18(5):377–84.

497. Chen X, Scholl TO, Stein TP. Association of Elevated Serum Ferritin Levels and the Risk of Gestational Diabetes Mellitus in Pregnant Women: The Camden Study. *Diabetes Care.* 2006;29(5):1077–82.

498. Asemi Z, Hashemi T, Karamali M, Samimi M, Esmaillzadeh A. Effects of vitamin D supplementation on glucose metabolism, lipid concentrations, inflammation, and oxidative stress in gestational diabetes: A double-blind randomized controlled clinical trial1-3. *Am J Clin Nutr*. 2013;98(6):1425–32.

499. Soheilykhah S, Mojibian M, Moghadam MJ, Shojaoddiny-Ardekani A. The effect of different doses of vitamin D supplementation on insulin resistance during pregnancy. *Gynecol Endocrinol*. 2013;29(4):396–9.

500. Asemi Z, Samimi M, Tabassi Z, Shakeri H, Esmaillzadeh A. Vitamin D Supplementation Affects Serum High-Sensitivity C-Reactive Protein, Insulin Resistance, and Biomarkers of Oxidative Stress in Pregnant Women. *J Nutr.* 2013;143(9):1432–8.

501. Poel YHM, Hummel P, Lips P, Stam F, van der Ploeg T, Simsek S. Vitamin D and gestational diabetes: A systematic review and meta-analysis. *Eur J Intern Med*. 2012;23(5):465–9.

502. Roth DE, Leung M, Mesfin E, Qamar H, Watterworth J, Papp E. Vitamin D supplementation during pregnancy: state of the evidence from a systematic review of randomised trials. bmj *BMJ*. 2017;359:j5237.

503. World Health Organization. Guideline: Vitamin D supplementation in pregnant women. World Heal Organ. 2012. Available from: http://apps.who.int/iris/bitstream/10665/85313/1/9789241504935_eng.pdf?ua=1 [Accessed 14 February 2018]

504. Urrutia RP, Thorp JM. Vitamin D in pregnancy. *Curr Opin Obstet Gynecol*. 2012;24(2):57–64.

505. Hollis BW. Vitamin D requirement during pregnancy and lactation. In: *Journal of Bone and Mineral Research*. 2007:22:39–44.

506. Lee JM, Smith JR, Philipp BL, Chen TC, Mathieu J, Holick MF. Vitamin D deficiency in a healthy group of mothers and newborn infants. *Clin Pediatr.* 2007;46(1):42–4.

507. Bodnar LM, Simhan HN, Powers RW, Frank MP, Cooperstein E, Roberts JM. High Prevalence of Vitamin D Insufficiency in Black and White Pregnant Women Residing in the Northern United States and Their Neonates 1. *J Nutr*. 2007;137(2):447–52.

508. ACOG Committee on Obstetric Practice. Committee opinion no. 495: Vitamin D: Screening and supplementation during pregnancy. *Obstetrics and Gynecology.* 2011;118:197–8.

509. Chiu KC, Chu A, Go VLW, Saad MF. Hypovitaminosis D is associated with

insulin resistance and β cell dysfunction. *Am J Clin Nutr*. 2004;25(4):820–5.

510. Norman A, Frankel J, Heldt A, Grodsky G. Vitamin D deficiency inhibits pancreatic secretion of insulin. *Science*. 1980;209(4458):823–5.

511. Sooy K, Schermerhorn T, Noda M, Surana M, Rhoten WB, Meyer M, et al. Calbindin-D(28k) controls [Ca2+](i) and insulin release. Evidence obtained from calbindin-D(28k) knockout mice and β cell lines. *J Biol Chem*. 1999;274(48):34343–9.

512. Alvarez JA, Ashraf A. Role of Vitamin D in Insulin Secretion and Insulin Sensitivity for Glucose Homeostasis. *Int J Endocrinol*. 2010:1–18.

513. Draznin B, Sussman KE, Eckel RH, Kao M, Yost T, Sherman NA. Possible role of cytosolic free calcium concentrations in mediating insulin resistance of obesity and hyperinsulinemia. *J Clin Invest*. 1988;82(6):1848–52.

514. Springer D, Jiskra J, Limanova Z, Zima T, Potlukova E. Thyroid in pregnancy: From physiology to screening. *Crit Rev Clin Lab Sci*. 2017;54(2):102–16.

515. Maratou E, Hadjidakis DJ, Kollias A, Tsegka K, Peppa M, Alevizaki M, et al. Studies of insulin resistance in patients with clinical and subclinical hypothyroidism. *Eur J Endocrinol*. 2009;160(5):785–90.

516. Garduño-Garcia JDJ, Alvirde-Garcia U, López-Carrasco G, Mendoza MEP, Mehta R, Arellano-Campos O, et al. TSH and free thyroxine concentrations are associated with differing metabolic markers in euthyroid subjects. *Eur J Endocrinol*. 2010;163(2):273–8.

517. Roos A, Bakker SJL, Links TP, Gans ROB, Wolffenbuttel BHR. Thyroid Function Is Associated with Components of the Metabolic Syndrome in Euthyroid Subjects. *J Clin Endocrinol Metab*. 2007;92(2):491–6.

518. Dimitriadis G, Mitrou P, Lambadiari V, Boutati E, Maratou E, Panagiotakos DB, et al. Insulin action in adipose tissue and muscle in hypothyroidism. J *Clin Endocrinol Metab*. 2006;91(12):4930–7.

519. Velasco I, Taylor P. Identifying and treating subclinical thyroid dysfunction in pregnancy: emerging controversies. *Eur J Endocrinol.* 2018;178(1):D1–12.

520. American College of Obstetricians & Gynecologists. ACOG practice bulletin no. 37: Thyroid disease in pregnancy. *Obs Gynecol*. 2002;100(2):387–96.

521. De Groot L, Abalovich M, Alexander EK, Amino N, Barbour L, Cobin RH, et al. Management of thyroid dysfunction during pregnancy and postpartum: An endocrine society clinical practice guideline. *Journal of Clinical Endocrinology and Metabolism*. 2012;97: 2543–65.

522. Baskin HJ, Cobin RH, Duick DS, Gharib H, Guttler RB, Kaplan MM, et al. American Association of Clinical Endocrinologists Medical Guidelines for clinical practice for the evaluation and treatment of hypogonadism in adult male patients--2002 update. *Endocr Pract Off J Am Coll Endocrinol Am Assoc Clin Endocrinol*. 2002;8(6):440–56.

523. Prasad V, Jorgenson J, Ioannidis JPA, Cifu A. Observational studies often make clinical practice recommendations: An empirical evaluation of authors' attitudes. *J Clin Epidemiol*. 2013;66(4)361-6.

524. Rayanagoudar G, Hashi AA, Zamora J, Khan KS, Hitman GA, Thangaratinam S. Quantification of the type 2 diabetes risk in women with gestational diabetes: a systematic review and meta-analysis of 95,750 women. *Diabetologia*. 2016;59:1403–11.

525. Burlina S, Dalfrà MG, Chilelli NC, Lapolla A. Gestational Diabetes Mellitus and Future Cardiovascular Risk: An Update. *International Journal of*

Endocrinology. 2016;2016:1–6.

526. Sullivan SD, Umans JG, Ratner R. Gestational diabetes: Implications for cardiovascular health. *Curr Diab Rep*. 2012;12(1):43–52.

527. Harreiter J, Dovjak G, Kautzky-Willer A. Gestational Diabetes Mellitus and Cardiovascular Risk after Pregnancy. *Women's Heal*. 2014;10(1):91–108.

528. Lau J. The case of the misleading funnel plot. *BMJ*. 2006;333(7568):597–600.

529. Guyatt GH, Meade MO, Jaeschke RZ, Cook DJ, Haynes RB. Practitioners of evidence based care. Not all clinicians need to appraise evidence from scratch but all need some skills. *BMJ*. 2000;320(7240):954–5.

530. Whittemore R, Chao A, Jang M, Minges KE, Park C. Methods for knowledge synthesis: An overview. *Hear Lung J Acute Crit Care*. 2014;43(5):453–61.

531. Lutjen P, Trounson A, Leeton J, Findlay J, Wood C, Renou P. The establishment and maintenance of pregnancy using in vitro fertilization and embryo donation in a patient with primary ovarian failure. *Nature*. 1984;307(5947):174–5.

532. Söderström-Anttila V, Sälevaara M, Suikkari AM. Increasing openness in oocyte donation families regarding disclosure over 15 years. *Hum Reprod.* 2010;25(10):2535–42.

533. Savasi VM, Mandia L, Laoreti A, Cetin I. Maternal and fetal outcomes in oocyte donation pregnancies. *Hum Reprod Update*. 2016;22(5):620–33.

534. Human Fertilisation and Embryology Authority. Fertility treatment 2014–2016: Trends and figures. 2018. Available from: https://www.hfea.gov.uk/media/2544/hfea-fertility-treatment-2014-2016-trends-and-figures.pdf [accessed 1/1/18]

535. Letur H, Peigné M, Ohl J, Cédrin-Durnerin I, Mathieu-D'Argent E, Scheffler

F, et al. Hypertensive pathologies and egg donation pregnancies: Results of a large comparative cohort study. *Fertil Steril*. 2016;106(2):284–90.

536. Clua E, Tur R, Coroleu B, Boada M, Rodríguez I, Barri PN, et al. Elective single-embryo transfer in oocyte donation programmes: should it be the rule? *Reprod Biomed Online*. 2012;25(6):642–8.

537. Pecks U, Maass N, Neulen J. Oocyte donation: a risk factor for pregnancy-induced hypertension: a meta-analysis and case series. *Dtsch Arztebl Int.* 2011;108(3):23–31.

538. Duley L, Henderson-Smart DJ, Meher S, King JF. Antiplatelet agents for preventing pre-eclampsia and its complications. *Cochrane Database of Systematic Reviews*. 2007. CD004659.

539. American Diabetes Association AD. 13. Management of Diabetes in Pregnancy:Standards of Medical Care in Diabetes-2018. *Diabetes Care*. 2018;41(Suppl 1):S137–43.

540. National Institute for Health and Care Excellence.Diabetes in pregnancy: management from preconception to the postnatal period. Available from: https://www.nice.org.uk/guidance/ng3 [Accessed 16/01/2018].

541. Wilkins-Haug L, Horton JA, Cruess DF, Frigoletto FD. Antepartum screening in the office-based practice: findings from the collaborative Ambulatory Research Network. *Obs Gynecol*. 1996;88(4 Pt 1):483–9.

542. Moore LE. Screening, Diagnosis, and Management of Gestational Diabetes. In: Diabetes in Pregnancy. Cham: Springer International Publishing; 2018; 45–59.

543. Morampudi S, Balasubramanian G, Gowda A, Zomorodi B, Patil AS. The challenges and recommendations for gestational diabetes mellitus care in

India: A review. *Frontiers in Endocrinology*. 2017;8:56.

544. Gupta Y, Kalra B, Baruah M, Singla R, Kalra S. Updated guidelines on screening for gestational diabetes. *Int J Womens Health*. 2015;7:539.

545. McKenzie JE, Brennan SE. Overviews of systematic reviews: Great promise, greater challenge. *Systematic Reviews*. 2017;6(1):185.

546. Bero LA, Jadad AR. How consumers and policymakers can use systematic reviews for decision making. *Annals of Internal Medicine*. 199;127:37–42.

547. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. The nuts and bolts of PROSPERO: An international prospective register of systematic reviews. *Syst Rev*. 2012;1(1):2

548. Wormald R, Evans J. What Makes Systematic Reviews Systematic and Why are They the Highest Level of Evidence? *Ophthalmic Epidemiology*. 2018;25(1):27–30.