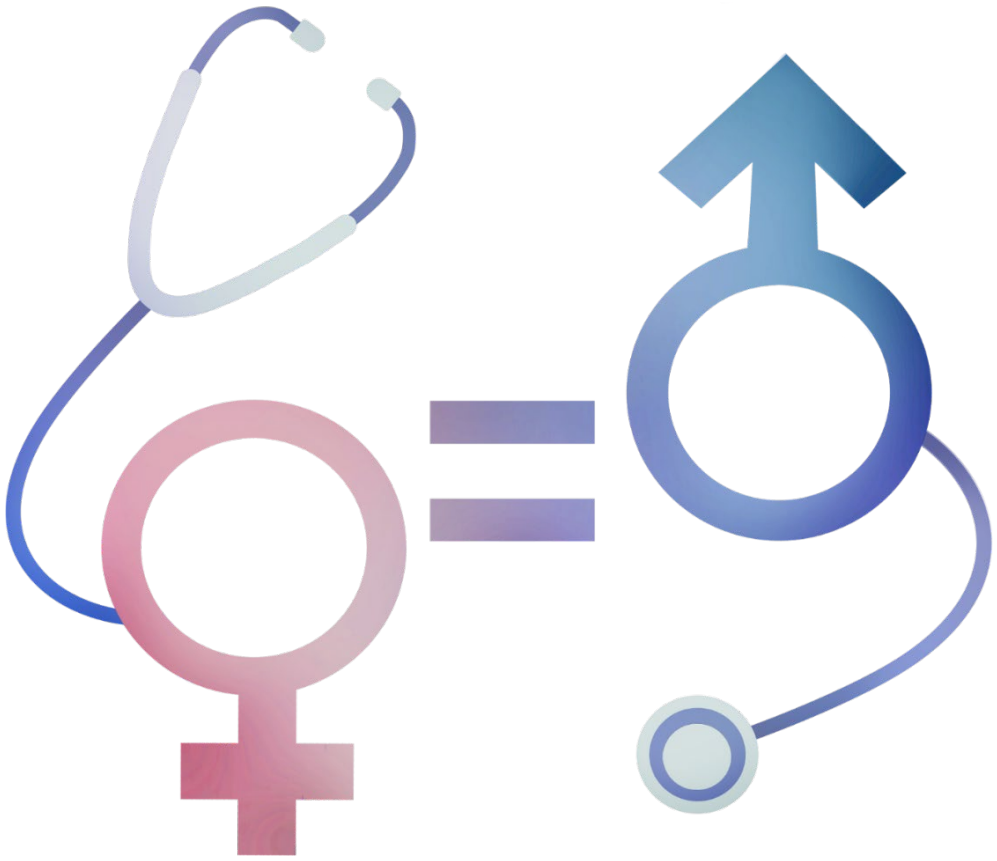
The role of sex in precision medicine: Novel methods in systematic reviews



Elena Stallings

Doctoral Programme in Health Sciences (2019-2021)

Universidad de Alcalá de Henares

Universidad de Alcalá

Doctoral Programme of Health Sciences

The role of sex in precision medicine: Novel methods in systematic reviews

Doctoral thesis presented by:

Elena Stallings

Doctoral thesis supervised by:

Dr. Javier Zamora

Dr. Agustín Albillos

Alcalá de Henares, 2021

Dedication of thesis

To my parents, Anne and Raymond, for always supporting me and being there to fall back on. Thank you for showing me what hard work is and allowing me to go on my own path.

Table of contents

Acknowledgements

First of all, I want to thank my thesis supervisors for all of their hard work and effort during these years. I want to thank my supervisor Dr. Javier Zamora for encouraging me to go down this PhD path. A few years ago, I never had any intention of doing a PhD and now look where I am. Thank you Javier for motivating me to continue on this track of constant learning. I could not have imagined having a better advisor and mentor for my PhD journey.

I would like to thank my co-supervisor, Dr. Agustin Albillos for his support and for helping with the logistics during the PhD.

I want to give a big thank you to Dr. Shakila Thangaratinam, Dr. John Allotey and the rest of the PregCOV19 team at the University of Birmingham and around the world. They welcomed me onboard the PregCOV19 team at the onset of the coronavirus pandemic. It was a very exciting project to be a part of and I learnt and grew a lot through it. Without this project, my thesis would not have shot to the finish line as quickly.

I would also like to give thanks to Karl Moons, Lotty Hooft and the rest of the Cochrane Prognosis Methods Group for inspiring me and sharing knowledge into the world of prognostic methods and prognosis systematic reviews.

To all of my colleagues in the Unit of Clinical Biostatistics at Hospital Ramon y Cajal, a big thank you. You have welcomed me wonderfully into the unit from the first day I arrived, even if in the beginning I probably didn´t understand even half of what you were saying. Don´t worry, my Spanish has improved with all those lunches and coffee breaks  You have been very important both academically and as my “work family”.

Alba, I moved to Barcelona and met you at just the right time! With our theses being so similar and in the same field it has been a pleasure being able to bounce ideas off of you and to have someone else to relate to in this PhD world, as you know just as well as me that it can get quite lonely at times. Thank you!

Saerah, thank you so much for helping with the graphic designs. They really help make the thesis come to life. Also, not only did you do the graphics, but you have been there from the start listening to me vent and let off steam. Thanks for cheering me on!

To my friends all over the world, you are too many to name individually, but you all know who you are and are always in my heart. Thanks for being there for me during these past few years.

Thank you to my parents, Anne and Raymond, and to my sister, Niamh, for always being there as my rocks and support. I always know I can count on you.

Finally, my gratitude and love to Henry for all his support. Thank you for helping to keep me sane, cooking nutritious meals and making me move from my desk at least a few times a day. Without you I may have moulded into one with my desk. Thank you!

List of Abbreviations

1. SDOH: Social determinant of Health
2. SR: Systematic review
3. LSR: Living systematic review
4. PF: Prognostic factor
5. SGBA: Sex and Gender Based Analysis
6. NIH: National Institute of Health
7. W.H.O.: World Health Organisation
8. COVID-19: Coronavirus disease 2019
9. PRISMA-P: Preferred reporting items for systematic reviews and metanalyses for protocols
10. CDSR: Cochrane Database of Systematic Reviews
11. ROB: Risk of Bias
12. NOS: Newcastle Ottawa Scale
13. QUADAS: Quality Assessment of Diagnostic Accuracy Studies
14. PROBAST: Prediction Model Risk of Bias Assessment Tool
15. QUIPS: Quality in Prognostic Studies
16. PROGRESS: Prognosis Research Strategy

List of figures

1. Introduction

1.1 Thesis summary

Title: The role of sex and gender in personalizing medicine: Using novel methods for systematic reviews

Author: Elena Stallings

Thesis supervisors: Javier Zamora and Agustin Albillos

Area of research: Health sciences, Universidad de Alcalá de Henares

Date of admission to PhD program: 29/03/2019

The thesis is presented in the form of a compendium of publications and includes four articles published in scientific journals with a good impact factor in general medicine. It also includes one article which is currently being considered for publication in a journal. The doctoral student was the first author of three of the published manuscripts and also of the article awaiting decision by the journal. The student was also the second author of one of the published manuscripts.

Publications arising from this Doctoral Thesis

1. **Sex as a prognostic factor for mortality in adults with acute symptomatic pulmonary embolism: Protocol.**

*López-Alcalde J, **Stallings EC,** Zamora J, Muriel A, Doorn S, Alvarez-Diaz N, Fernandez-Felix BM, Quezada Loaiza CA, Perez R, Jimenez D. Sex as a prognostic factor for mortality in adults with acute symptomatic pulmonary embolism. Cochrane Database of Systematic Reviews 2021, Issue 1. Art. No.: CD013835. DOI: 10.1002/14651858.CD013835.*

-The Cochrane database of systematic reviews

-Impact factor: 7.890 (2019)

-Publication date: 13th January 2021

2. Development and evaluation of a prognostic factor search filter for Ovid Medline.

-Journal of Medical Internet Research

-Impact factor: 5.03

-Publication date: Awaiting decision by journal

3. Sex as a Prognostic Factor in Systematic Reviews: Challenges and Lessons Learned

Stallings, E.; *Antequera, A.; López-Alcalde, J.; García-Martín, M.; Urrútia, G.; Zamora, J. Sex as a Prognostic Factor in Systematic Reviews: Challenges and Lessons Learned.* J. Pers. Med. *2021,* 11*, 441. https://doi.org/10.3390/jpm11060441*

-Journal of Personalized Medicine

-Impact factor: 4.433 (2019)

-Publication date: 21st May 2021

4. Clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy: living systematic review and meta-analysis.

Allotey J, **Stallings E*,** Bonet M, Yap M, Chatterjee S…..Mofenson L, Zamora J, Thangaratinam S; for PregCOV-19 Living Systematic Review Consortium. Clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy: living systematic review and meta-analysis. BMJ. 2020 Sep 1;370:m3320. doi: 10.1136/bmj.m3320. PMID: 32873575; PMCID: PMC7459193.*

**Joint first author*

-The BMJ (British Medical Journal)

-Impact factor: 30.223 (2019)

-Publication date: 1st September 2020

5. Collaborations in times of coronavirus: reflections on a living systematic review of covid19 in pregnancy.

Stallings E, *Allotey J, Van Wely M, Thangaratinam S, Zamora J. Collaborations in times of coronavirus: reflections on a living systematic review of covid19 in pregnancy. In: Collaborating in response to COVID-19: editorial and methods initiatives across Cochrane. Cochrane Database of Systematic Reviews 2020;(12 Suppl 1): [77-78].* *https://doi.org/10.1002/14651858.CD202002*

-The Cochrane database of systematic reviews

-Impact factor: 7.890 (2019)

-Publication date: 10th December 2020

1.2 Background to project

This thesis arose from the linking of two separate projects, one studying the sex differences in various clinical disorders and the other studying the effects of covid-19 in pregnancy. Thus, the overarching theme of the thesis relates to issues involving women's healthcare.

SEXCOMPLEX project

The original inspiration for this thesis came from the SEXCOMPLEX project (grant number PIE16/00050). This project aimed to investigate the influence of sex hormones and sex differences on the pathophysiology, clinical presentation and long-term course of clinical disorders of complex aetiology. Sex and gender medicine is a relatively new discipline, especially in Spain. Therefore, this project was considered a pioneering initiative.

This was the first project I was involved in while working on my PhD. The Cochrane systematic review protocol "Sex as a prognostic factor in patients with acute pulmonary embolism" was published early in 2020, whilst the systematic review itself is still being carried out.

Preg-COV19 project

During the covid-19 pandemic, I became involved with the preg-COV19 project, which is a large collaboration between researchers, mostly from the University of Birmingham and the World Health Organization (WHO). In this project we undertook living systematic reviews (LSR) involving pregnant and postnatal women at risk, suspected, and diagnosed to have COVID-19, and to synthesise the relevant evidence on prevalence, risk factors, mother-to-child transmission, diagnosis, and treatment of the disease. I have been leading the section of this project related to the identification of risk factors for pregnant women to develop COVID-19 and to develop adverse outcomes due to COVID- 19.

The living systematic review (LSR) "Clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy: living systematic review and meta-analysis" was the first article published from this project. The Cochrane short report "Collaboration in times of Coronavirus: Reflections on a living systematic review of COVID-19 in pregnancy" was also published from this project describing our experience of this LSR during the pandemic.

1.3 Precision medicine

The official definition of precision medicine, defined by the united states national research council is as follows: “The tailoring of medical treatment to the individual characteristics of each patient…to classify individuals into subpopulations that differ in their susceptibility to a particular disease or their response to a specific treatment. Preventative or therapeutic interventions can then be concentrated on those who will benefit, sparing expense and side effects for those who will not” (1). In other words, it is a term used to describe the treatment or interventions focusing on patients based on their individual characteristics, such as disease severity, genetic traits or sex and gender. With the advancements in technology scientists can now personalise drugs to make them more specific and effective for certain populations. This approach is highly effective in improving the prognosis of many patients. Precision medicine is a relatively new field of research as it only began to lift off in 2001 after the completion of the human genome project (2, 3).

Clinical areas where personalised approaches are promising include cancer treatments, cardiovascular diseases, psychiatric disorders, diabetes, and pain. For example, people with melanoma who have a certain mutation in the BRAF gene can be targeted with a specific drug that aims directly for BRAF (4). With this highly specific treatment, patient prognosis significantly increases. Breast cancer is another type of cancer that benefits from precision medicine. Breast cancer is generally treated according to biomarkers in the patient such as oestrogen receptors or HER2 gene status (5). This allows the drug to specifically target these biomarkers, thus improving patient prognosis. Within the same field of research, breast cancer, precision medicine has another application. It can also be applied as part of prevention methods. Screening women for variants of the BRCA1/2 genes, which give more than an 80% chance of developing breast cancer, can allow for preventative surgery and other intervention (6, 7).

Genetic testing has its advantages along with its downfalls as it can be useful in the prevention of diseases, but unfortunately there are a lot of diseases that cannot be cured. Currently, there are approximately 75,000 different genetic tests available on the market (3). This is where the ethical debate of genetic testing enters into precision medicine (8). As discussed above, in breast cancer it can be used as a method of prevention, but for other

diseases such as Huntington's disease if a genetic variation is detected there is not treatment available (9). Therefore, genetic testing in precision medicine is something that must be thoroughly considered from all aspects.

Sex plays an influential role in health and disease and using the characteristic of sex in precision medicine can improve outcomes for both males and females (10). Throughout the remainder of this thesis introduction I will discuss further in depth the differences between the sexes and how and why sex is an important factor to consider in precision medicine.

1.4 Sex as a determinant of health

1.4.1 Determinants of health

Many factors determine the health of individuals and communities. The range of factors that combine to influence health status are known as determinants of health. As can be seen in figure 1, they are categorised into five broad areas including policymaking, social factors, health services, individual behaviour and biology and genetics (11, 12). The term 'determinants of health' was coined in the 1970s and it refers to the factors that can have a significant positive or negative influence on health. Addressing sex and gender as determinants of health is a step in the right direction in personalizing medicine. The social factors category of the determinants of health is also known as the social determinants of health (SDOH). These factors include race, ethnicity, income and wealth, educational status, and sex and gender identification (11, 13).

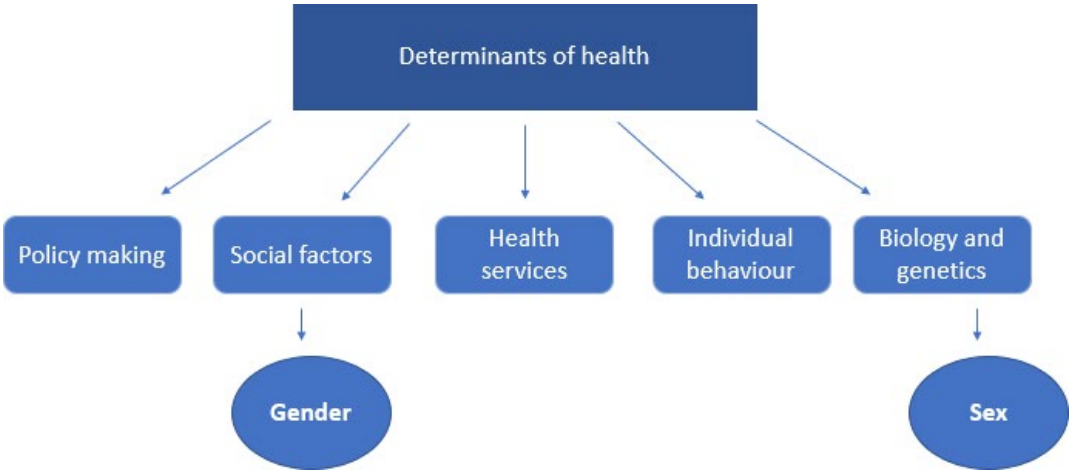


Figure 1: Determinants of health

Medical care has its limits of what it can do and without taking into account the determinants of health, key measures of health, such as life expectancy and infant mortality, will not improve (14). Studying the determinants of health can improve health outcomes for socially disadvantaged populations. For example, children who grow up in a neighbourhood that is socioeconomically disadvantaged are more likely to be obese. This is

due to the higher number of fast-food restaurants and lower availability of fresh produce which leads to poorer nutrition (14). To combat this, health promotion strategies must reach beyond just the clinical aspect and look further afield to investigate the deeper causes, which are often rooted in SDOH.

1.5 Sex and gender terminology

There is a need to increase literacy around the terms of sex and gender. Most people, including many researchers, do not realise that the terms of sex and gender are different and therefore use the words interchangeably (15). When researchers are writing an article, they prefer to use all of the possible terms to refer to sex (which they think includes gender), so as not to be repetitive. However, this is wrong and an incorrect usage of the words. This misuse leads to confusion when reading and analysing studies as it is not clear whether sex or gender is being studied. The majority of the time the authors will be referring to sex, but mistakenly use the word gender. A perfect example is the study by Lee 2017, where they refer to the "male gender" (16).

One of the causes of this problem could be that, unlike English, not all languages have different terms to refer to sex and gender. For example, in German there are no separate words (eg. Geschlecht in german) and the only distinction is made through the context of the sentence (17). Thus, the words sex and gender when translated from German, are done so identically or interchangeably. This also happens in the Greek language where the two terms cannot be distinguished easily as both sex and gender use the same word "φύλο" (18).

The term gender is actually derived from the term G-I/R which refers to gender identity role (18). The term "gender" has been spreading through the literature in recent years with people mistakenly using it as a synonym for sex. Many times, rats, mice, guinea pigs or isolated cells are described using the gender terms (gender, man, woman) (19). This is incorrect and not possible. Animals are not seen to have a gender as we cannot judge their genders based on the social and cultural roles played, only their biological sex.

Accurate use of sex and gender terminology is an important step to work through the complexities of these words.

1.6 Sex differences in health

The increase of studies on sex and gender differences and the development of sex and gender specific medicine represents a significant milestone in the progress of personalizing medicine. In order to continue progressing, we must understand that the concepts of sex and gender are interconnected, however differ in many aspects. The term distinction between sex and gender was introduced by the sexologist John Money in 1955 (17).

Definition of sex

Sex is the attribute of being male or female in organisms such as humans or other living beings that have the ability to reproduce sexually. It refers to a person's biological and physiological characteristics, such as reproductive organs and genetic differences (20, 21). The words male and female are used to describe the two different sexes.

Sex differences between males and females

Sex determination:

Sex is determined solely by the presence or absence of a Y chromosome. Genetic factors are used to define the sex of an individual. Females have 46 chromosomes made up of 22 pairs of autosomes and a pair of Xs, whilst males also have 22 pairs of autosomes, but in contrast to females, they have a single X chromosome and a Y chromosome that determines maleness (22, 23). These chromosomal factors and other factors further downstream in the process, such as the gonadal hormones, directly act on tissues producing sexual differences. Gonadal hormones are hormones produced by the primary reproductive organs, which are the testes in males and ovaries in females. The major hormones are oestradiol and progesterone from the ovaries and testosterone from the testes (22). Hormone levels differ between the sexes. Females have higher levels of progesterone whilst males have increased levels of testosterone. Males being more prone to physical aggression than females is a sex difference, which can be explained by the higher levels of testosterone in males. The most obvious physical differences between males and females are the features related to reproductive roles (including the endocrine systems, genital and breast organs). Other differences include the differentiation of muscle mass and height.

Figure 2: Physical sex differences between males and females

Sex differences in health:

Due to sex differences in many diseases, the best course of treatment for illnesses could differ between the sexes. Examples of diseases and illnesses influenced by sex include breast cancer, osteoporosis, and pulmonary embolism.

Breast cancer is one of the most common neoplasms worldwide. Although it affects both females and males, it is extremely rare in males. Breasts do not develop in the same way in males as in females, however males do have a small amount of breast tissue (24). Male breast cancer accounts for only 0.6% to 1% of all breast cancer cases (25). Thus, male breast cancer is much lesser studied than female breast cancers. Due to this, studies have shown that male patients with breast cancer have higher mortality rates than female patients (26, 27).

Osteoporosis is a growing worldwide health problem leading to complications such as bone fractures. Osteoporosis is a bone disease that causes bones to become weak and brittle, due to a decrease in bone density (28). It affects both sexes, but with different incidence rates and different ages (29). In general, females start losing bone density at an earlier age and at a faster rate than males. Osteoporosis is four times more common in females than in males aged 50 years or older (28).

Pulmonary embolism is a blockage of one of the pulmonary arteries in the lungs. This blockage can be caused by a blood clot which travels from another site in the body to the lungs (30). There has been a higher incidence of pulmonary embolism in females, compared

to males (31). Females also tend to be older than males and females appear to be at higher risk for short term mortality after acute pulmonary embolism (31, 32).

1.7 Sex inequities in health research

1.7.1 In the laboratory

Health care inequities exist throughout the system all the way from the laboratory to clinical trials to research positions. To help combat these inequities in the laboratory, in 2016, the US National Institute of Health (NIH) started requiring grant proposals to include sex as a biological variable in research (33). This means using both male and female animals, cells, and tissues in pre-clinical studies. However, this is not always feasible and easy. For example, when working with human cell lines it is often not feasible to include cells derived from both males and females as such lines were established, in some instances, from patients with rare disorders, many years ago (34).

In most areas of research, male animals are used more often than females and there are very few studies that use both males and females. The main reasons for this are cost and complexity. To include both male and female animals in an experiment means double the financial and human resource requirements (34). Including female animals in studies can also be seen as more complex due to their hormonal cycle. It has been suggested that in mouse studies four different groups of female mice should be included in order to account for the stages of the murine oestrus cycle. It is difficult to justify this approach due to the costs and complexity involved (34, 35).

1.7.2 In clinical trials

Just like a child is not a small adult and there is a special branch of medicine known as paediatrics for children, a female is not a mere copy of a male. There needs to be more focus on sex specific medicine to investigate these issues and a starting point is including females in all clinical trials. Similarly, like in the laboratories, for decades females had been mostly excluded from clinical trials. The principal reason for this was the same as with the female lab animals; they were considered to be too complex due to their monthly hormonal cycles. There are also a few other possible explanations for the exclusion of females in trials, including the fear of causing foetal damage in pregnant women or increased costs due to the higher number of participants that must be enrolled to cater for female involvement. However, in 1994 this began to change as the NIH mandated that females and minorities be included in NIH funded clinical research (35). This was one step in the right direction, but

even though females now make up half of clinical trial participants, there are many published studies which still fail to conduct a sex and gender-based analysis (SGBA). Due to their exclusion from many trials, women remain vulnerable to the adverse effects of pharmacological therapies. Thus, women tend to experience higher rates (almost double) of adverse events in comparison to men (36). Greater attention is needed in the area of studying the frequency and severity of adverse events amongst women trial participants.

131 million women worldwide give birth annually (37). These pregnant women are at an increased risk of infections and other illnesses due to alterations in their immune, respiratory, and cardiovascular systems during pregnancy. In relation to the current covid-19 pandemic, pregnant women have been almost entirely excluded from therapeutic and vaccine trials, even though many of these trials are just repurposing drugs that are already used in pregnancy (37). This is not unique to covid-19, this is the case with many illnesses when it comes to including pregnant women in trials. However, the pandemic has been highlighting the vulnerability of pregnant women when excluded from trials. The difficulty of including pregnant women lies in that they are a high-risk population, as there are always adverse outcomes in pregnancy. Therefore, even if the drug is repurposed and widely used for other conditions, if something goes wrong, the patient could blame the drug and sue the manufacturer or university running the trial.

Females tend to use more medications than males as they suffer more from chronic illnesses and they also usually pay more attention to their health (48). Unfortunately, females still represent a small percentage (22%) of the participants in the first phases of clinical trials (48). For this reason, it is necessary to test drugs in all populations (females, males, pregnant women, children) that they may possibly be used in. This ensures the appropriate dose regimens and will minimise the likelihood of adverse effects.

1.8 Systematic reviews

1.8.1 Introduction to systematic reviews

Publications in health research are continuously increasing which makes it difficult for clinicians to keep up to date with the best evidence. To address this challenge, systematic reviews (SRs) were developed (38). Another reason for SRs is to reduce research waste, which is a significant problem as many studies are carried out unnecessarily. SRs help prevent such waste by highlighting what is already significantly documented from multiple independent sources and what still needs to be researched (39). SRs are also key to making informed health choices and informing health policy guidelines. Since they synthesise all the studies on a given topic and appraise the quality of evidence of said studies, they are a highly significant form of evidence available to health care providers and policy makers (40).

A systematic review (SR) uses a systematic and explicit method to collect secondary data, critically appraise studies, synthesize findings, and answer a specific research question (41, 42). In 1753 James Lind published the first example of an SR in his paper "A treatise of the scurvy" (43). In this piece of work, he aimed to synthesize the evidence on scurvy in an unbiased manner. It included a critical and chronological view of what had already been published on the subject. He wrote about his search strategy for relevant material, just like reviewers do in today's SRs (43). Nowadays there is wide acceptance of Linds´ idea that decisions in healthcare and healthcare policy need to be informed by up to date systematic reviews. However, applying this principle is often proven difficult as reviews often reveal that more studies are needed, and the information sought is not yet available.

Although James Lind took an early step in the direction towards SRs, it was not until many years later that it was acknowledged that more attention was needed in the area of evidence synthesis. The Cochrane Collaboration was founded in 1993 and is responsible for regularly updating and maintaining the Cochrane Database of Systematic Reviews (CDSR). The CDSR is a journal with a collection of reviews and protocols maintained by various review groups (44). It includes 53 review groups and has approximately 30,000 experts who volunteer on SRs from around the world. The organisation was named after Archie Cochrane who was a field epidemiologist and made his name and reputation through his 1971 monograph "Effectiveness and efficiency. Random reflections on health services" (45). In

this monograph, Cochrane called for an international register of controlled trials and criteria to be drawn up to appraise the quality of published research (45).

1.8.2 Systematic review process

Once the review topic has been chosen, the next step should be to create a review protocol. This protocol should describe the rationale, the hypothesis, and the planned methods for the review (46). Protocols aid in detecting modifications made to the methods and also selective reporting. The Preferred Reporting Items for Systematic reviews and Meta-Analyses for Protocols 2015 (PRISMA-P 2015) exists to act as a guideline in developing the SR protocol. It consists of a 17-item checklist (46).

Figure 3: Systematic review process

All SRs, whether Cochrane or not, are based on a specific research question using the PICOS format. P is for the population of interest; I is for the intervention being studied or in the case of prognostic reviews for the index factor; C is for the comparator or in prognostic reviews comparator factors; O is for the outcomes of interest; S is for the setting of the studies. All of these elements must be accounted for when formulating the research question. SRs can be divided into different types depending on the research question. These classifications include effects of interventions, diagnosis, and prognosis.

After the question is defined, the next stage of a SR is searching for relevant studies. This means running searches in databases using key words and MeSH terms from the research question. Once the studies are retrieved from the searches, the next task is to screen them using the inclusion and exclusion criteria, which was set at the beginning of the review process. The eligibility criteria are based on the PICO question. The exclusion criteria are for studies that are withdrawn, duplicated, abstract on or of an unrelated topic. An example of a list of exclusion criteria is as follows: 1) withdrawn study, 2) duplicated study, 3) wrong population (patients not with investigated disease or health condition), 4) wrong intervention (not the correct treatment or other type of intervention that we are investigating in the SR). It best to write the list of exclusion criteria in certain order, for example, the way in which it is done above, as it makes it easier to discard irrelevant studies.

When the studies to be included in the review have been selected, the relevant data must be extracted. A data extraction template is drawn up either in excel or in a SR software such as Eppi- reviewer or Covidence. Data such as sex, gender, age of participants and details of the health condition, details of the intervention and numerical data are extracted. The numerical data will then be used in a meta-analysis if possible.

Then the quality of the studies must be assessed. It is just as important to assess the quality of the included studies as it is to analyse the data. Poorly conducted studies can include biases from the research methodology. Thus, these studies should be interpreted with caution. There are many different tools available to assess study quality and risk of bias. These include the following: Cochrane risk of bias (ROB 2), Newcastle ottawa-scale (NOS), Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2), Quality In Prognosis Studies (QUIPS) and Prediction model Risk Of Bias ASsessment Tool (PROBAST). The ROB 2 tool

should be used to assess quality and risk of bias in randomised clinical trials (RCTs) in Cochrane systematic reviews (or others). The NOS is used for non-observational studies of cohort and case control studies. The QUADAS-2 tool is used for diagnostic accuracy studies. QUIPS should be used in reviews of prognostic factors (47), while PROBAST needs to be used with reviews of prognostic model studies (48).

Once the data has been extracted and quality assessed the analysis can begin and a meta-analysis carried out if feasible. Data synthesis can be done narratively and if possible, through meta-analysis. If the populations are mixed, and there are enough studies, subgroup analyses can be carried out.

To deal with poor reporting in systematic reviews an international group of experienced authors developed the PRISMA (preferred reporting items for systematic reviews and meta-analyses) guidelines. The PRISMA guidelines consists of a 27-item checklist and a 4-phase flow diagram (49). These guidelines help to ensure consistent and reliable reporting in systematic reviews.

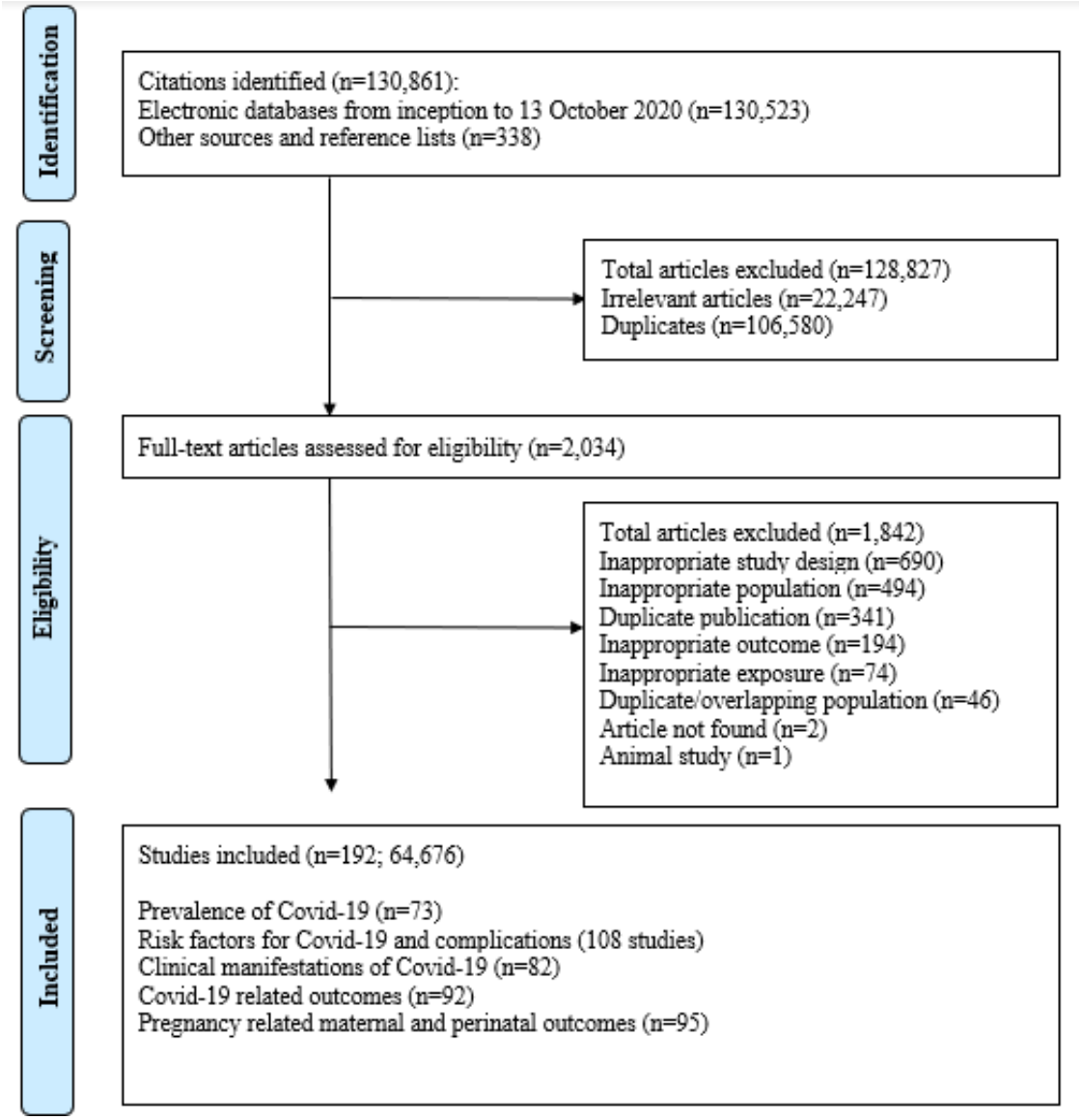


Figure 4: PRISMA flow diagram example taken from our SR of risk factors, clinical manifestations, maternal and perinatal outcomes in pregnant and recently pregnant women with suspected or confirmed coronavirus disease (COVID-19).

1.8.3 Novel types of systematic reviews

Prognostic systematic reviews

Prognosis research is a relatively new area of research and is vital for personalizing medicine. The PROGRESS (PROGnosis RESearch Strategy) framework was developed in 2013 by an international group of experts on prognosis research. It classifies prognosis research into four main types including: overall prognosis, prognostic factors (PFs), prognostic models

and predictors of treatment effect (50-53). Thus, prognostic systematic reviews, either of PFs or prognostic models, are also a novel area of research synthesis.

A prognostic factor is “any measure that, among people with a given health condition (that is, a start point), is associated with a subsequent clinical outcome (an endpoint)” (52). A prognostic model is “a formal combination of multiple predictors from which risks of a specific endpoint can be calculated for individual patients” (53). Prognosis SRs are formed a little differently than other types of SR's. Instead of forming the review question using PICO, they form the question using PICOTS. The main difference is found in the I, which is the index prognostic factor or model in prognostic SR's versus the intervention in other SR's. Another difference is in the C as in prognostic factor studies it is the comparative prognostic factors (confounders) and in other SR's it is the comparator interventions or placebos. Also, the T (timing) and S (setting) are added into the review question for prognostic SR's, which means the timing of measurement of the prognostic factor and the setting must be considered(54).

Systematic review speeds

A regular systematic review usually takes from one to three years to complete. However, there are times when we cannot wait one year to compile the evidence to inform guidelines and clinical practice. Times such as an epidemic or pandemic require evidence to be produced fast and efficiently. There are two relatively new types of SR's to deal with this: rapid reviews and living systematic reviews.

Rapid reviews are a form of evidence synthesis in which stages of the systematic review process are simplified or omitted to produce results in a timely manner (55). A rapid review can cut some corners by, for example, searching in three databases instead of five or six and conducting each step of the review by only one reviewer instead of the usual two reviewers independently. A rapid review can be carried out in approximately 6 weeks. However, due to the omission of certain parts of review stages, a rapid review may be susceptible to bias such as sampling bias, choosing studies bias, and obtaining accurate data bias (56).

A living systematic review (LSR) is a review that is continuously updated over a defined period of time. While being updated it maintains the same systematic review rigor and methodological quality. This living systematic review format should be used when the

current literature is lacking in the review topic, but new evidence will be appearing soon. In times of epidemics or pandemics research production can be greatly accelerated. A LSR allows access to high quality up-to-date syntheses of this new evidence (57). LSRs are associated with practical challenges, as continuous updating requires a lot of resources. A larger than normal review team might also be needed in the case of an LSR.

1.8.4 Search filters

As previously mentioned, an essential part of any systematic review is the extensive search for eligible studies. For many topics and areas of research, such as diagnosis or prognosis, studies are not indexed properly and often thousands of irrelevant studies are returned in the search. Search filters, or search strategies, are collections of search terms designed to aid in this process and to retrieve certain types of records topic. Search filters may be designed to retrieve references using a specific study design or by topic or by a different feature of the research question. They can be designed to maximise sensitivity (or recall) or to maximise sensitivity or precision (and reduce the number of irrelevant records that need to be assessed for relevance). Methodological search filters have been proven to be particularly effective in identifying intervention studies. A highly sensitive search strategy (HSSS) is widely used within the Cochrane Collaboration for identifying randomised trial studies in MEDLINE (58).

There are 4 common methodological key elements within search filter design: identification of a 'gold standard'; search term selection; evaluation of the search filter and validation (59). When identifying a gold standard there are two ways of doing so which include hand searching for gold standard studies or relative recall. The three most common sources of bias in filter development come from: 1) If systematic reviews were used to compile the reference set, these SRs must not have used a filter or terms for which they are trying to develop the filter 2) The choice of the gold standard records are subject to high bias if they are topic specific and not generalizable; 3) The validation of the filter should be carried out using a separate reference set (validation set) (60). In order to ensure quality search filters being developed, Glanville et al created the search filter appraisal checklist (61). It is composed of seven sections and aids researchers in the development process of their filters.

1.8.5 Sex and gender-based analysis

Integration and acknowledgement of sex and gender in all studies as well as in systematic reviews is integral for understanding the applicability of evidence. The importance of carrying out sex and gender-based analysis and properly considering sex and gender in research is slowly being recognized (62, 63). However, evidence suggests that sex and gender reporting in Cochrane reviews, among others, is inadequate (64, 65). Reviews do not use the terms consistently and often use them interchangeably. There are various ways for SR authors to consider sex and gender in their reviews. This includes carrying out a subgroup analysis based on sex and gender if sufficient evidence is available, or to at least plan for such an analysis in the protocol if there is not sufficient evidence. Authors could also note if the included studies have a sex or gender imbalance and how the included studies used the sex and gender terminology (64).

Tools exist to aid in carrying out sex and gender-based analysis and reporting. The Sex and Gender Equity in Research (SAGER) guidelines have been created to guide authors in preparing manuscripts with consideration for reporting on sex and gender information. The guidelines include procedures for reporting sex and gender in study design, data analyses, and results (66). Both the sex and gender of participants in studies should be reported, or at least the sex, if gender is not available (67). In order for those characteristics to be included, researchers and/ or clinicians should ask patients both their sex assigned at birth and the gender to which they currently identify with (67).

To enhance equity, proper sex and gender-based analysis in every study is needed. Without it, the evidence is more limited in its applicability.

2. Objectives

2.1 Primary Objectives

The primary objective of this PhD thesis is to highlight sex inequities in healthcare.

- Through an analysis of risk factors for pregnant women with Covid-19 in a living systematic review.
- By assessing the potential role of sex (being a male or a female) as a prognostic factor in patients with pulmonary embolism

2.2 Secondary Objectives

As a secondary objective I looked at developing novel methods to aid in these reviews of risk factors and prognostic factors:

- Development of a prognostic factor search filter for use in systematic review searches.
- Modification and discussion of tools and methods employed in a prognostic factor systematic review.

3. 1st Article: Sex as a prognostic factor for mortality in adults with acute symptomatic pulmonary embolism (Protocol)

This project started in 2017 but was put on hold various times due to Cochrane and then the Covid-19 pandemic. This SR “Sex as a prognostic factor in patients with symptomatic acute pulmonary embolism” was the only prognostic review that the Cochrane Collaboration accepted worldwide during 2017, which shows that we have the support of the “Cochrane prognosis methods group” and that our review, once published, will be considered as a “model” for other Cochrane reviewers. The aim of the SR is to synthesize the available evidence to determine whether sex is a prognostic factor for mortality in adults with acute symptomatic pulmonary embolism.

By publishing a Cochrane protocol before embarking on our SR we can be sure that all methods that we will employ are correct as it has been peer reviewed externally and by Cochrane editors. I am leading this SR and it is almost finished, so the preliminary results have been mentioned in the discussion and included in the appendices.

Cochrane Library

Cochrane Database of Systematic Reviews

Sex as a prognostic factor for mortality in adults with acute symptomatic pulmonary embolism (Protocol)

López-Alcalde J, Stallings EC, Zamora J, Muriel A, van Doorn S, Alvarez-Diaz N, Fernandez-Felix BM, Quezada Loaiza CA, Perez R, Jimenez D

López-Alcalde J, Stallings EC, Zamora J, Muriel A, van Doorn S, Alvarez-Diaz N, Fernandez-Felix BM, Quezada Loaiza CA, Perez R, Jimenez D.
Sex as a prognostic factor for mortality in adults with acute symptomatic pulmonary embolism (Protocol).
Cochrane Database of Systematic Reviews 2021, Issue 1. Art. No.: CD013835.
DOI: 10.1002/14651858.CD013835.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Prognosis Protocol]

Sex as a prognostic factor for mortality in adults with acute symptomatic pulmonary embolism

Jesús López-Alcalde[1,2,3,4], Elena C Stallings[2], Javier Zamora[2], Alfonso Muriel[2,5], Sander van Doorn[6], Noelia Alvarez-Diaz[7], Borja Manuel Fernandez-Felix[2], Carlos A Quezada Loaiza[8], Raquel Perez[9], David Jimenez[8]

[1]Cochrane Associate Centre of Madrid, Madrid, Spain. [2]Clinical Biostatistics Unit, Hospital Universitario Ramón y Cajal (IRYCIS). CIBER Epidemiology and Public Health (CIBERESP), Madrid, Spain. [3]Faculty of Health Sciences, Universidad Francisco de Vitoria, Pozuelo de Alarcón, Spain. [4]Institute for Complementary and Integrative Medicine, University Hospital Zurich, University of Zurich, Zurich, Switzerland. [5]Department of Nursing and Physiotherapy, Universidad de Alcalá, Alcalá De Henares, Spain. [6]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands. [7]Library, Hospital Universitario Ramón y Cajal (IRYCIS), Madrid, Spain. [8]Respiratory Department, Hospital Universitario Ramón y Cajal (IRYCIS), Madrid, Spain. [9]Respiratory Department, Hospital Universitario 12 de Octubre; Universidad Complutense Madrid, Madrid, Spain

Contact address: Jesús López-Alcalde, cochrane.madrid@ufv.es, suso77@gmail.com.

Editorial group: Cochrane Vascular Group.
Publication status and date: New, published in Issue 1, 2021.

Citation: López-Alcalde J, Stallings EC, Zamora J, Muriel A, van Doorn S, Alvarez-Diaz N, Fernandez-Felix BM, Quezada Loaiza CA, Perez R, Jimenez D. Sex as a prognostic factor for mortality in adults with acute symptomatic pulmonary embolism (Protocol). *Cochrane Database of Systematic Reviews* 2021, Issue 1. Art. No.: CD013835. DOI: 10.1002/14651858.CD013835.

ABSTRACT

Objectives

This is a protocol for a Cochrane Review (prognosis). The objectives are as follows:

To determine whether sex (i.e. being a male or a female) is an independent (i.e. autonomous) prognostic factor for predicting mortality in adults with acute symptomatic PE. See Table 1 for a formulation of the review question in population, index prognostic factor, comparator, outcome(s), timing, and setting (PICOTS) format.

BACKGROUND

Description of the condition

Venous thromboembolism (VTE) is a common cardiovascular disease that involves the formation of a blood clot (thrombus) in a vein (Bartholomew 2017). VTE can manifest as pulmonary embolism (PE) or deep vein thrombosis (DVT). Pulmonary embolism (PE) occurs when venous thrombi dislodge from their sites of formation and embolize to the pulmonary artery circulation system (Konstantinides 2014). About 50% of the patients with pelvic vein thrombosis or proximal leg DVT develop PE, which is oNen asymptomatic (Kearon 2012). About 90% of pulmonary emboli originate from the lower extremities, with most involving the proximal veins (Lee 2016). Acute PE is the most severe clinical presentation of VTE (Konstantinides 2014).

Diagnosis

Clinical recognition of PE is oNen inaccurate due to the signs and symptoms of PE being non-specific. Unfortunately, there is no test available that is sensitive and specific enough to confirm or exclude an acute symptomatic PE diagnosis. Therefore, in order to diagnose the disease it is necessary to combine clinical probability, D-dimer results, and imaging testing. In patients assessed for suspected PE, it has been shown that adherence to proven diagnostic algorithms improves patient prognosis (Kearon 2012; Konstantinides 2014; Roy 2006). The clinical probability of the patient having PE is the first thing to be assessed and with this information physicians can identify patients who require anticoagulant treatment whilst waiting on the results of the diagnostic tests. Clinical decision rules (CDR) have been proposed that combine items from the patient's clinical history, initial examination and sometimes from the chest x-rays or laboratory tests. The Wells and Geneva scores are the most extensively validated CDRs (Wells 2000).

Predisposing factors

PE is now recognised as a complex (multifactorial) disease. It involves both environmental exposures (e.g. clinical risk factors) as well as genetic and environmental interactions. When PE is associated with precipitating risk factors (such as surgery, cancer, trauma, immobilisation, pregnancy, or oral contraceptive use), it is classified as provoked or secondary (Kearon 2016a; Konstantinides 2014). When there are no precipitating factors, it is known as unprovoked (Kearon 2016a), spontaneous, or idiopathic. On the other hand, there are several conditions, only present in females, that are well-established risk factors for VTE. Relevant examples include pregnancy and the postpartum period (James 2006; Kujovich 2004; Marik 2008; Morris 2010), the use of oral contraceptives, which are the most common cause of thrombosis in young women (Peragallo Urrutia 2013; Stegeman 2013), and hormone replacement therapy (Cushman 2004).

Risk stratification

According to the short-term prognosis, PE can be classified as low-risk, intermediate-risk or high-risk (Merli 2017). High-risk PE is an acute PE with obstructive shock or systolic blood pressure (SBP) lower than 90 mmHg. Intermediate-risk PE is an acute PE without systemic hypotension (SBP ≥ 90 mm Hg), but with either right ventricle dysfunction or myocardial necrosis (Murphy 2018). If a PE has none of these severe features, it is called low-risk PE.

In patients with acute symptomatic PE, initial treatment decisions should be driven by their risk of short-term mortality and other adverse outcomes. When patients have a high-risk for PE associated complications (i.e. haemodynamically unstable patients) they need to be admitted to an intensive care unit (ICU) and be given early recanalisation (i.e. thrombolysis, percutaneous or surgical embolectomy) in addition to standard anticoagulation (Konstantinides 2014). In normotensive patients, there is a need for further risk stratification in order to diRerentiate patients who have a low-risk of early PE complications from those with an intermediate-high risk of PE complications. Low-risk PE patients may not need to be admitted to hospital, and thus could take advantage of either full outpatient anticoagulant therapy or a shortened hospital stay. Conversely, intermediate- or high-risk PE patients have a higher risk of PE complications due to preserved systemic arterial pressure; therefore, these patients could benefit from an intensification of therapy (Barrios 2018). Several issues should be taken into consideration during risk stratification, including the risk of bleeding from anticoagulants or thrombotic therapy, the risk of early venous thromboembolism recurrence and the consequences of these risks.

Treatment

In patients with high-risk PE, the primary cause of death is acute right ventricle failure. Therefore, the first stage of treatment is providing haemodynamic and respiratory support. The next step of treatment is usually anticoagulation for at least three months, as this can prevent premature death (Kearon 2016b; Konstantinides 2014).

Epidemiology

PE is relatively common worldwide, and its incidence is increasing (Alotaibi 2016; Belohlavek 2013). PE represents the most common cause of vascular death aNer myocardial infarction and stroke, and is the leading preventable cause of death in hospitalised patients (Tapson 2008). No exact worldwide epidemiological data are available, and most PE cases are undiagnosed and thus untreated (Cohen 2007). In addition, many countries, especially those classed as developing countries, lack population-based estimates for thrombotic conditions (Wendelboe 2016). However, the incidence of PE is estimated to be approximately 60 to 70 per 100,000 of the general population in Europe (Belohlavek 2013). In the US, the frequency of PE increased from 1998 to 2006, with the rate of PE detection nearly doubling without any change in mortality (Murphy 2017). With better technology, clinicians are better equipped to detect previously missed pulmonary emboli, but these are not necessarily clinically relevant (Doherty 2017; Wiener 2013). The US Centers for Disease Control and Prevention (CDC) estimate between 60,000 to 100,000 deaths per annum from PE in the US (CDC 2015), which represents 0.4% of all deaths in the country per annum (Murphy 2017). The mortality data from Australia and the UK show a similar frequency to the US, representing 0.2% (Australian Bureau of Statistics 2015), and 0.4% (British Lung Foundation 2015; ORice for National Statistics 2013) of all deaths, respectively.

Description of the prognostic factor

A prognostic factor is a characteristic in people with a given health condition (a start point) that is associated with a subsequent clinical outcome (an endpoint) (Hemingway 2013; Riley 2013). Therefore, prognostic factors distinguish groups of people with

a diRerent average prognosis (Riley 2013). The importance of prognosis research is increasingly recognised, as chronic health conditions and diseases are increasingly common and costly.

Health equity is the absence of avoidable and unfair diRerences in health (Welch 2020). Sex, gender, and sexual orientation may contribute to health inequalities and health inequities (Evans 2003; Welch 2020). 'Sex' refers to "the biological, genetic and physiological processes that generally distinguish females from males, while 'gender' refers to the roles, relationships, behaviours" and other traits that societies ascribe typically to women, men, and people of diverse gender identities (e.g. transgender) (CIHR 2012; Heidari 2016).

In this review, we will assess the potential role of sex (i.e. being a male or a female) as a prognostic factor in patients with PE. This review will not evaluate the association between gender or sexual orientation and the outcomes of patients with PE.

Health outcomes

We will assess the association between sex (being a male or a female) and mortality in patients with PE by evaluating the outcomes of all-cause mortality and PE-related mortality. All-cause mortality is death from any cause following the diagnosis of PE. PE-related mortality is defined as death confirmed by autopsy, or those deaths following a clinically severe PE, in the absence of any alternative diagnosis (Muriel 2014).

Why it is important to do this review

PE is the most common cause of vascular death aNer myocardial infarction and stroke, and the leading preventable cause of death in hospitalised patients (Tapson 2008). Therefore, the eRective management of PE is among the top priorities for improving survival rates in patients with thromboembolic disorders.

Prognostic factor research aims to identify factors associated with clinical outcomes in people with a particular disease or health condition (Hemingway 2013; Riley 2013). There can be diRerent uses of the evidence on individual prognostic factors:

1. to define modifiable targets for interventions to improve outcomes;
2. to build blocks for prognostic models; and
3. to determine predictors of diRerential treatment response (Riley 2013).

Prognostic factors are relevant to patient management as they help to stratify patients by diRerent risk groups, thus helping to reduce morbidity and mortality (Riley 2013). The identification of prognostic factors is a crucial step within the current drive towards personalised medicine (Riley 2013; Trusheim 2007).

Biological diRerences between the sexes can result in diRerential health risks, disease incidence, and health service needs (O'Neill 2014). Sex diRerences in the presentation and clinical course of conditions may dictate diRerent approaches to detection and management. Although sex diRerences in arterial disease have received substantial attention, there are still very few studies that have explored sex diRerences within VTE (Blanco-Molina 2014). There are inconsistent data in studies of patients with proven acute PE, in regard to the relationship between sex and adverse outcome rates. For example, in a study of 276,484 patients with acute PE, in-hospital mortality was significantly higher in females compared to males (Agarwal 2015). However, in another study, male patients were seen to have a higher risk of 30-day death compared to female patients (Aujesky 2005). Conversely, three other studies found no significant association between sex and prognosis (Jimenez 2010; Keller 2019; Panigada 2016).

Therefore, it is critical to determine if there are sex diRerences in the clinical course of patients treated for PE, as this may inform diRerent approaches for its detection, monitoring and management between males and females. The determination of the prognostic value of sex can be particularly important to support decisions when the benefit-risk balance of an intervention is not clear. Some examples identified in recent clinical guidelines (Kearon 2016b; Konstantinides 2019) are as follows:

- the choice of the optimal anticoagulant drug(s) and regimen (Kearon 2016b), particularly in patients with renal insuRiciency and creatinine clearance greater than 30 mL/min (Konstantinides 2019);
- the decision to administer reduced-dose thrombolysis and catheter-based reperfusion modalities in patients with intermediate- or high-risk PE (Konstantinides 2019);
- the criteria for selecting patients for early discharge (Konstantinides 2019).

In addition, the predictors of early PE-related death remain to be determined, and these predictors would be useful to identify possible candidates for reperfusion treatment among patients with intermediate-risk PE (Konstantinides 2019). To know the role of sex as a prognostic factor in patients with PE is also essential for professionals involved in drug discovery and development and for authorities responsible for the regulation and implementation of drug development programmes.

OBJECTIVES

To determine whether sex (i.e. being a male or a female) is an independent (i.e. autonomous) prognostic factor for predicting mortality in adults with acute symptomatic PE. See Table 1 for a formulation of the review question in population, index prognostic factor, comparator, outcome(s), timing, and setting (PICOTS) format.

METHODS

This protocol follows the methods proposed in other Cochrane prognosis reviews (Hayden 2014; Skoetz 2017; Westby 2018). Moreover, we followed the guidance provided in Riley 2019 and the general protocol template of the Cochrane Prognosis Methods Group (Cochrane Prognosis Methods Group 2019). Our protocol report adheres to the guideline for Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) Statement (Shamseer 2015). The review report will conform to the guidance of the PRISMA Statement (Liberati 2009), supplemented with the CHARMS-Prognostic factor checklist (Moons 2014). We will also follow the guidance for systematic reviews and meta-analyses of prognostic factor studies (Riley 2019).

Criteria for considering studies for this review

We have formulated the review question according to the PICOTS system. This format is based on the CHARMS checklist and informs

the objective and the eligibility criteria for the review (Debray 2017; Moons 2014; Riley 2019). See Table 1.

Types of studies

We will include any longitudinal study, randomised or non-randomised, investigating the prognostic significance of sex in adults with PE for predicting mortality. In practical terms, the following study designs will be eligible (Foroutan 2020): a) observational studies (e.g. cohort studies, case-control studies, or database linkage studies); and b) secondary analyses of experimental studies (randomised or non-randomised) providing evidence regarding prognosis. For an experimental study to be eligible, it must have used either the control group alone or the entire study cohort adjusted for the intervention.

We will exclude the following study designs, but we will report them in the 'Characteristics of excluded studies' table if the remaining eligibility criteria were met:

- **Descriptive studies describing the course of the condition/ disease**
- **Phase-1 exploratory prognostic studies ('exploratory studies'):** studies aimed at investigating all associations, usually in univariate analyses, of potential prognostic factors and outcomes. These studies are necessary to identify new prognostic factors, but they will not be eligible for our review because they provide the least conclusive information regarding the independence of a variable as a valid prognostic factor. Moreover, due to the high number of factors explored, exploratory studies oNen have widely varying results with common spurious associations, which may overstate their conclusions (Hayden 2008; Hayden 2014).
- **Other studies reporting univariate associations**
- **Phase-3 prognostic studies:** studies to understand prognostic pathways. We will exclude these studies because our review aims to determine the prognostic role of just one factor.
- **Cross-sectional studies**
- **Prognostic model studies:**
  - Studies to develop a prediction model (independently, if it reports any association of sex with any of our review outcomes)
  - Studies to validate a prediction model (that is, to validate the model in patient data not used in the development process)
  - Studies to evaluate the impact of a prognostic model on clinical practice and outcomes
- **Studies evaluating only the interactions between intervention and prognostic factors:** for example, a randomised controlled trial (RCT) or other study reporting only treatment eRect modification data

We will not exclude any study based on sample size, duration of follow-up, publication status, publication year or language. We will exclude studies that fulfil all our review eligibility criteria, but do not assess or report our outcomes of interest (see 'Selection of studies,' below).

Appendix 1 details the study design features (i.e. more than the reported study design labels) that we have considered to define study design eligibility.

Types of participants

We will include all adults, hospitalised or not, treated for acute symptomatic PE confirmed by objective testing, such as pulmonary angiography, ventilation/perfusion lung scan, or another validated measurement.

- **Adult:** person aged 16 years or older (in many settings, age 16 is when patients leave paediatric care and enter adult care)
- **PE:** defined as the dislodgement of venous thrombi from their site of formation and their embolization to the pulmonary artery circulation system (Konstantinides 2014)
- **Acute**: the follow-up should start no later than fiNeen days aNer diagnosis
- **Symptomatic:** at least chest symptoms must be present, such as dyspnoea or chest pain
- **Objective testing confirmation:** we will consider the following as valid examples of objective testing: high probability ventilation-perfusion scintigraphy; positive contrast-enhanced, PE protocol; helical chest computerised tomography for PE; or lower limb compression ultrasonography, positive for proximal DVT

We will include studies regardless if the patients were treated for PE or not, providing the diagnosis for PE was confirmed.

We will exclude studies with at least one of the following characteristics.

- Studies conducted in animals, cadavers or *in vitro*
- Studies conducted in females or males only, as they do not allow determination of the role of sex
- Studies conducted with healthy volunteers
- Studies where all the participants were children or adolescents (younger than age 16). We will exclude these studies because PE presents clinical and prognostic peculiarities in these age groups, as compared with in adults (Navanandan 2019; Zaidi 2017)
- Studies where the participants did not have confirmed PE
- Studies including only a subset of the participants relevant to our review question will not be eligible but will be listed in the 'Characteristics of excluded studies' table if they meet the remaining review criteria, but we are unable to extract the data of interest.

Types of prognostic factors

Index prognostic factor

We will include studies that assess the role of sex as a prognostic factor. Sex, categorised as female or male, relates to a set of biological attributes in humans and animals (Heidari 2016). In particular, sex refers to the biological, genetic and physiological processes that generally distinguish females from males, and is associated with features including chromosomes, gene expression, hormone function and reproductive/sexual anatomy (Heidari 2016). We will preferably include studies ascertaining sex by genotyping of a blood sample (Clayton 2016). However, we will accept any assessment of sex as provided by the study authors.

The concepts of sex and gender are distinct but interrelated (Doull 2010). However, this review will not assess the role of gender as a prognostic factor. Gender refers to the roles, relationships,

behaviours, relative power, and other traits that societies generally ascribe to women and men, as well as people of diverse gender identities (e.g. transgender persons) (Heidari 2016).

We acknowledge that 'sex' and 'gender' are poorly described and reported in published articles (Doull 2010; Lopez-Alcalde 2019; Runnels 2014; Welch 2017). If the reporting is unclear or incorrect, we will try to contact the authors for clarification. If no additional information is provided, we will generally assume that the study is considering sex, unless the authors explicitly state that they have evaluated the social aspect.

Other covariates

The focus of this review will be on the adjusted prognostic value of sex, that is, its prognostic eRect aNer adjusting for other covariates. Adjustment for the following key covariates, most taken from the scale of the Simplified PESI (sPESI) (Jimenez 2010) for mortality in patients with PE, will be of interest: age, history of cancer, current cancer, history of chronic cardiopulmonary disease, current chronic cardiopulmonary disease, heart rate, systolic blood pressure, and O_2 saturation. We will consider this list to assess the adjustment domain in the 'Risk of bias' tool (see 'Assessment of risk of bias in included studies').

Please note that we anticipate that we may modify the draN list, if and when we find new evidence that justifies any changes. Appendix 2 describes the process that we followed in selecting the covariates for adjustment.

Type of outcomes to be predicted, and timing

We will consider all-cause mortality and PE-related mortality measured at diRerent time points, all of them defined as primary outcomes. We provide the complete definition for each outcome according to the criteria adapted from Saldanha 2014.

Outcome	Definition	Specific measurement[a]	Specific metric[b]	Type of data[c]	Method of aggregation[d]	Timing		
						Time of prognostication[e]	Over what period the outcomes are predicted by these factors[f]	Minimum follow-up of the study participants to consider the outcome in the review
1. All-cause hospital mortality	Death from any cause occurring at the hospital	Any, as reported by the study authors	Value at a time-point	•Dichotomous •Event of interest: death	Proportion	•Index prognostic factor (sex): to be measured at the start of PE diagnosis •Other covariates: to be measured preferably at the start of PE diagnosis[g]	The longest follow-up provided by the study authors	None
2. All-cause hospital mortality at 30 days	Death from any cause occurring at the hospital during the first 30 days following the start of PE diagnosis	Any, as reported by the study authors	Value at a time-point	•Dichotomous •Event of interest: death	Proportion		30 days from PE diagnosis	All the participants must be followed for at least 30 days after PE diagnosis[i]
3. All-cause mortality at 90 days	Death from any cause occurring at the hospital or after discharge during the first 90 days following the start of PE diagnosis	Any, as reported by the study authors	Value at a time-point	•Dichotomous •Event of interest: death	Proportion		90 days from PE diagnosis	All the participants must be followed for at least 90 days after PE diagnosis[i]
4. Early hospital mortality (during the first 48 hours)	Death from any cause occurring at the hospital during the 48 hours following the start of PE diagnosis	Any, as reported by the study authors	Value at a time-point	•Dichotomous •Event of interest: death	Proportion		48 hours from PE diagnosis	All the participants must be followed for at least 48 hours after PE diagnosis[i]
5. All-cause mortality at one year	Death from any cause occurring at the hospital or after discharge during the first year following the start of PE diagnosis	Any, as reported by the study authors	Value at a time-point	•Dichotomous	Proportion		One year from PE diagnosis	All the participants must be followed for at least one year after PE diagnosis[i]

				•Event of interest: death			
6. PE-related hospital mortality	Death due to PE occurring at the hospital	Preferably, death confirmed by autopsy or death following a clinically severe PE, either initially or shortly after an objectively confirmed recurrent event, in the absence of any alternative diagnosis (Muriel 2014)[h]	Value at a time-point	•Dichotomous •Event of interest: death	Proportion	The longest follow-up provided by the study authors	None
7. PE related hospital mortality at 30 days	Death due to PE occurring at the hospital during the first 30 days following the start of PE diagnosis		Value at a time-point	•Dichotomous •Event of interest: death	Proportion	30 days from PE diagnosis	All the participants must be followed for at least 30 days after PE diagnosis[j]
8. Early PE-related hospital mortality (during the first 48 hours)	Death due to PE occurring at the hospital during the 48 hours following the start of PE diagnosis	Any, as reported by the study authors	Value at a time-point	•Dichotomous •Event of interest: death	Proportion	48 hours from PE diagnosis	All the participants must be followed for at least 48 hours after PE diagnosis[i]

Footnotes:

[a]The specific measurement or technique/instrument used to make the measurement

[b]The specific format of the outcome data from each participant that will be used for analysis (e.g., value at a time-point or change from baseline)

[c]Type of data: dichotomous, continuous, ordinal, counts and rates, or time-to-event (survival)

[d]How data from each group will be summarised (e.g., mean, percentage/proportion)

[e]The time point from which the outcome will be predicted

[f]The time-point that will be used for analysis

[g]We anticipate that the studies may use different starting points to define the follow-up. For example, from the recruitment, from the diagnosis of PE, from the allocation to the study arm, from the admission to the hospital, from the admission to the ICU or from the start of the treatment. We will preferably use the start of the PE diagnosis, but if this information is not available, we will consider the time as provided by the study authors. We will assess the impact of this decision by sensitivity analysis.

[h]However, we will admit any definition as provided by the authors

[i]Except for those participants that died or were discharged within this period

jExcept for those participants that died within this period

PE: pulmonary embolism

We attempted to select a 'Core Outcomes Set' for this review by searching the COMET initiative database (www.comet-initiative.org). We found one defined and published set, but this focused on trials in children and therefore was not addressing our review question. As a consequence, we have selected the outcomes listed above based on the following criteria:

1. the outcome must be critical from a patient perspective;
2. the outcome must support decision-making in the management of patients with PE.

We chose 'all-cause mortality' as a primary outcome because it has the greatest clinical relevance and is the most important outcome for individuals with PE. Furthermore, all-cause mortality is an objective endpoint and is not susceptible to be biased by the outcome assessor. We have also defined diRerent follow-up durations because we expect delayed eRects of PE.

We defined all mortality outcomes as binary variables (dead or alive), instead of using survival methods. We took this decision as the quality of life of patients in the hospital can be very poor, so patients who die in the hospital do not benefit if the duration of their survival is prolonged (Schoenfeld 2005); thus, the critical outcome is mortality and not patient survival. Secondly, some PE patients may be treated in the ICU: survival analysis should be avoided in the ICU context (Schoenfeld 2005) because Kaplan-Meier survival analysis assumes that censoring is non-informative; that is, it considers that the hazard of death remains unchanged when a censoring event occurs (Wolkewitz 2014). However, this assumption is incorrect in the ICU, as discharged patients are usually in a better health condition than patients who stay. The assumption that censoring is non-informative therefore generates artificially reduced survival plots (Schoenfeld 2005). There are statistical solutions to treat discharge as a competing event for death in the ICU (Wolkewitz 2014), but we believe that from a clinical point of view, the relevant outcome is mortality and not survival.

We will not consider all-cause mortality in the ICU or PE mortality in the ICU because they would only be useful if the majority of patients were still in the ICU at the time of analysis (Finkelstein 1994; Schoenfeld 2005). Thus, we will consider all-cause mortality to include all deaths at the hospital, inclusive of ICU deaths.

Setting

We will include studies involving patients with PE managed in any setting. Summaries of prognosis are not meaningful unless associated with a particular strategy for treatment so that prognostic studies can aid decisions about treatment. This implies that ideally, prognostic factors should be evaluated either in a cohort of patients treated the same way, or in a randomised trial (Altman 2001). We acknowledge that combining studies with patients with PE managed in any setting assumes that all the treatments are equally eRective and that the prognosis of patients is independent of the setting. This may not be true. Thus, the variation in the eRects of the treatments may be a relevant source of heterogeneity in this review. We also acknowledge that diRerences in hospital admission rates are likely to be related to the hospital- and country-specific availability of hospitals, admission policies, insurance systems, and other factors. Therefore, the patients admitted may not be homogenous. However, we consider that our synthesis will still provide relevant information. Moreover, we will try to explore the role of the region where the studies were carried out by subgroup analysis.

Search methods for identification of studies

Electronic searches

The Cochrane Vascular Information Specialist aims to identify all relevant studies regardless of language or publication status (published, unpublished, in press, or in progress).

The Information Specialist will search the following databases for relevant studies:

- the Cochrane Vascular Specialised Register via the Cochrane Register of Studies (CRS-Web);
- the Cochrane Central Register of Controlled Trials (CENTRAL) via the Cochrane Register of Studies Online (CRSO);
- Medline (Ovid MEDLINE® Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE® Daily and Ovid MEDLINE®) (1946 onwards);
- Embase Ovid (from 1974 onwards);
- CINAHL Ebsco (from 1982 onwards).

The Information Specialist has devised a draN search strategy for MEDLINE which is displayed in Appendix 3.This will be used as the basis for search strategies for the other databases listed.

The Information Specialist will search the following trials registries:

- The World Health Organization International Clinical Trials Registry Platform (who.int/trialsearch);
- ClinicalTrials.gov (clinicaltrials.gov).

Searching other resources

We will screen the reference lists of retrieved included trials and of systematic reviews on our topic.

We will contact experts on the topic (including authors of included studies, authors of systematic reviews) to identify any additional, unreported or ongoing studies.

We will handsearch documents of the Organization for the Study of Sex DiRerences (OSSD).

We will use the Web of Science database from Clarivate (clarivate.com/products/web-of-science) to track articles that have cited the primary reference for each study included in this review. We will also search the publisher web sites, PubMed (www.ncbi.nlm.nih.gov/pubmed) and the Retraction Watch database (www.retractionwatch.com) for retractions and comments related to references of included studies.

We will search for conference abstracts of major symposia from 2010.

1. Meetings of the OSSD: 5th edition (2010) to 14th edition (2019)
2. European Respiratory Society (ERS): 2010 to 2019
3. International Society of Thrombosis and Haemostasis (ISTH): 2010 to 2019
4. American Thoracic Society (ATS): 2010 to 2019
5. American Society of Hematology (ASH): 2010 to 2019
6. CHEST congress (CHEST): 2010 to 2019

7. Acute Cardiovascular Care (ACC): 2010 to 2019
8. European Society of Cardiology (ESC): 2010 to 2019

Data collection and analysis

Selection of studies

Two of six review authors (BF, CAQL, DJ, ES, JLA, RP), will independently check all titles and abstracts for inclusion. We will classify the titles and abstracts into four groups: 'obviously irrelevant', 'potentially eligible', 'potentially excluded' or 'unclear'. We will obtain the full-text version of those records classified as 'potentially eligible', 'potentially excluded' or 'unclear'. Two of six review authors (BF, CAQL, DJ, ES, JLA, RP) will independently assess the eligibility of each selected full-text article. We will resolve disagreements by consensus. In the case of disagreement, a third review author (one of AM, DJ, JLA or JZ) will serve as a neutral arbiter. There will be no restriction on language or date of publication of the papers.

If necessary, we will ask the study authors for clarification. If we cannot clarify the issues and we cannot exclude the study for any reason we will put these studies into 'awaiting classification'.

We will use the EPPI-Reviewer web-based soNware (Park 2018) to implement the selection process. We will complete a PRISMA flow chart to describe the selection process (Liberati 2009). We will also create tables describing the characteristics of excluded studies. These tables will detail the main reason for exclusion for studies that a reader might otherwise expect to see included in the review.

If there are multiple reports of the same study or data sets that overlap, we will collate them so that each study (not each report), is the unit of interest in the review. We will extract data from the data set with the largest sample size, most detailed results and the most appropriate follow-up.

We will exclude studies that fulfil all our review eligibility criteria, except the outcomes, i.e. studies in which no outcome of interest for the review was assessed or reported. For example, we will exclude a phase-2 prognostic study that aimed to determine whether sex is an independent prognostic factor for predicting the length of stay of patients with PE. We acknowledge that the exclusion of studies based on the reporting of the outcomes will hamper our evaluation of the risk of bias derived from selective outcome reporting. However, we anticipate that including all prognostic studies independently of the outcome reported will generate a workload that unaRordable to the team resources. On the other hand, we will not exclude studies based on their timing. For studies reporting several follow-ups for the same outcome, we will choose the most appropriate one for analysis.

Data extraction and management

Two of five review authors (BF, CAQ, DJ, ES and JLA) will independently extract data of each included study. We will use a consensus method to agree on the final extraction. A third review author (JZ or JLA) will intervene if there are disagreements. A third review author (AM) will check the accuracy of the numeric data in the review. We will try to obtain crucial missing information or clarification from study authors or organisations. If necessary, we will translate the included reports. We will examine any relevant retraction statements and errata for relevant information regarding each included study.

We will use the CHARMS-PF guidance to extract data (Riley 2019). This form adapts the original CHARMS checklist (CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies) for prognostic factor studies (Moons 2014), based on the experience of conducting systematic reviews of prognostic factor studies (Riley 2019). We will extract key information from each primary study.

- Dates, country and setting in which the study was conducted
- Study design
- Eligibility criteria
- Participants details
- Pulmonary embolism diagnostic criteria
- Treatment details
- Details of the prognostic factor:
  - Sex definition
  - Sex measurement (for example, self-reported or by genotyping of blood sample)
- Definition of start points
- Outcomes reported
- For each review outcome, we will extract the information as described in the 'Types of outcome measures' section (Saldanha 2014)
- Duration of study follow-up
- Type of analysis:
  - Explanatory/confirmatory
  - Presence of a valid study registration
  - Presence of a valid protocol
  - Logistic regression/Cox regression
  - Adjustment done for other prognostic factors (if any) to estimate the prognostic association
  - The covariates used in the adjusted analysis
  - Age limit used to dichotomise age or other variables (if adopted)
- Association measures for the prognostic factor and each review outcome:
  - Type of association measure, e.g. odds ratios (ORs), risk ratios (RRs) hazard ratios (HRs)
  - Confidence interval (CI), variance and standard error (SE)
  - Details on any adjustment factors used
  - We plan to extract the unadjusted and the adjusted measure of association (if available)
- Methods used to handle missing data
- Attrition:
  - Loss to follow-up
  - Reasons
- Information to assess applicability
- Information to assess risk of bias
- Data needed to perform the meta-analyses, such as the estimates, and their corresponding standard errors or confidence intervals.

We will use the online EPPI-Reviewer soNware (Park 2018), to build the data extraction templates and extract the data. We will pilot the data extraction form with five studies for usability. We will summarise the information retrieved in a table detailing the characteristics of each included study.

Transformations of reported data and assumptions made

The two key elements that must be extracted from each primary study to estimate the eRect of a prognostic factor with a meta-analysis are the prognostic factor eRect estimate and its precision (that is, the SE or the 95% CI) (Riley 2019). If needed, we plan to undertake transformations of reported data to use data from as many studies as possible. Thus, we will attempt to restore the missing information and to standardise the data to our desired format.

To convert the data, we plan to follow the guidance described in Westby 2018 ('Measures of association' section), Riley 2019 ('Methods to restore the missing information upon data extraction' section), and the *Cochrane Handbook* Section 7.7 (Higgins 2011) and Section 12.5.4 (Schünemann 2011). If needed, we will perform the conversions with the calculator available in Review Manager 5.3 (Review Manager 2014). Before concluding that the necessary information to calculate a prognostic association is not available, we will consult Cochrane Prognosis Methods.

We will present the associations consistently, that is, associations above one will indicate a worse prognosis for women (higher mortality). If necessary, we will recalculate the associations to be in the same direction.

As stated below in 'Type of measure of association', we will attempt to consider the OR and its 95% CI as the common measure of prognostic association in all the studies. We will also try to convert the combined OR to an absolute risk reduction (ARR) to facilitate its interpretation. To compute the ARR from an OR, we will use the Absolute Risk Calculator provided by the Health Information Research Unit at McMaster University (hiru.mcmaster.ca/AbsoluteRiskCalculator). We will also obtain the lower and upper limits of the CI 95% of the ARR by applying the same formula to the lower and upper confidence limits of the adjusted OR.

Assessment of risk of bias in included studies

Tool to assess the risk of bias

We will use the QUIPS (Quality In Prognosis Studies) tool to assess the risk of bias (RoB) (Hayden 2013; Riley 2019). The tool has six domains (with signalling questions related to each domain that can inform judgments of RoB in prognostic research):

1. Study participation
2. Study attrition
3. Prognostic factor measurement
4. Outcome measurement
5. Adjustment for other prognostic factors
6. Statistical analysis and reporting

For each study, we will label the six domains for each prognostic factor-outcome combination. Therefore, we will assess the RoB per outcome. We will make a judgement for each domain choosing one of the following options (Riley 2019):

- Low risk: the criterion is adequately fulfilled in the study
- High risk: the criterion is not adequately fulfilled in the study
- Moderate risk: there is not suRicient information provided to be able to make a clear judgement on the RoB.

We will detail and justify judgements on RoB in a 'Risk of bias table' for each included study. We will also generate RoB graphs and figures.

Overall assessment of the risk of bias and incorporation into analyses

All the tool domains will be 'key domains' for RoB. Thus, we will summarise the RoB for each prognostic factor-outcome combination in two diRerent manners, 'within each study' and 'across studies' (Higgins 2011).

	Interpretation	Risk of bias for each prognostic factor-outcome combination	
		Within each study across different domains	Across studies
Low risk of bias	Plausible bias unlikely to seriously alter the results	Low risk of bias for all key domains	Most information is from studies at low risk of bias
Moderate risk of bias	Plausible bias that raises some doubt about the results	Moderate risk of bias for one or more key domains (and no domain is rated as high risk)	Most information is from studies at low or moderate risk of bias
High risk of bias	Plausible bias that seriously weakens confidence in the results	High risk of bias for one or more key domains	The proportion of information from studies at high risk of bias is sufficient to affect the interpretation of results

We will describe the RoB among the included studies in the results section. Also, we will consider the RoB across studies for each prognostic eRect estimation, as part of the determination of the quality of the evidence with the GRADE system (Guyatt 2011).

We will meta-analyse studies independently of their RoB, but we will explore the eRect of this decision by carrying out a sensitivity analysis.

Procedure to assess the risk of bias

Two of five review authors (AM, BF, DJ, ES, JLA) will independently appraise all the domains of the QUIPS tool for each included study. We will agree on the final judgements for each domain via consensus. A second review author (JLA or JZ) will intervene if there are disagreements. A third review author (AM) will check the final decisions. If the study report does not provide information for a domain, or this information is not clear, we will follow a three-stage process. First, we will consult other publications that may have used the same data set (which is frequent in prognostic studies based on large existing cohorts) (Riley 2019). Second, if we cannot solve the doubt, we will attempt to contact the authors for clarification. Third, if we do not clarify the issue, we will make judgments based on the available information and the consensus between the review authors. We will not be blinded to study authors, institution or journal of publication.

As suggested in Riley 2019, we will define in advance criteria to assess the signalling items and domains for our specific review question, as this will probably facilitate reproducibility in our judgements. In particular, we will use our data extraction template in EPPI-reviewer (Park 2018) to define the following key aspects, many of them already pre-defined in this protocol:

1. Study participation
2. Attrition
3. Definitions of suRiciently valid and reliable measurement of the index prognostic factors (see 'Types of prognostic factors')
4. Definitions of suRiciently valid and reliable measurement of the outcomes (see 'Types of outcomes')
5. The core set of other (adjustment) prognostic factors that are deemed necessary for the primary studies to adjust for (see 'Comparator prognostic factors' and Appendix 3).

Measures of association to be extracted

Type of measure of association

We will attempt to consider the OR and its 95% CI as the measure of prognostic association in all the studies. We have chosen this measure because we anticipate that the OR will be the most common measure used in the primary studies: it is the only measure for dichotomous outcomes that can be estimated from case-control studies, and OR is obtained when logistic regression is used to adjust for confounders (Reeves 2011). If results from multivariable analyses in the primary studies are reported in another form, we will attempt to convert these to ORs at a particular time point (See 'Data extraction and management' above). If we find a study reporting a hazard ratio (HR), we will not attempt to convert the HR to OR and we will perform meta-analysis based on HRs.

Adjusted prognostic effect estimates

We will extract the adjusted measure of association for each study and prognostic eRect estimate. We acknowledge that the studies providing the adjusted prognostic eRect of a particular factor can diRer in the set of adjustment covariates or in the cut-oR used to dichotomise the covariates. This makes the interpretation of the meta-analysis challenging (Riley 2019). We agree that age, history of cancer, history of chronic cardiopulmonary disease, heart rate, systolic blood pressure, and O_2 saturation will be the core set of adjustment factors for each review outcome.

If a study provides adjusted estimates but not adjusted for our minimal set of adjustment factors, we will meta-analyse the study, but we will 'penalise' the estimate as part of the RoB assessment (we will assess the impact of this decision by sensitivity analysis). If less than four of the key factors are adjusted for in the study, it will be assessed as high risk of bias in the adjustment domain of the risk of bias tool. However, if four or more of the key factors are adjusted for, the study will be defined as low risk of bias for this domain. If the study only adjusted for PESI/sPESI but did not detail for which individual factors they had adjusted, we will mark the RoB domain as moderate.

If the same study presents diRerent estimates for the same outcome, each of them adjusted for diRerent factors, we will extract for meta-analysis the estimate that has adjusted for the maximum number of our key covariates. If there are several estimations, all of them having adjusted for our key covariates, we will consider the estimate adjusted for more of our key covariates in total. We assume that this will minimise the risk of confounding bias in the estimation.

Concerning the dichotomisation of our key covariates, we will accept any cut-oR used by the primary authors. We acknowledge that diRerent cut-oRs for the same covariate will occur among studies and that this situation may aRect the prognostic estimate obtained in our review. Thus, we will perform sensitivity analysis to assess the impact of our decision by excluding studies that have adjusted for PESI (or PESI simplified) measured as a categorical variable.

Direction of the associations

We will present the associations consistently, that is, associations above one will indicate a worse prognosis for women (higher mortality). See 'Data extraction and management' for how we will recalculate associations to be in the same direction.

Unit of analysis issues

The prognostic factor (sex) and outcome (mortality) will both be considered at the patient level. Thus, we do not anticipate that there will be unit of analysis errors (Deeks 2011). However, in the case that we find any unit of analysis error which cannot be handled, we will meta-analyse the estimation, but take into account the associated RoB as part of the domain 'Statistical analysis and reporting' of the RoB assessment.

Dealing with missing data

We plan to include all the studies that investigated the role of sex as a prognostic factor in patients with PE regardless of the presence of missing data. We plan to contact study authors to request missing data. For all the review outcomes we will consider the follow-up to start aNer PE diagnosis. However, if the study reports only the follow-up from other time points, such as the start of the treatment or the start of the symptoms, we will use this data for the analyses.

We acknowledge that the presence of diRerent strategies in the included studies to handle missing participant data may introduce heterogeneity in the results. We plan to repeat the meta-analysis to assess the eRect of excluding studies that did not adopt multiple imputation techniques to deal with missing values.

Assessment of heterogeneity

We expect that heterogeneity between the included studies will be common (Riley 2013). We plan to synthesise all the associations found about the prognostic eRect of sex with mortality outcomes in patients with PE. We do not expect to meta-analyse the prognostic within relevant subgroups. However, we will assess the presence of heterogeneity following a two-step process.

- Assessment of clinical and methodological heterogeneity

We plan to meta-analyse all the studies regardless of their clinical characteristics and their study design (as we plan to evaluate a potential association and not causation). However, we will attempt to use subgroup analyses to explore possible causes of heterogeneity that are clinical or methodological (see 'Subgroup analysis and investigation of heterogeneity').

- Assessment of statistical heterogeneity of the results

We will assess the statistical heterogeneity across the meta-analysed results considering the following factors:

- Identification of heterogeneity
  * Visual inspection of the prognostic eRect estimates: we will display graphically the results of clinically and methodologically comparable studies with forest plots, and we will assess the possibility of statistical heterogeneity visually.
  * The Chi^2 P value: we will use the chi-squared test for identifying heterogeneity (Chi^2 P value < 0.10 will be significant) (Deeks 2011).
- Quantification of heterogeneity
  * Use of the I^2 statistic: the IZ statistic describes the percentage of the total variation across studies that is due to heterogeneity rather than sampling error (chance) (Higgins 2003). We will define an IZ estimate greater than or equal to 50% and accompanied by a statistically significant Chi^2 P value as evidence of substantial statistical heterogeneity (Chapter 9. *Cochrane Handbook*) (Section 9.5.2; Deeks 2011).
  * Use of the Tau^2 and the 95% prediction interval: we will also measure the heterogeneity using the estimate of between-study variance (Tau^2) in a random-eRects model, as reliance on the IZ statistic in assessing heterogeneity may be misleading (Rucker 2008). We will also report an approximate 95% prediction interval indicating the potential true prognostic eRect of a factor in a new population (Riley 2011; Riley 2019).

We will try to explain heterogeneity by conducting subgroup analyses (if the number of studies found is suRicient). See 'Subgroup analysis and investigation of heterogeneity'.

Assessment of reporting biases

We plan to examine the presence of 'small-study eRects', that is, the presence of a systematic diRerence in prognostic eRect estimates for small studies and large studies (Riley 2019; Sterne 2011). We will assess publication bias for each meta-analysis (if the meta-analysis includes at least ten studies) by:

- Visual inspection of the funnel plot: we will interpret as a strong potential for small-study eRects the apparent asymmetry of the funnel plot with a higher proportion of smaller studies in one particular direction (Riley 2019).
- Use of test for asymmetry; we will also test for asymmetry at the 10% level using the Peters' test for ORs (Peters 2006; Riley 2019; Sterne 2011).
- Interpretation of small-study eRects: we will interpret the presence of small-study eRects with caution as it may be due chance, heterogeneity, publication bias and selective reporting. All these situations are frequent in prognosis research (Kyzas 2007a; Kyzas 2007b; Riley 2019) and it is diRicult to disentangle them (Riley 2019). We will consider that small-study eRects are caused by heterogeneity rather than by publication bias if the smaller studies used fewer adjustment factors for the analysis. This may explain why these small studies presented larger prognostic eRects.

Data synthesis

Data synthesis and meta-analysis approaches

We plan to combine the results from individual studies in a meta-analysis to provide a pooled prognostic eRect estimate only if the following criteria are met:

- there are enough studies (at least two studies);
- the studies are suRiciently homogeneous:
  * the studies are clinically similar in terms of population and sex measurement;
  * the studies are methodologically similar: we will consider that all phase-2 prognostic factor studies are methodologically comparable studies to determine a prognostic association (independently of their design). However, we plan to explore if the study design explains heterogeneity (see subgroup analysis);
  * the outcomes are measured at similar follow-up points;
  * the outcomes are measured with similar measurement tools;
  * the studies have the same type of prognostic eRect estimate measure, that is, an OR and 95% CI (or, at least, this information can be obtained);
  * the prognostic eRect estimate has been adjusted for at least one factor (independently of the factors considered for adjustment). If a study presents the unadjusted measure only (raw data), we will not include this data for analysis.

Statistical model for meta-analysis

We will not assume a common (fixed) prognostic eRect of sex on mortality. We anticipate that the prognostic eRect estimates will vary among studies due to several reasons, in particular, due to the presence of diRerent study populations, designs, prognostic eRect measures (OR and RR), unavailability of SE, diRerent time points and measurement of the outcomes, various sets of adjustment factors and due to missing data (Riley 2019). Thus, we will assume that there is not a single underlying prognostic eRect to estimate and therefore the heterogeneity among the study eRects cannot be explained by chance alone and follow a distribution across studies (Deeks 2011). However, we still consider that the underlying clinical questions will be similar enough and pooling will be meaningful if the extra uncertainty due to that heterogeneity is adequately represented (Cornell 2014). Therefore, we will apply

a random-eRects model, which is an approach for meta-analysis that incorporates study-to-study variability beyond what would be expected by chance (Cornell 2014), and that allows for unexplained heterogeneity across studies (Riley 2019).

The DerSimonian and Laird (DL) method is the most commonly used random-eRects model, and is available in Review Manager 5 statistical soNware (Review Manager 2014). However, this method has long been challenged (Veroniki 2019) because it produces a 95% CI that is too narrow (and P values that are typically too small) under two circumstances that this review will probably meet: a small number of studies and the presence of substantive diRerences among study estimates (Cornell 2014; Riley 2019). Therefore, we plan to use the Hartung-Knapp-Sidik-Jonkman (HKSJ) method for random eRects meta-analysis, as it has shown to consistently result in more adequate error rates than the DL method, especially when the number of studies is small (IntHout 2014). However, we will take into consideration that even with the HKSJ method, extra caution is needed when there are less than six studies of very unequal sizes (IntHout 2014). We plan to use the Cochrane Review Manager 5 soNware (Review Manager 2014) for organising the text of the review. We will use the 'metareg' command in Stata to perform the meta-analysis with the HKSJ method (Harbord 2008).

We plan to combine results in a meta-analysis independently of their RoB and the factors considered for adjustment. However, we will assess the impact of this decision by sensitivity analysis. We also plan to evaluate the influence of the statistical model used to pool data on the prognostic eRect estimate (see 'Sensitivity analysis').

If we find relevant unexplained statistical heterogeneity, we will still meta-analyse the data, but we will downgrade the certainty of the prognostic eRect estimate as part of the GRADE assessment (see below). If we detect that the meta-analysis is inappropriate for other reasons, we will not combine results. However, we will undertake a narrative analysis of studies, providing a descriptive presentation of results with supporting tables.

If there are enough studies, we will follow the guidance in Riley 2019, which states that if restricting the analysis to the subset of studies at low RoB resolves previous issues of small-study eRects, then it gives even more credence to focus conclusions on the meta-analysis results based only on the studies with low risk of bias.

Presentation of results

For the meta-analysis of each prognostic eRect estimate we plan to provide the pooled estimate based on the random-eRects approach (the average prognostic eRect of sex), its Hartung-Knapp 95% CI, the I^2, the estimate of Tau^2 (between-study variance) and the 95% prediction interval for the prognostic eRect in a single population, as done in Westby 2018 and suggested in Riley 2011 and Riley 2019.

An OR larger than one will suggest that female sex is associated with higher odds of mortality. For relative eRects, we will define the clinical importance of the observed prognostic associations as follows: small: OR < 1.2; moderate: OR between 1.2 and 2; large: OR > 2. For absolute risk diRerences, we will consider an absolute risk of 5% (50 per 1000) as the threshold for identifying an important prognostic factor.

The meaning of OR is diRicult to understand (Boissel 1999; Deeks 2011; Sackett 1996; Sinclair 1994). Moreover, ORs tend to be interpreted as RRs by clinicians (Deeks 2000). This can be misleading, as the OR is similar to the RR for outcomes with a low incidence (< 10%), but the OR exaggerates the eRect when the incidence of the outcome increases (Zhang 1998). This may be the case in our review, because all-cause mortality in patients who are treated for PE is 30% in high-income countries (Klok 2010; Ng 2011), while the PE-related mortality in patients treated for PE is estimated between 2% and 10% (Belohlavek 2013; den Exter 2013; Konstantinides 2016). To facilitate interpretation of the results, we will undertake each meta-analysis based on ORs, and express the meta-analysis as an OR. However, the 'Summary of findings' table(s) will also present illustrative comparative risks and the absolute risk reductions (ARR) for the eRect of the prognostic factor. To calculate the ARRs we will consider a range of diRerent prevalences of the prognostic factor (being a female) and diRerent risks of the outcome in the entire cohort. See 'Transformations of reported data and assumptions made' for details on the formula we will use to convert the data.

Subgroup analysis and investigation of heterogeneity

We plan to investigate if the following prespecified factors can explain heterogeneity If there are at least two studies per subgroup:

- Assessment of clinical heterogeneity
  - **Mean participants' age:** less than 45 years versus older than 45
  - **Setting:** patients managed at the hospital versus patients managed at the outpatient setting
  - **Measurement of the prognostic factor (sex):** measured at the start of PE diagnosis versus measured at the start of PE treatment
  - **Treated for PE:** participants treated for PE versus participants not treated for PE. It is estimated that in Europe around 30% of PE-related deaths occur before receiving any treatment for PE (Belohlavek 2013). Moreover, these numbers can be even higher in low resource settings (Wendelboe 2016)
  - **Reperfusion treatment for PE:** patients who received reperfusion treatment for PE (thrombolysis or surgical embolectomy) versus patients who did not
  - **Haemodynamic status:** stable versus unstable (as defined by the study authors)
  - **Geographic region:** Europe and North America versus other regions
- Assessment of methodological heterogeneity
  - **Study design:** experimental studies versus cohort studies versus case control studies
  - **Study design:** experimental studies versus observational studies
  - **Risk of bias:** studies with high RoB versus studies with low or moderate RoB

Sensitivity analysis

We plan to undertake the following sensitivity analysis if there are suRicient studies.

- We will repeat the meta-analysis to assess the eRect of including only studies with prospective assessment of outcomes.

- We will repeat the meta-analysis to assess the eRect of including only observational studies.
- We will repeat the meta-analysis to assess the eRect of excluding the studies that the Index prognostic factor (sex) was measured at the start of PE treatment (instead of diagnosis).
- We will repeat the meta-analysis to assess the eRect of excluding studies that have used routinely collected hospital administrative databases.
- We will repeat the meta-analysis to assess the eRect of excluding studies that have adjusted for PESI (or PESI simplified) measured as a categorical variable.
- We will repeat the meta-analysis to assess the eRect of excluding studies with high RoB.
- We will repeat the meta-analysis to assess the eRect of excluding studies that have provided an adjusted estimate but that did not adjust for all our core set of covariates.
- We will repeat the meta-analysis to assess the eRect of excluding studies that did not adopt multiple imputation techniques to deal with missing participant data.
- We will repeat the meta-analysis to assess the eRect of using a fixed-eRect model.
- We will repeat the meta-analysis based on the DL method.

Conclusions and summary of findings

We will assess the certainty of the body of evidence for each prognostic eRect estimation according to the recommendations of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group (GRADE 2013). We will use the adapted GRADE approach for questions on prognostic factors (Foroutan 2020; Huguet 2013; Iorio 2015; Westby 2018). GRADE initially considers evidence from phase 2 studies as high certainty. However, this initial certainty of evidence can be modified, based on the following criteria:

- Criteria for downgrading confidence in the prognostic eRect estimate: RoB, inconsistency, imprecision, indirectness and publication bias
- Criteria for upgrading confidence in the prognostic eRect estimate: large eRect

We will consider that the best evidence regarding a prognostic factor normally comes from observational studies (cohort studies, registries, or database linkage studies). Thus, we will provide an initial high-certainty rating to the body of the evidence based on these studies (Foroutan 2020; Iorio 2015). On the other hand, the certainty of the evidence for secondary analyses of RCTs will be probably lower due to the presence of restrictions of patients relevant for our review questions (Foroutan 2020). We will assess these restrictions as part of the assessment of indirectness with GRADE.

We will not consider the phase of investigation of studies in our assessment of the strength of the evidence available, as only phase 2 studies will be eligible.

We will use GRADEproGDT soNware (GRADEpro-GDT 2015) to create 'Summary of findings' tables with the main results of the review, including the certainty of the body of evidence related to each outcome. All the review outcomes are critical for decision making, so they will be included in the table. The 'Summary of findings' table will contain all decisions to down- or upgrade the certainty of the evidence with footnotes, and provide explanations to help the reader's understanding of the review where necessary. Two review authors (JLA, ES) will assess the certainty of the evidence found for each outcomes. Another review author (AM) will check the assessments. We have included a template 'Summary of findings' table in Table 2. We will create one table for each of the main comparisons of the review (if there are more than one).

ACKNOWLEDGEMENTS

We would like to thank the Cochrane Prognostic Methods group, particularly Karl Moons, Lotty HooN and Thomas Debray, for the support and guidance in the methodology and for the deliverance of a course on carrying out prognostic systematic reviews at the Cochrane Associate Centre of Madrid (June 2018).

We would like to thank the Cochrane Vascular Information Specialist, Candida Fenton.

The Cochrane Vascular editorial base wish to thank Elizabeth Stovold, Information Specialist, Cochrane Airways, for peer reviewing the search strategy.

The review authors and the Cochrane Vascular editorial base wish to thank the following peer reviewers for their comments:

Professor Alfonso Iorio, McMaster University, Hamilton, Canada
H Dirk Sostman MD FACR, Weill Cornell Medical College and Methodist Hospital, New York, USA
Dr Valdis Gībietis, Riga Stradins University, Riga, Latvia

We also thank the peer reviewer who chose to remain anonymous.

REFERENCES

Additional references

Agarwal 2015

Agarwal S, Clark D 3rd, Sud K, Jaber WA, Cho L, Menon V. Gender disparities in outcomes and resource utilization for acute pulmonary embolism hospitalizations in the United States. *American Journal of Cardiology* 2015;**116**(8):1270-6. [PMID: 26341183]

Alotaibi 2016

Alotaibi GS, Wu C, Senthilselvan A, McMurtry MS. Secular trends in incidence and mortality of acute venous thromboembolism: the AB-VTE population-based study. *American Journal of Medicine* 2016;**129**(8):879.e19-25. [PMID: 26925811]

Altman 2001

Altman DG. Systematic reviews of evaluations of prognostic variables. *BMJ (Clinical research ed.)* 2001;**323**(7306):224-8. [PMID: 11473921]

Aujesky 2005

Aujesky D, Obrosky DS, Stone RA, Auble TE, Perrier A, Cornuz J, et al. Derivation and validation of a prognostic model for pulmonary embolism. *American Journal of Respiratory and Critical Care Medicine* 2005;**172**(8):1041-6. [PMID: 16020800]

Australian Bureau of Statistics 2015

Australian Bureau of Statistics. Causes of death, Australia, 2015. Canberra: ABS, 2016. www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3303.02015?OpenDocument (accessed 2 February 2019).

Barrios 2018

Barrios D, Morillo R, Yusen RD, Jimenez D. Pulmonary embolism severity assessment and prognostication. *Thrombosis Research* 2018;**163**:246-51. [PMID: 28911787]

Bartholomew 2017

Bartholomew JR. Update on the management of venous thromboembolism. *Cleveland Clinic Journal of Medicine* 2017;**84**(12 Suppl 3):39-46. [PMID: 29257737]

Belohlavek 2013

Belohlavek J, Dytrych V, Linhart A. Pulmonary embolism, part II: management. *Experimental and Clinical Cardiology* 2013;**18**(2):139-47. [PMID: 23940439]

Blanco-Molina 2014

Blanco-Molina A, Enea I, Gadelha T, Tufano A, Bura-Riviere A et al. Sex diRerences in patients receiving anticoagulant therapy for venous thromboembolism. *Medicine* 2014;**93**(17):309-17. [PMID: 25398066]

Boissel 1999

Boissel JP, Cucherat M, Li W, Chatellier G, GueyRier F, Buyse M, et al. The problem of therapeutic eRicacy indices. 3. comparison of the indices and their use [Aperçu sur la problématique des indices d'eRicacité thérapeutique, 3: comparaison des indices et utilisation. Groupe d'Etude des Indices D'eRicacite]. *Therapie* 1999;**54**(4):405-11. [PMID: 10667106]

British Lung Foundation 2015

British Lung Foundation. Pulmonary embolism statistics. London: British Lung Foundation, 2015. www.statistics.blf.org.uk/pulmonary-embolism (accessed 20 February 2019).

Burles 2017

Burles K, Innes G, Senior K, Lang E, McRae A. Limitations of pulmonary embolism ICD-10 codes in emergency department administrative data: let the buyer beware. *BMC Medical Research Methodology* 2017;**17**(1):89. [PMID: 28595574]

CDC 2015

Centers for Disease Control and Prevention. Venous thromboembolism (blood clots). Atlanta: CDC, 2015. www.cdc.gov/ncbddd/dvt/data.html (accessed 2 February 2019).

Cho 2014

Cho JH, Kutti Sridharan G, Kim SH, Kaw R, Abburi T, Irfan A, et al. Right ventricular dysfunction as an echocardiographic prognostic factor in hemodynamically stable patients with acute pulmonary embolism: a meta-analysis. *BMC Cardiovascular Disorders* 2014;**14**:64. [PMID: 24884693]

CIHR 2012

Canadian Institute of Health Research, Institute of Gender and Health. What a diRerence sex and gender make: a gender, sex and health research casebook - CIHR. Vancouver, British Columbia: Canadian Institute of Health Research, 2012.

Clayton 2016

Clayton JA, Tannenbaum C. Reporting sex, gender, or both in clinical research? *JAMA* 2016;**316**(18):1863-4. [PMID: 27802482]

Cochrane Prognosis Methods Group 2019

Cochrane Prognosis Methods Group. Review tools: templates. https://methods.cochrane.org/prognosis/our-publications (accessed 13 July 2018).

Cohen 2007

Cohen AT, Agnelli G, Anderson FA, Arcelus JI, Bergqvist D, Brecht JG, et al. Venous thromboembolism (VTE) in Europe. The number of VTE events and associated morbidity and mortality. *Thrombosis and Haemostasis* 2007;**98**(4):756-64. [PMID: 17938798]

Cornell 2014

Cornell JE, Mulrow CD, Localio R, Stack CB, Meibohm AR, Guallar E, et al. Random-eRects meta-analysis of inconsistent eRects: a time for change. *Annals of Internal Medicine* 2014;**160**(4):267-70. [PMID: 24727843]

Cushman 2004

Cushman M, Kuller LH, Prentice R, Rodabough RJ, Psaty BM, StaRord RS, et al. Estrogen plus progestin and risk of venous thrombosis. *JAMA* 2004;**292**(13):1573-80. [PMID: 15467059]

Debray 2017

Debray TP, Damen JA, Snell KI, Ensor J, HooN L, Reitsma JB, et al. A guide to systematic review and meta-analysis of prediction model performance. *BMJ (Clinical research ed.)* 2017;**356**:i6460. [PMID: 28057641]

Deeks 2000

Deeks J. Issues in the selection for meta-analyses of binary data. In: Evidence for action: challenges for The Cochrane Collaboration in the 21st century. Abstracts of the 8th Cochrane Colloquium; 2000 25-29 Oct; Cape Town, South Africa. The Cochrane Collaboration, 2000.

Deeks 2011

Deeks J, Higgins J, Altman D. Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org (accessed 02 February 2019).

den Exter 2013

den Exter PL, van Es J, Klok FA, KroN LJ, Kruip MJ, Kamphuisen PW, et al. Risk profile and clinical outcome of symptomatic subsegmental acute pulmonary embolism. *Blood* 2013;**122**(7):1144-9. [PMID: 23736701]

Doherty 2017

Doherty S. Pulmonary embolism: an update. *Australian Family Physician* 2017;**46**(11):816-20. [PMID: 29101916]

Doull 2010

Doull M, Runnels VE, Tudiver S, Boscoe M. Appraising the evidence: applying sex- and gender-based analysis (SGBA) to Cochrane systematic reviews on cardiovascular diseases. *Journal of Women's Health* 2010;**19**(5):997-1003. [PMID: 20384450]

Evans 2003

Evans T, Brown H. Road traRic crashes: operationalizing equity in the context of health sector reform. *Injury Control and Safety Promotion* 2003;**10**(1-2):11-2. [PMID: 12772480]

Finkelstein 1994

Finkelstein DM, Schoenfeld DA. Analysing survival in the presence of an auxiliary variable. *Statistics in Medicine* 1994;**13**(17):1747-54. [PMID: 7997708]

Foroutan 2020

Foroutan F, Guyatt G, Zuk V, Vandvik PO, Alba AC, Mustafa R, et al. GRADE Guidelines 28: Use of GRADE for the assessment of evidence about prognostic factors: rating certainty in identification of groups of patients with diRerent absolute risks. *Journal of Clinical Epidemiology* 2020;**121**:62-70.

Gibietis 2019

Gibietis V, Kigitovica D, Vitola B, Strautmane S, Skride A. Glomerular filtration rate as a prognostic factor for long-term mortality aNer acute pulmonary embolism. Medical Principles and Practice: International Journal of the Kuwait University, Health Science Centre 2019. [PMID: 30716739]

GRADE 2013

Schünemann H, Brozek J, Guyatt G, Oxman A, editor(s). GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. The GRADE Working Group, 2013. Available from guidelinedevelopment.org/handbook/.

GRADEpro-GDT 2015 [Computer program]

McMaster University (developed by Evidence Prime, Inc) GRADEpro GDT. Version (accessed 17 May 2019). Hamilton, ON: McMaster University (developed by Evidence Prime, Inc), 2015. Available at gradepro.org.

Guyatt 2011

Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *Journal of Clinical Epidemiology* 2011;**64**(4):380-2. [PMID: 21185693]

Harbord 2008

Harbord RM, Higgins JP. Meta-regression in Stata. *Meta* 2008;**8**(4):493–519.

Hayden 2008

Hayden JA, Cote P, Steenstra IA, Bombardier C. Identifying phases of investigation helps planning, appraising, and applying the results of explanatory prognosis studies. *Journal of Clinical Epidemiology* 2008;**61**(6):552-60. [PMID: 18471659]

Hayden 2013

Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. *Annals of Internal Medicine* 2013;**158**(4):280-6. [PMID: 23420236]

Hayden 2014

Hayden JA, Tougas ME, Riley R, Iles R, Pincus T. Individual recovery expectations and prognosis of outcomes in non-specific low back pain: prognostic factor exemplar review. *Cochrane Database of Systematic Reviews* 2014, Issue 9. Art. No: CD011284. [DOI: 10.1002/14651858.CD011284]

Heidari 2016

Heidari S, Babor TF, Castro PD, Tort S, Curno M. Sex and gender equity in research: rationale for the SAGER guidelines and recommended use. *Research Integrity and Peer Review* 2016;**1**:2.

Hemingway 2013

Hemingway H, CroN P, Perel P, Hayden JA, Abrams K, Timmis A, et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ (Clinical research ed.)* 2013;**346**:e5595. [PMID: 23386360]

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ (Clinical research ed.)* 2003;**327**(7414):557-60. [PMID: 12958120]

Higgins 2011

Higgins JP, Deeks JJ. Chapter 7: Selecting studies and collecting data. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Huguet 2013

Huguet A, Hayden JA, Stinson J, McGrath PJ, Chambers CT, Tougas ME, et al. Judging the quality of evidence in reviews of prognostic factor research: adapting the GRADE framework. *Systematic Reviews* 2013;**2**:71. [PMID: 24007720]

IntHout 2014

IntHout J, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random eRects meta-analysis is straightforward and considerably out performs the standard DerSimonian-Laird method. *BMC Medical Research Methodology* 2014;**14**:25. [PMID: 24548571]

Iorio 2015

Iorio A, Spencer FA, Falavigna M, Alba C, Lang E, Burnand B, et al. Use of GRADE for assessment of evidence about prognosis: rating confidence in estimates of event rates in broad categories of patients. *BMJ (Clinical research ed.)* 2015;**350**:h870. [PMID: 25775931]

James 2006

James AH, Jamison MG, Brancazio LR, Myers ER. Venous thromboembolism during pregnancy and the postpartum period: incidence, risk factors, and mortality. *American Journal of Obstetrics and Gynecology* 2006;**194**(5):1311-5. [PMID: 16647915]

Jimenez 2010

Jimenez D, Aujesky D, Moores L, Gomez V, Lobo JL, Uresandi F, et al. Simplification of the pulmonary embolism severity index for prognostication in patients with acute symptomatic pulmonary embolism. *Archives of Internal Medicine* 2010;**170**(15):1383-9. [PMID: 20696966]

Jo 2013

Jo JY, Lee MY, Lee JW, Rho BH, Choi WI. Leukocytes and systemic inflammatory response syndrome as prognostic factors in pulmonary embolism patients. *BMC Pulmonary Medicine* 2013;**13**:74. [PMID: 24325351]

Kearon 2012

Kearon C, Akl EA, Comerota AJ, Prandoni P, Bounameaux H, Goldhaber SZ, et al. Antithrombotic therapy for VTE disease: Antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* 2012;**141**(2 Suppl):e419S-96S. [PMID: 22315268]

Kearon 2016a

Kearon C, Ageno W, Cannegieter SC, Cosmi B, Geersing GJ, Kyrle PA. Categorization of patients as having provoked or unprovoked venous thromboembolism: guidance from the SSC of ISTH. *Journal of Thrombosis and Haemostasis* 2016;**14**(7):1480-3. [PMID: 27428935]

Kearon 2016b

Kearon C, Akl EA, Ornelas J, Blaivas A, Jimenez D, Bounameaux H, et al. Antithrombotic therapy for VTE disease: CHEST guideline and expert panel report. *Chest* 2016;**149**(2):315-52. [PMID: 26867832]

Keller 2019

Keller K, Rappold L, Gerhold-Ay A, Hobohm L, Hasenfuss G, Konstantinides SV, et al. Sex-specific diRerences in pulmonary embolism. *Thrombosis Research* 2019;**178**:173-81. [PMID: 31039477]

Klok 2010

Klok FA, Zondag W, van Kralingen KW, van Dijk AP, Tamsma JT, Heyning FH, et al. Patient outcomes aNer acute pulmonary embolism. A pooled survival analysis of diRerent adverse events. *American Journal of Respiratory and Critical Care Medicine* 2010;**181**(5):501-6. [PMID: 19965808]

Konstantinides 2014

Konstantinides SV, Torbicki A, Agnelli G, Danchin N, Fitzmaurice D. 2014 ESC guidelines on the diagnosis and management of acute pulmonary embolism. *European Heart Journal* 2014;**35**(43):3033-69. [PMID: 25452462]

Konstantinides 2016

Konstantinides SV. Trends in incidence versus case fatality rates of pulmonary embolism: good news or bad news? Thrombosis and Haemostasis 2016;**115**(2):233-5. [PMID: 26632213]

Konstantinides 2019

Konstantinides SV, Meyer G, Becattini C, Bueno H, Geersing GJ, Harjola VP, et al. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS): The Task Force for the diagnosis and management of acute pulmonary embolism of the European Society of Cardiology (ESC). *European Heart Journal* 2020;**41**:543-603. [DOI: 10.1093/eurheartj/ehz405]

Kujovich 2004

Kujovich JL. Hormones and pregnancy: thromboembolic risks for women. *British Journal of Haematology* 2004;**126**(4):443-54. [PMID: 15287937]

Kyzas 2007a

Kyzas PA, Denaxa-Kyza D, Ioannidis JP. Almost all articles on cancer prognostic markers report statistically significant results. *European Journal of Cancer* 2007;**43**(17):2559-79. [PMID: 17981458]

Kyzas 2007b

Kyzas PA, Denaxa-Kyza D, Ioannidis JP. Quality of reporting of cancer prognostic marker studies: association with reported

prognostic eRect. *Journal of the National Cancer Institute* 2007;**99**(3):236-43. [PMID: 17284718]

Lee 2016

Lee JS, Moon T, Kim TH, Kim SY, Choi JY, Lee KB, et al. Deep vein thrombosis in patients with pulmonary embolism: prevalence, clinical significance and outcome. *Vascular Specialist International* 2016;**32**(4):166-74. [PMID: 28042556]

Liberati 2009

Liberati A, Altman DG, TetzlaR J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ (Clinical Research ed.)* 2009;**339**:b2700. [PMID: 19622552]

Lopez-Alcalde 2018

López-Alcalde J, Stallings E, Cabir S, Zamora J. Definition of eligibility for non-randomised studies in 113 Cochrane reviews of interventions to prevent healthcare-associated infections. Abstracts of the Cochrane Colloquium, Edinburgh, Scotland. *Cochrane Database of Systematic Reviews* 2018;**9**(Suppl 1):1052.

Lopez-Alcalde 2019

López-Alcalde J, Stallings E, Cabir S, Fernández A, Daheron M, Bonfill X, Zamora J. Consideration of sex and gender in Cochrane reviews of interventions for preventing healthcare-associated infections: a methodology study. *BMC Health Services Research* 2019;**19**:169. [PMID: 30876452]

Marik 2008

Marik PE, Plante LA. Venous thromboembolic disease and pregnancy. *New England Journal of Medicine* 2008;**359**(19):2025-33. [PMID: 18987370]

Merli 2017

Merli GJ. Pulmonary embolism in 2017: how we got here and where are we going? *Techniques in Vascular and Interventional Radiology* 2017;**20**(3):128-34. [PMID: 29029706]

Moons 2014

Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Medicine* 2014;**11**(10):e1001744. [PMID: 25314315]

Morris 2010

Morris JM, Algert CS, Roberts CL. Incidence and risk factors for pulmonary embolism in the postpartum period. *Journal of Thrombosis and Haemostasis* 2010;**8**(5):998-1003. [PMID: 20128859]

Muriel 2014

Muriel A, Jimenez D, Aujesky D, Bertoletti L, Decousus H, Laporte S, et al. Survival eRects of inferior vena cava filter in patients with acute symptomatic venous thromboembolism and a significant bleeding risk. *Journal of the American College of Cardiology* 2014;**63**(16):1675-83. [PMID: 24576432]

Murphy 2017

Murphy SL, Xu J, Kochanek KD, Curtin SC, Arias E. Deaths: final data for 2015. *National Vital Statistics Reports* 2017;**66**(6):1-75. [PMID: 29235985]

Murphy 2018

Murphy E, Lababidi A, Reddy R, Mendha T, Lebowitz D. The role of thrombolytic therapy for patients with a submassive pulmonary embolism. *Cureus* 2018;**10**(6):e2814. [PMID: 30397555]

Navanandan 2019

Navanandan N, Stein J, Mistry RD. Pulmonary embolism in children. *Pediatric Emergency Care* 2019;**35**(2):143-51. [PMID: 30702542]

Ng 2011

Ng AC, Chung T, Yong AS, Wong HS, Chow V, Celermajer DS, et al. Long-term cardiovascular and non-cardiovascular mortality of 1023 patients with confirmed acute pulmonary embolism. *Circulation* 2011;**4**(1):122-8. [PMID: 21098781]

O'Neill 2014

O'Neill J, Tabish H, Welch V, Petticrew M, Pottie K, Clarke M, et al. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *Journal of Clinical Epidemiology* 2014;**67**(1):56-64. [PMID: 24189091]

OFice for National Statistics 2013

UK ORice for National Statistics. Deaths registered in England and Wales: 2012. London: ORice for National Statistics, 2013. www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsregistrationsummarytables/2013-07-10 (accessed 20 February 2019).

OSSD

The Organization for the Study of Sex DiRerences (OSSD). Organization for the Study of Sex DiRerences. www.ossdweb.org (accessed 13 December 2018).

Panigada 2016

Panigada G, Masotti L, Rosi C, Teghini L, Cimolato B, Bertieri MC, et al. Thromboembolic burden, prognostic assessment and outcomes of females compared to males in acute pulmonary embolism. *Acta Clinica Belgica* 2016;**71**(3):142-8. [PMID: 27075811]

Park 2018

Park SE, Thomas J. Evidence synthesis soNware. *BMJ Evidence-Based Medicine* 2018;**23**(4):140-1. [PMID: 29880698]

Peragallo Urrutia 2013

Peragallo Urrutia R, Coeytaux RR, McBroom AJ, Gierisch JM, Havrilesky LJ, Moorman PG, et al. Risk of acute thromboembolic events with oral contraceptive use: a systematic review and meta-analysis. *Obstetrics and Gynecology* 2013;**122**(2 Pt 1):380-9. [PMID: 23969809]

Peters 2006

Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of two methods to detect publication bias in meta-analysis. *JAMA* 2006;**295**(6):676-80. [PMID: 16467236]

Polus 2017

Polus S, Pieper D, Burns J, Fretheim A, Ramsay C, Higgins JPT, et al. Heterogeneity in application, design, and analysis characteristics was found for controlled before-aNer and interrupted time series studies included in Cochrane reviews. *Journal of Clinical Epidemiology* 2017;**91**:56-69. [PMID: 28750849]

Reeves 2011

Reeves BC, Deeks JJ, Higgins JP, Wells GA. Chapter 13: Including non-randomised studies. In: Higgins JP, Green S editor(s). Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Review Manager 2014 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Riley 2011

Riley RD, Higgins JP, Deeks JJ. Interpretation of random eRects meta-analyses. *BMJ* 2011;**342**:d549. [PMID: 21310794]

Riley 2013

Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA, et al. Prognosis research strategy (PROGRESS) 2: prognostic factor research. *PLoS Medicine* 2013;**10**(2):e1001380. [PMID: 23393429]

Riley 2019

Riley RD, Moons KGM, Snell KIE, Ensor J, HooN L, Altman DG, et al. A guide to systematic review and meta-analysis of prognostic factor studies. *BMJ (Clinical research ed.)* 2019;**364**:k4597. [PMID: 30700442]

Roy 2006

Roy PM, Meyer G, Vielle B, Le Gall C, Verschuren F, Carpentier F, et al. Appropriateness of diagnostic management and outcomes of suspected pulmonary embolism. *Annals of Internal Medicine* 2006;**144**(3):157-64. [PMID: 16461959]

Rucker 2008

Rucker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. *BMC Medical Research Methodology* 2008;**8**:79. [PMID: 19036172]

Runnels 2014

Runnels V, Tudiver S, Doull M, Boscoe M. The challenges of including sex/gender analysis in systematic reviews: a qualitative survey. *Systematic Reviews* 2014;**3**:33. [PMID: 24720875]

Sackett 1996

Sackett DL, Deeks JJ, Altman DG. Down with odds ratios! *Evidence Based Medicine* 1996;**1**:164-6. [DOI: 10.1136/ebm.1996.1.164]

Saldanha 2014

Saldanha IJ, Dickersin K, Wang X, Li T. Outcomes in Cochrane systematic reviews addressing four common eye conditions: an evaluation of completeness and comparability. *PloS One* 2014;**9**(10):e109400. [PMID: 25329377]

Scherz 2010

Scherz N, Labarere J, Mean M, Ibrahim SA, Fine MJ, Aujesky D. Prognostic importance of hyponatremia in patients with acute pulmonary embolism. *American Journal of Respiratory and Critical Care Medicine* 2010;**182**(9):1178-83. [PMID: 20595225]

Schoenfeld 2005

Schoenfeld D. Survival methods, including those using competing risk analysis,are not appropriate for intensive care unit outcome studies. *Critical Care* 2006;**10**:103. [PMID: 16420653]

Schünemann 2011

Schünemann HJ, Oxman AD, Vist GE, Higgins JP, Deeks JJ, Glasziou P, Guyatt GH. Chapter 12: Interpreting results and drawing conclusions. In: Higgins JP, Green S editor(s), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org 2011.

Sen 2014

Sen HS, Abakay O, Tanrikulu AC, Sezgi C, Taylan M, Abakay A, et al. Is a complete blood cell count useful in determining the prognosis of pulmonary embolism? *Wiener Klinische Wochenschri* 2014;**126**(11-12):347-54. [PMID: 24664312]

Shamseer 2015

Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ (Clinical research ed.)* 2015;**350**:g7647. [PMID: 25555855]

Sinclair 1994

Sinclair JC, Bracken MB. Clinically useful measures of eRect in binary analyses of randomized trials. *Journal of Clinical Epidemiology* 1994;**47**(8):881-9. [PMID: 7730891]

Skoetz 2017

Skoetz N, Will A, Monsef I, Brillant C, Engert A, von Tresckow B. Comparison of first-line chemotherapy including escalated BEACOPP versus chemotherapy including ABVD for people with early unfavourable or advanced stage Hodgkin lymphoma. *Cochrane Database of Systematic Reviews* 2017, Issue 5. Art. No: CD007941. [DOI: 10.1002/14651858.CD007941.pub3]

Stegeman 2013

Stegeman BH, de Bastos M, Rosendaal FR, van Hylckama Vlieg A, Helmerhorst FM, Stijnen T, et al. DiRerent combined oral contraceptives and the risk of venous thrombosis: systematic

review and network meta-analysis. *BMJ (Clinical research ed.)* 2013;**347**:f5298. [PMID: 24030561]

Sterne 2011

Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ (Clinical research ed.)* 2011;**343**:d4002. [PMID: 21784880]

Tapson 2008

Tapson VF. Acute pulmonary embolism. *The New England Journal of Medicine* 2008;**358**(10):1037-52. [PMID: 18322285]

Trusheim 2007

Trusheim MR, Berndt ER, Douglas FL. Stratified medicine: strategic and economic implications of combining drugs and clinical biomarkers. *Nature reviews: Drug discovery* 2007;**6**(4):287-93. [PMID: 17380152]

Tugwell 2017

Tugwell P, Knottnerus JA. Current study design labels are confusing! Call for consensus on better terminology that clearly reflects specific features. Journal of Clinical Epidemiology 2017;**91**:1-3. [PMID: 29169708]

Veroniki 2019

Veroniki AA, Jackson D, Bender R, Kuss O, Langan D, Higgins JPT, et al. Methods to calculate uncertainty in the estimated overall eRect size from a random-eRects meta-analysis. *Research Synthesis Methods* 2019;**10**(1):23-43. [PMID: 30129707]

Welch 2017

Welch V, Doull M, Yoganathan M, Jull J, Boscoe M, Coen SE, et al. Reporting of sex and gender in randomized controlled trials in Canada: a cross-sectional methods study. *Research Integrity and Peer Review* 2017;**2**:15. [PMID: 29451565]

Welch 2020

Welch VA, Petkovic J, Jull J, Hartling L, Klassen T, Kristjansson E, et al. Chapter 16: Equity and specific populations. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (updated September 2020). Cochrane, 2020. Available from training.cochrane.org/handbook.

Wells 2000

Wells PS, Anderson DR, Rodger M, Ginsberg JS, Kearon C, Gent M, et al. Derivation of a simple clinical model to categorize patients probability of pulmonary embolism: increasing the models utility with the SimpliRED D-dimer. *Thrombosis and Haemostasis* 2000;**83**(3):416-20. [PMID: 10744147]

Wendelboe 2016

Wendelboe AM, Raskob GE. Global burden of thrombosis: epidemiologic aspects. *Circulation Research* 2016;**118**(9):1340-7. [PMID: 27126645]

Westby 2018

Westby MJ, Dumville JC, Stubbs N, Norman G, Wong JK, Cullum N, et al. Protease activity as a prognostic factor for wound healing in venous leg ulcers. *Cochrane Database of Systematic Reviews* 2018, Issue 9. Art. No: CD012841. [DOI: 10.1002/14651858.CD012841.pub2]

Wiener 2013

Wiener RS, Schwartz LM, Woloshin S. When a test is too good: how CT pulmonary angiograms find pulmonary emboli that do not need to be found. *BMJ (Clinical research ed.)* 2013;**347**:f3368. [PMID: 23820021]

Wolkewitz 2014

Wolkewitz M, Sommer H. Evaluating mortality in an intensive care unit requires extended survival models. *Critical Care* 2014;**18**:415. [PMID: 25029237]

Zaidi 2017

Zaidi AU, Hutchins KK, Rajpurkar M. Pulmonary embolism in children. *Frontiers in Pediatrics* 2017;**5**:170. [PMID: 28848725]

Zhang 1998

Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. *JAMA* 1998;**280**(19):1690-1. [PMID: 9832001]

ADDITIONAL TABLES

Table 2. DraK 'Summary of findings' table

Does being a female compared with being a male help predict mortality in adults with acute symptomatic PE?

Patient or population: adults with acute symptomatic PE (confirmed by objective testing)

Settings: any

Index prognostic factor: being a female

Comparator: age, history of cancer, current cancer, history of chronic cardiopulmonary disease, current chronic cardiopulmonary disease, heart rate, systolic blood pressure, and O_2 saturation

Prognostic factor	Outcome	Study results and measurements	Absolute effect estimates* (95% CI)	Certainty in the effect estimates (Qual-	Plain text

Table 2. DraK 'Summary of findings' table *(Continued)*

			Assumed risk in men	Corresponding risk in women	ity of the evidence)	summary
Sex (female versus male)	**All-cause hospital mortality** (follow up: the longest follow-up provided by the study authors)	**OR [value]** (95% CI [value] to [value]) Based on data from XXXX patients in XX studies	**X per 1000**	**X per 1000** Difference: XX more/less per 1000 (CI 95% X more/less - X more/less)	**Very / Low / Moderate / High** Due to	
Sex (female versus male)	**All-cause hospital mortality** (at 30 days)	**OR [value]** (95% CI [value] to [value]) Based on data from XXXX patients in XX studies	**X per 1000**	**X per 1000** Difference: XX more/less per 1000 (CI 95% X more/less - X more/less)	**Very / Low / Moderate / High** Due to	
Sex (female versus male)	**All-cause mortality** (at 90 days)	**OR [value]** (95% CI [value] to [value]) Based on data from XXXX patients in XX studies	**X per 1000**	**X per 1000** Difference: XX more/less per 1000 (CI 95% X more/less - X more/less)	**Very / Low / Moderate / High** Due to	
Sex (female versus male)	**Early hospital mortality** (during the first 48 hours)	**OR [value]** (95% CI [value] to [value]) Based on data from XXXX patients in XX studies	**X per 1000**	**X per 1000** Difference: XX more/less per 1000 (CI 95% X more/less - X more/less)	**Very / Low / Moderate / High** Due to	
Sex (female versus male)	**All-cause mortality** (at one year)	**OR [value]** (95% CI [value] to [value]) Based on data from XXXX patients in XX studies	**X per 1000**	**X per 1000** Difference: XX more/less per 1000 (CI 95% X more/less - X more/less)	**Very / Low / Moderate / High** Due to	
Sex (female versus male)	**PE-related hospital mortality** (follow up: the longest follow-up provided by the study authors)	**OR [value]** (95% CI [value] to [value]) Based on data from XXXX patients in XX studies	**X per 1000**	**X per 1000** Difference: XX more/less per 1000 (CI 95% X more/less - X more/less)	**Very / Low / Moderate / High** Due to	
Sex (female versus male)	**PE-related hospital mortality** (at 30 days)	**OR [value]** (95% CI [value] to [value]) Based on data from XXXX patients in XX studies	**X per 1000**	**X per 1000** Difference: XX more/less per 1000 (CI 95% X more/less - X more/less)	**Very / Low / Moderate / High** Due to	
Sex (female ver-	**Early PE-related hospital mortality** (during the first 48 hours)	**OR [value]** (95% CI [value] to [value]) Based on data from XXXX patients in XX studies	**X per 1000**	**X per 1000** Difference: XX more/less per 1000 (CI 95% X more/less - X more/less)	**Very / Low / Moderate / High** Due to	

Table 2. DraK 'Summary of findings' table *(Continued)*

sus male)				

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative eFect** of the intervention (and its 95% CI).
Abbreviations:
CI: confidence interval; **OR:** odds ratio; **PE**: pulmonary embolism
GRADE Working Group grades of evidence:
High certainty: Further research is very unlikely to change our confidence in the estimate of eRect.
Moderate certainty: Further research is likely to have an important impact on our confidence in the estimate of eRect and may change the estimate.
Low certainty: Further research is very likely to have an important impact on our confidence in the estimate of eRect and is likely to change the estimate.
Very low certainty: We are very uncertain about the estimate.

APPENDICES

Appendix 1. Study design features

There is no standardised nomenclature for non-randomised studies (NRS), and this can cause problems when defining the types of studies to include in a systematic review and when deciding on the eligibility of the primary studies (Lopez-Alcalde 2018; Polus 2017; Reeves 2011; Tugwell 2017). We consider here explicit study design features (not only the study design labels) to define the design eligibility. Moreover, we will take these features into account when assessing studies for selection. The Cochrane Non-randomised Studies Methods Group (NRSMG) (Reeves 2011), proposes to define items 1 to 5. We will also consider additional criteria relevant for prognostic studies (items 6 to 9).

1. **Unit of allocation (individual or group level)**: not applicable as there is no allocation of an intervention in our review question
2. **Comparison**: between two groups of participants (males and females)
3. **Method of allocation of study participants to groups (randomised or not randomised)**: not applicable as there is no allocation of an intervention in our review question
4. **Prospective or retrospective character of each study part:** any. We will also include studies that did not describe if they were prospective or retrospective (as these aspects are rarely reported):
   a. **Identification of participants:** prospective, retrospective or unclear
   b. **Assessment of baseline:** prospective, retrospective or unclear
   c. **Evaluation of outcomes:** prospective, retrospective or unclear
   d. **Generation of hypothesis:** prospective, retrospective or unclear
5. **Variables to assess the comparability between study groups:**
   a. **Potential additional prognostic factors**
      i. For a study to be eligible, we will require that the study has tried to determine the adjusted prognostic value of sex - that is, its prognostic value independently of other existing prognostic factors such as age, or history of cancer. Thus, for a study to be eligible it should have taken into consideration additional prognostic factors (apart from sex) by using a particular design approach to control for confounding, or by using a specific method to measure and adjust for confounding in the analysis. We will not require the consideration of specific covariates, the use of a particular design approach to control for confounding, or the use of a particular method to measure and adjust for confounding in the analysis. Our data extraction and risk of bias assessments will consider the covariates that were measured, controlled (by the study design) and adjusted (by the analysis). See below 'Comparator' and Appendix 2 for additional prognostic factors
   b. **Baseline assessment of outcome:** not applicable, as we will not require this criterion for inclusion
6. **Temporal sequence:** we will only include longitudinal studies, that is, studies that collect data over a period of time. Thus, we will exclude cross-sectional studies (studies that collect data only once and in one short period of time). We considered admitting cross-sectional studies for two reasons. First, our review question does not aim to test a causal association between sex and the outcomes. Second, we know the temporal sequence as the potential prognostic factor (sex) always comes before any outcome. However, we excluded cross-sectional studies because they do not allow the assessment of the proper temporal sequence for the study covariates.
7. **Phase of prognostic factor investigation:** phase 2-confirmatory. That is, explanatory research aimed to confirm an independent association between a potential prognostic factor (sex) and the outcome of interest. A phase-2 study seeks to measure the independent eRect of a prognostic factor while controlling for other factors (Hayden 2008; Hayden 2014), and is recognisable by its objective statement that outlines a specific prognostic factor of interest (Hayden 2008).
8. **Follow-up period to measure the outcome:** as defined for each outcome (see below).

9. **Data sources used in the study:** studies will be eligible independently of their data origin (data collected exclusively for research purposes or based on administrative databases). For example, a phase-2 prognostic study based on a database obtained for a randomised controlled trial would be eligible. On the other hand, we acknowledge that there is an ongoing controversy about the accuracy of administrative databases for the identification of PE cases (Burles 2017); these studies will be eligible as well, but we will assess the impact of this decision by sensitivity analysis.

Appendix 2. Key covariates for the adjustment of mortality estimates in patients with pulmonary embolism

We identified the key covariates for adjustment both from non systematic review of the literature, and in discussion with clinicians of the review team according to the following process.

Step	Method	Potential additional prognostic factors	Source
1. Preliminary searches to identify potential prognostic factors on mortality in patients with pulmonary embolism	1. PubMed search:"pulmonary embolism"[Title]) AND "prognostic factor"[Title] 2. Embase search: 'prognostic factor':ti AND 'pulmonary embolism':ti 3. Initial discussion with review team members	Red cell distribution width	Sen 2014
		Right ventricular dysfunction (RVD)	Cho 2014
		Glomerular filtration rate	Gibietis 2019
		Hyponatremia	Scherz 2010
		Leukocytes	Jo 2013
		SIRS	Jo 2013
2. Identify prognostic models for mortality in patients with pulmonary embolism	We considered the factors considered in the simplified PESI prognostic model (Jimenez 2010)	• Age • History of cancer • History of chronic cardiopulmonary disease • Heart rate • Systolic blood pressure • O_2 saturation	Jimenez 2010
3. Prioritisation of additional prognostic factors in GRADEPro GDT (GRADEpro-GDT 2015)	a. We circulated the preliminary list of prognostic factors to our systematic review team. b. The review authors commented on the factors already listed and/or added new ones to the list. c. The review team received a new revised list and were asked to prioritise the factors, ranking them from 1 to 9, with 1 being of least importance and 9 of the highest importance. d. We sent a new list of potential prognostic factors to group the factors according to their relative importance (1 to 3 points: not relevant; 4 to 6 points: important; 7 to 9 points: critical). e. We asked the review team to confirm the final list of key additional prognostic factors.		
4. Final decision	We agreed the final list of covariates		

Appendix 3. MEDLINE search strategy

1 Pulmonary Embolism/

2 Thromboembolism/

3 Thrombosis/

4 exp Venous Thromboembolism/

5 exp Venous Thrombosis/

6 ((vein* or ven*) adj thromb*).ti,ab.

7 (blood adj3 clot*).ti,ab.

8 deep vein thrombosis.ti,ab.

9 (lung adj3 clot*).ti,ab.

10 (DVT or VTE).ti,ab.

11 peripheral vascular thrombosis.ti,ab.

12 post-thrombotic syndrome.ti,ab.

13 pulmonary embolism.ti,ab.

14 (pulmonary adj3 clot*).ti,ab.

15 (thrombus* or thrombopro* or thrombotic* or thrombolic* or thromboemboli* or thrombos* or embol* or microembol*).ti,ab.

16 venous thromboembolism.ti,ab.

17 or/1-16

18 exp Sex Factors/

19 exp Sex Characteristics/

20 exp Sex Distribution/

21 exp Sex/

22 exp Sex Ratio/

23 exp Women's Health/

24 exp Men's Health/

25 boy*.ti,ab.

26 female*.ti,ab.

27 gender.ti,ab.

28 girl*.ti,ab.

29 male*.ti,ab.

30 maternal.ti,ab.

31 men.ti,ab.

32 postnatal.ti,ab.

33 pregnan*.ti,ab.

34 sex.ti,ab.

35 women.ti,ab.

36 or/18-35

37 17 and 36

38 exp Mortality/

39 exp Follow-Up Studies/

40 exp Incidence/

41 exp Survival Analysis/

42 prognos*.ti,ab.

43 predict*.ti,ab.

44 course*.ti,ab.

45 "disease history".ti,ab.

46 mortality.ti,ab.

47 outcome*.ti,ab.

48 or/38-47

49 37 and 48

50 exp animals/ not humans.sh.

51 49 not 50

HISTORY

Protocol first published: Issue 1, 2021

CONTRIBUTIONS OF AUTHORS

JLA: guarantor of the review, conceiving the review, designing the review, coordinating the review, designing search strategies, methodological input, providing a policy perspective, writing the protocol, editing the protocol, study selection, data extraction, risk of bias assessment, statistical analysis, writing of the review
ES: conceiving the review, designing the review, designing search strategies, methodological input, providing a policy perspective, writing the protocol, editing the protocol, study selection, data extraction, risk of bias assessment, statistical analysis, writing of the review
JZ: conceiving the review, designing the review, methodological input, providing a policy perspective, writing the protocol, risk of bias assessment, statistical analysis
AM: methodological input, providing a policy perspective, writing the protocol, risk of bias assessment, statistical analysis
SvD: designing the review, providing a policy perspective, writing the protocol
NAD: designing search strategies, providing a policy perspective, writing the protocol
BF: providing a policy perspective, writing the protocol, study selection, data extraction, risk of bias assessment, statistical analysis
CAQ: clinical input, providing a policy perspective, writing the protocol, study selection, data extraction
RP: clinical input, providing a policy perspective, writing the protocol, study selection
DJ: conceiving the review, designing the review, methodological input, clinical input, providing a policy perspective, writing the protocol, study selection, data extraction, statistical analysis, writing of the review

DECLARATIONS OF INTEREST

JLA: none known
ES: none known. This review will form part of the thesis of Elena Stallings, who is enrolled in doctoral studies with the Department of Health Sciences at the Universidad de Alcala, Madrid
JZ: none known
AM: none known
SvD: none known
NAD: none known
BF: none known
CAQL: none known
RP: none known
DJ: has declared that he received payment for consultancy (Bayer, Bristol-Meiers-Squibb, Daiichi-Sankyo, Pfizer, ROVI, Sanofi), lecture and education presentations (Bayer, Bristol-Meiers-Squibb, Daiichi-Sankyo, Merck-Sharp-Dome, Pfizer, ROVI, Sanofi) and from grants (Daiichi-Sankyo)

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- SEXCOMPLEX, Spain

  This study is part of the SEXCOMPLEX project ('Influence of sex and sex hormones on human chronic disorders of complex etiology'), a two-year project (2017–2019) coordinated by Hospital Ramón y Cajal (Madrid, Spain). The SEXCOMPLEX project was supported by Instituto de Salud Carlos III (Plan Estatal de I + D + i 2013–2016) and co-financed by the European Development Regional Fund 'A way to achieve Europe' (ERDF) grant number PIE16/00050. These funding sources had/will have no role in the design of this review, it's execution, analyses, interpretation of the data, or decision to submit results.
- Chief Scientist ORice, Scottish Government Health Directorates, The Scottish Government, UK

  The Cochrane Vascular editorial base is supported by the Chief Scientist ORice.

4. 2nd Article: Development of a prognostic factor search filter for Ovid Medline

After carrying out a preliminary search in Medline for studies to be included in the "Sex as a prognostic factor in patients with symptomatic acute pulmonary embolism" review and also after carrying out another PF SR I realised that it was necessary to create a search filter for prognostic factor studies. To my knowledge, a filter for PF studies does not yet exist. There are general prognosis filters available such as the Haynes clinical queries filter in PubMed, but these are not specific for prognostic factors and end up retrieving a lot of studies that are not relevant.

Prognostic factor systematic review searches that are carried out without the use of a prognostic filter can retrieve up to 110,000 studies that must be manually screened. In the SR of pulmonary embolism our search retrieved 113,000 references. In our PF SR on sepsis the search retrieved 30,000 before we decided to add a general prognosis filter. This many references to screen is not feasible when you have a small research team and limited resources to carry out the project, so, the development of the PF filter arose from the necessity to reduce the number of references to screen to a more manageable size for systematic reviewers.

Development and evaluation of a search filter to identify prognostic factor studies in Ovid MEDLINE.

Elena Stallings [1,2], Andrea Gaetano Gil [1], Noelia Alvarez-Diaz [3], Ivan Solà [2,4], Jesús López-Alcalde[1,2,5,6], Daniel Molano[7], Javier Zamora [1,2,8]

Affiliations:

[1]Clinical Biostatistics Unit, Instituto Ramón y Cajal de Investigación Sanitaria,28034 Madrid, Spain

[2]CIBER Epidemiología y Salud Pública (CIBERESP), 28029 Madrid, Spain

[3]Library, Hospital Universitario Ramón y Cajal (IRYCIS), Madrid, Spain.

[4] Sant Pau Biomedical Research Institute (IIB Sant Pau), 08041 Barcelona, Spain

[5]Faculty of Health Sciences, Universidad Francisco de Vitoria, 28223 Madrid, Spain;

[6]Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, CH-8091 Zurich, Switzerland.

[7]Department of Critical Care, Fundacion Universitaria de Ciencias de la Salud, Hospital de San José, Carrera 19 # 8-32, Bogota, Colombia, 11001

[8]WHO Collaborating Centre for Global Women's Health, Institute of Metabolism and Systems Research, University of Birmingham, B15 2TT Birmingham, UK

Abstract

Background: Systematic reviews (SRs) are valuable resources as they address specific clinical questions by searching and summarizing all existing relevant studies. However, finding all information to include in systematic reviews can be challenging. To facilitate a comprehensive identification of studies, methodological search filters have been developed to find articles related to specific clinical questions. To our knowledge, no filter exists for finding studies on the role of prognostic factor (PF). We aimed to develop and evaluate a search filter to identify PF studies in Ovid MEDLINE that has maximum sensitivity.

Methods: We followed current recommendations for the development of a search filter by first identifying a reference set of PF studies included in relevant systematic reviews on the topic, and by selecting search terms using a word frequency analysis complemented with an expert panel discussion. We evaluated filter performance using the relative recall methodology.

Results: We constructed a reference set of 73 studies included in six systematic reviews from a larger sample (91 reviews). After completing the word frequency analysis using the reference set studies, we compiled a list of 80 of the most frequent methodological terms. This list of terms was evaluated

by the Delphi panel for inclusion in the filter, resulting in a final set of 8 appropriate terms. The consecutive connection of these terms with the Boolean operator OR produced the filter. We then evaluated the filter using the relative recall method against the reference set, comparing the references included in the SRs with our new search using the filter. The overall sensitivity of the filter was calculated to be 95%, while the overall specificity was 41%. The precision of the filter varied considerably, ranging from 0.36 to 17%. The NNR (number needed to read) value depends on the total number of hits in the search and varied largely from 6 to 278.

Conclusions: We developed a search filter for OVID-Medline with acceptable performance that could be used in systematic reviews of PF studies. Using this filter could save as much as 40% of the title and abstract screening task. The specificity of the filter could be improved by defining additional terms to be included, although it is important to evaluate any modification to guarantee the filter is still highly sensitive.

Keywords: prognostic factor, search filter, systematic review

Introduction

It is essential to carry out a systematic and extensive search for any type of systematic review. However, searches can often retrieve an overwhelming number of studies (1, 2). To overcome this, methodological search filters have been developed to find articles related to specific clinical questions. A search filter is a pre-defined combination of search terms combined into a search strategy using the "AND" Boolean operator. Dozens of search filters exist for retrieving randomized controlled trials (RCTs) (3, 4). These filters have been successful in reducing the number of references needed to screen in systematic reviews, however this is difficult to reproduce for prognostic factor studies, as the literature pertaining to non-interventional studies is more variable. Unlike RCTs, non-interventional investigations have heterogenous, non-standardized study designs (5). These studies also suffer from poorer indexing of terms, thus making them more difficult to find in the database. Due to these limitations, the use of filters in diagnostic or prognostic studies is not widely recommended (6-8).

Prognosis research focuses on identifying variables that allow the estimation of the possibilities of improvement or worsening in a given health problem. This area of clinical research is becoming significantly more important, as throughout the world, people are living longer, but with more chronic health conditions and diseases. Prognosis research can be classified into four different themes or areas of research: fundamental prognostics, prognostic models, stratified medicine, and prognostic factors (9-12). A prognostic factor (PF) "is any measure that, among people with a given health condition (that is, a start point), is associated with a subsequent clinical outcome (an endpoint)" (11). Generic filters exist for finding prediction and prognosis studies such as the Haynes broad filter, Ingui filter

and the Yale prognosis and natural history filter (13). These published prognostic search filters have lower sensitivity and precision than other types of search filters such as those for medical intervention studies (14). While carrying out various PF systematic reviews we explored the possibility of using a PF filter (15, 16), however, to the best of our knowledge, no filter exists for these studies. The aim of this paper is to develop and evaluate a search filter for prognostic factor studies to be used in SRs. The main objective of the filter is to achieve maximum sensitivity so as not to lose any relevant studies when using the filter, while maintaining specificity to make the search more efficient.

Methods

We developed a search filter partially based on methods described by Rietjens et al., Sampson et al., and also on the criteria of the filter appraisal tool developed by Glanville et al (17-19). The completed filter appraisal checklist is available as supplementary material. We completed the study in three phases as outlined below:

1. Identification of a reference set (relative recall)
2. Search term selection
3. Filter evaluation

1. Identification of a reference set (relative recall):

The first step of search filter development is to create the reference set list, which is most often referred to as the gold standard (20). The reference set is a known set of studies that are relevant to the general type of studies under review, in our case, prognostic factor studies. We used the relative recall method, which involves replicating the searches of systematic reviews and using the included studies in these reviews as the reference standard (19). Relative recall is useful as it allows for the inclusion of a broader range of journals and publication years than otherwise could be included practically by manual searching (7, 19). This approach is also more generalizable to topics that are important for our filter, as the literature is spread across a broad range of journals.

We searched for prognostic factor systematic reviews and only included those which: carried out a search on Ovid MEDLINE, did not include a prognosis filter or prognosis terms in the search strategy, and that used a search strategy that was publicly available and reproducible. Additionally, we made sure that the SR´s were related to different topics to allow for generalizability.

2. Search term selection:

Frequency analysis

Search term selection was partially based on the objective method used by Rietjens 2019 (18). A word frequency analysis of PF articles was carried out using the free online software systematic review

accelerator. We separately analyzed the language of both the included and excluded studies of the SRs used for relative recall to create two distinct lists of terms.

Calculate chi square values

Chi square values were calculated for terms generated from the word frequency analysis. From this, we determined the significance of the difference in relative frequencies of the terms in positive studies (the studies that are included in the review) and negative studies (studies not included in the review). As expected, given the small number of studies included, all terms showed non-significant results. Thus, we complemented this frequency analysis with a Delphi panel of experts to reach a consensus on the terms selected for the filter.

Delphi Panel

The Delphi panel consisted of 15 members of various specialties, in particular systematic reviewers, statisticians, clinicians and information retrieval specialists. Each panelist had to evaluate the appropriateness of including each term in the filter. We used the RAND definitions of agreement to classify the terms as appropriate, neutral or inappropriate for use in the filter and also to decide whether this qualification was agreed on by a majority of the panel members (21). The Delphi method consisted of three rounds, the first two being individual ratings and the last round was a panel meeting where a discussion took place on the ratings given to each term. The most relevant methodological terms were extracted from the frequency analysis and made into a list of 80 terms. This list was given to the panel to rate on a scale of 1-9, with 1 being the least appropriate for inclusion and 9 being most appropriate. The terms scoring between 7 and 9 on the Delphi were defined as potentially eligible for inclusion in the filter (21). The consecutive connection of these terms with the Boolean operator OR produced the final search strategy (filter).

3. Filter evaluation:

An essential component of the search filter development process is the evaluation of how well the search filter performs in retrieving relevant records in a systematic review. To carry out the evaluation the filter was combined using the Boolean operator AND with the broad search strategy for Ovid MEDLINE that was used in each included SR.

During the evaluation we tested the sensitivity, specificity, precision, and number needed to read (NNR) of the filter. We used table 1 below to guide us in the evaluation:

Table 1: Table to calculate sensitivity, specificity, precision and NNR of the filter.

	Reference set articles	Non-reference set articles
Retrieved	**A** (True Positive)	**B** (False Positive)
Not retrieved	**C** (False Negative)	**D** (True Negative)

Sensitivity is the number of references in the reference set retrieved by the filter as a proportion of the total number of references in the bibliography (22-24). If the search had low sensitivity, it would miss a large proportion of relevant articles. In contrast, a highly sensitive search is constructed so that it can pick up most of the relevant articles. It was calculated by: **(A/(A+C)) x 100**

Specificity is the number of references that are not relevant and are not retrieved as a proportion of the total number of non-relevant references (22, 24). It was calculated by: **(D/(B+D)) x 100.**

Precision (or positive predictive value PPV) is the number of relevant records retrieved as a proportion of the total number of records retrieved by the filter (22, 24). It was calculated by:
(A/(A+B)) x 100

The number needed-to read (NNR) is a measure of the usability of the filter, as it indicates how many records a searcher must screen for each relevant record retrieved (22-24). In the context of searching, NNR refers to the number of references that have to be manually screened to find one additional relevant article. A relatively high NNR means a lot of references would have to be checked, thus having important resource implications in terms of time and cost, whereas a low NNR means that relevant articles can be identified quicker without having to check large numbers of titles and abstracts. It was calculated by: **(1/precision) x 100**

The measure of workload saved is the percentage of studies that could be screened but can be saved by using the filter. When using the filter, as compared to without the filter, less articles should be retrieved thus saving workload during the screening process. It was calculated by:
((C+D)/(A+B+C+D)) x 100
Table 2 provides a summary of the different performance measures and formulas used in our study.

Table 2: Summary of performance measures and formulas

Performance measure	Formula
Sensitivity	(A/(A+C)) x 100
Specificity	(D/(B+D)) x 100
Precision	(A/(A+B)) x 100
Number needed to read (NNR)	(1/precision) x 100
Workload saved	((C+D)/(A+B+C+D)) x 100

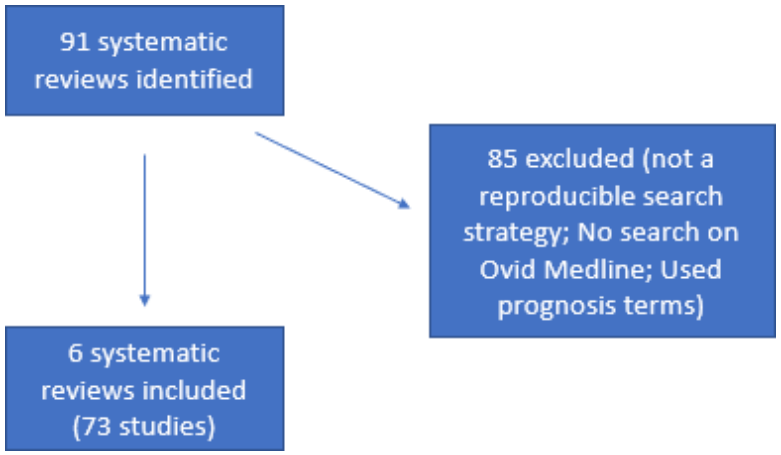
We computed a pooled average of sensitivity, and specificity over the 6 reviews used for evaluation by means of a random effects meta-analysis of proportions using Stata v. 16.0.

Results

1. Identification of a reference set (relative recall):

As outlined in figure 1, our search on PubMed yielded ninety-one SRs of prognostic factors of various topics. We excluded eighty-five SRs due to not having a publicly available and reproducible search strategy, not having carried out a search on Ovid MEDLINE, or for having used prognosis terms in the search strategy. Finally, we formed our reference set with six SRs that met all of our criteria (25-30). Each reference set included between 3 and 22 studies, with a total of 73 studies included in the reference set as a whole. The prognostic factors assessed in these reviews were the following: symptoms of depression, protease activity, sarcopenia, interstitial pneumonia, controlling nutritional status score, and interim PET results.

Figure 1. Flow diagram of reference set search



2. Selection of search terms:

After completing the word frequency analysis, we compiled a list of 80 of the most frequent methodological terms in the prognostic factor reference set. This list of terms was evaluated by the Delphi panel for inclusion in the filter. At the end of the last round of the Delphi we had a list of 8 terms which were deemed appropriate and agreed upon by the panel to include in the filter. We consulted the information retrieval specialists from the Delphi panel about the best way to combine them using MeSH and free text title/abstract terms. We truncated the terms prognostic (prognos*) and predictive (predict*) to be as inclusive as possible in the search. The consecutive connection of these terms with the Boolean operator OR produced the final search strategy (filter) and it is shown below in table 3.

Table 3: Terms included in prognostic factor filter

1	exp Risk/
2	risk.tw.
3	exp Cohort Studies/
4	cohort.tw.
5	exp Prognosis/
6	"prognos*".tw.
7	"predict*".tw.
8	exp Incidence/
9	incidence.tw.
10	exp Survival Analysis/
11	survival.tw.
12	"causal factor".tw.
13	course.tw.
14	or/1-13

3. Filter evaluation:

We evaluated the filter using the relative recall method with the six systematic reviews in our reference set. The filter was added to the end of the search strategy of each SR using the Boolean operator “AND”. The complete search strategy was entered into Ovid MEDLINE and the number of references retrieved was recorded and downloaded into Endnote. To measure the performance of the filter we compared the references retrieved from the original search in the review with our new search using the filter. The performance of the filter in each review is shown in table 4.

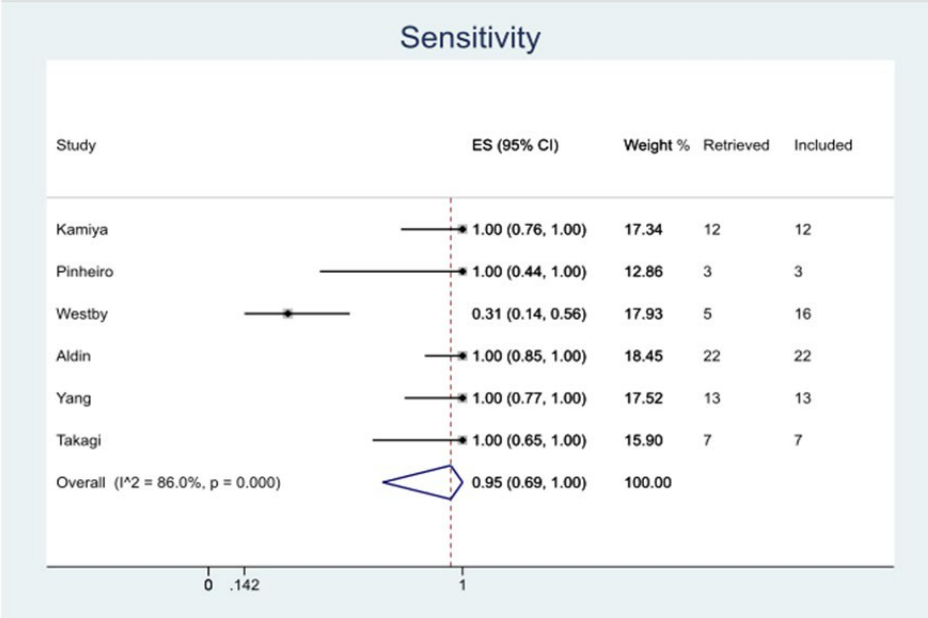
Table 4: Results for sensitivity, specificity, precision, NNR and NNS of the filter evaluated in each review.

Study	Number of included studies in SR	Number of studies retrieved in original SR search	Sensitivity (%)	Specificity (%)	Precision (%)	NNR	Workload saved (%)
Kamiya	12	120	100	46	17	6	42
Pinheiro	3	1314	100	37	0.4	278	37
Westby	16	784	31.25	70	2	47	70
Aldin	22	5562	100	35	0.6	164	35
Yang	13	565	100	44	4	26	41
Takagi	7	100	100	14	8	12	13

Note: SR- systematic review; NNR- number needed to read.

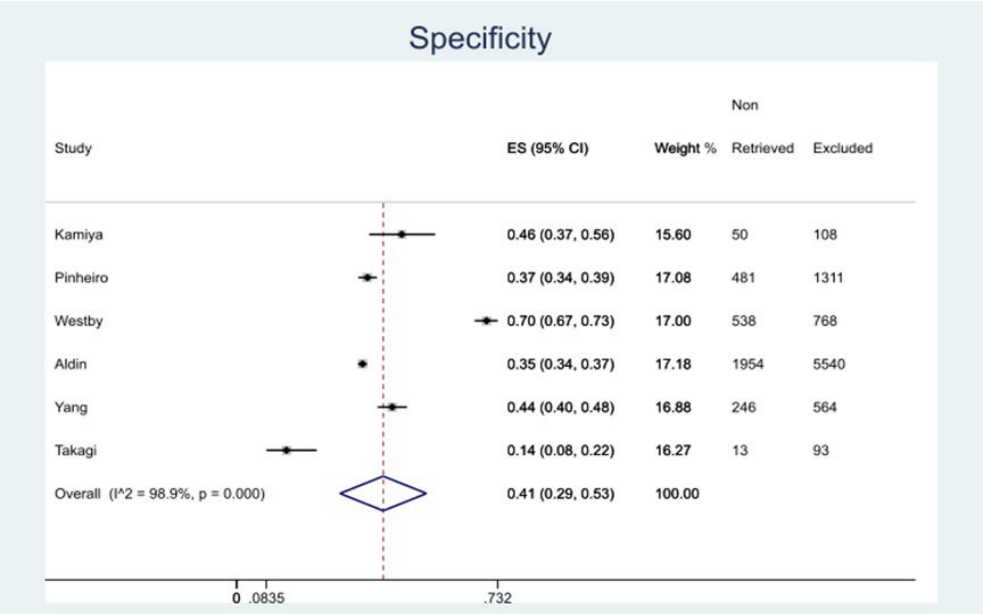
The filter had a sensitivity of 100% in all reference sets except for Westby 2018, which had a low sensitivity of 31%. As can be seen below in figure 2, the filters overall sensitivity was calculated to be 95% (95% CI 69%-100%).

Figure 2: Sensitivity of the filter in various systematic review searches.



The specificity varied from 14-70%, with the highest performance of specificity being in Westby 2018 and the lowest in Takagi 2019. As seen below in figure 3, the overall specificity was calculated to be 41% (95% CI 29-43%). The precision performance also varied considerably ranging from 0.4 (30) to 17% (29). The NNR value varied largely among reviews ranging from 6 to 278. Time saving was substantial ranging from 13% (Takagi) to 70% (Westby).

Figure 3: Specificity of filter in various systematic review searches



Discussion

Main findings

We aimed to develop and test a search filter for finding studies about the role of PFs in Ovid MEDLINE. Overall, the obtained filter showed an excellent sensitivity to retrieve studies from a reference set constructed from studies included in relevant systematic reviews in the field. Specificity was much lower with an overall combined specificity of 41%. Precision ranged from 0.36 to 17%, but it is important to note that efforts to optimise recall has a direct impact on the screening burden (total number of references retrieved) and may not be an appropriate indicator to measure performance of approaches focusing on sensitivity. Resulting from these statistics, the number of references required to screen to retrieve a relevant article varied hugely, from 6 to 277. We calculated that, when using the filter, the screening workload would be lower in all reference sets (13 to 70%).

Out of the six reviews in which we tested the filter, Westby 2018 was the only review where the filter was not effective in retrieving all of the reference set studies. It was a Cochrane review on protease activity as a prognostic factor for healing wounds (25). After examining the studies that weren´t retrieved, we observed that they did not use any of the search terms attributable to prognosis and their approach was not obvious for usual prognostic factor studies. Those studies had terms such as influence or associated that could make them in some way related to prognosis. Another possible explanation for the low sensitivity in Westby 2018 could be that the review authors were generous or lenient with the studies that they included in the review. When examining the flow diagram of Westby 2018, they screened a lot of full texts (10% of the titles and abstracts screened were passed on to the full text stage). In comparison, most of the other reviews in the reference set only passed on 2-3% of

studies to the full text stage, thus they were seemingly stricter with the prognostic factor study criteria. When we added our filter to the other systematic review strategies, the sensitivity was 100%, as all included studies were retrieved.

Comparison with available prognosis filters

There are a few published filters for prognosis studies which focus on prognostic models and prediction rules. We compared our prognostic factor filter with the Haynes broad prognosis filter (31): (incidence[MeSH:noexp] OR mortality[MeSH Terms] OR follow up studies[MeSH:noexp] OR prognos*[Text Word] OR predict*[Text Word] OR course*[Text Word]). We chose this filter as a comparison since it is the one that is most available to people who use PubMed. In general, the filter is known to have a sensitivity of 90% and specificity of 80%. We evaluated this filter in our reference set. As can be seen below in table 5 the filter was less sensitive overall than our PF filter (74% 95%CI (0.45 - 0.96)), but it was more specific (0.63 95%CI (0.51 - 0.74)). All of the SR´s in our reference set had a similar precision performance as the Haynes filter. This is because the reference sets had very low numbers of included studies and this statistic is dependent on this. More workload can be saved using the Haynes filter, but that is at a risk of losing potentially relevant studies to include in the review.

Table 5: Results from Haynes sensitive broad filter in our reference set

Study	Sensitivity (%)	Specificity (%)	Precision (%)	NNR	Workload saved (%)
Kamiya 2019	75	63	18.4	5	59
Pinheiro 2015	100	67	0.7	146	67
Westby 2018	13	87	1.9	52	87
Aldin 2019	80	56	1.5	68	56
Yang 2019	92	62	5.4	19	60
Takagi 2019	86	35	9.1	11	34

Note: NNR- number needed to read

Strengths and limitations

Our relative recall references included various topics, thus allowing us to evaluate the filter over many different clinical situations. If the references in the reference set are from one area only it can lead to subject bias in the filter (working well in some subjects, but not others). Through using the relative recall method, we were able to ensure that each study in the reference set was in fact a prognostic factor study. It can be difficult to decipher prognostic factor studies from other studies at times, so since we were using studies that were included in prognostic factor systematic reviews, we could be assured that they were truly prognostic factor studies.

When developing the protocol for this study, we realized that there were many different methods that researchers have used in the past to create a search filter. We examined all the published methods and weighed up our options before deciding on which methods to follow. If we had more resources, time, and manpower available there are more robust methods that we could employ in the future. These other methods include creating a larger reference set of PF studies or creating a traditional gold standard through manually searching for studies.

Implications for research

This filter has a high sensitivity so we can be assured that the risk of missing a study is very low. However, as we noted with the studies in Westby 2018 (25), not all PF studies include typical prognostic words, so we still need to think carefully about what kind of studies we might be searching for and if they will include the correct terms. The use of the filter in search strategies could decrease the number of studies needed to be manually screened. Many times, search strategies for PF systematic reviews yield large numbers of studies from the search, for example 20,000-100,000 references. Thus, it can take a lot of time (6-8 weeks) and resources to screen through them all, making the NNR an important statistic. This PF filter needs to be evaluated in rapid reviews, as time constraints in these reviews make efficient searches even more necessary.

Conclusions

To the best of our knowledge, no search filter exists for locating PF studies in Ovid MEDLINE, nor in any other online database. Our filter had a high sensitivity of 95% overall in the systematic reviews in which we tested it. Its specificity on the other hand, was lower at 41% overall. Our aim was to create a sensitive filter as we feel the most important part of search filter development is to not lose any relevant studies in the search. Further research is still needed on this topic to increase the specificity of the filter, while keeping its high sensitivity.

Acknowledgements: We would like to thank Karl Moons for his input and suggestions. We would like to acknowledge the Delphi panel members whose input has been invaluable to the project. Delphi panel: Alfonso Muriel, Borja Félix Fernández, Alba Antequera, John Allotey, Jo Jordan, Candida Fenton, Federico Gordo, Marta Roque, Ana Halperin, José Ignacio Pijoan Zubizarreta, José Molina Pérez, Esther García, Sergio Serrano.

Conflicts of interest: The authors declare that they have no competing interests.

Funding: This work has been partially funded by the SEXCOMPLEX project (grant number PIE16/00050) and Biomarkers in sepsis project (grant PI19/00481) form the Instituto de Salud Carlos III (Plan Estatal de I + D + i 2013–2016) and co-financed by the European Development Regional Fund. "A way to achieve Europe" (ERDF). These funding sources had no role in the conception, design, execution or decision to submit this work.

Authors' contributions: All authors read and approved the final manuscript

References

1. Prady SL, Uphoff EP, Power M, Golder S. Development and validation of a search filter to identify equity-focused studies: reducing the number needed to screen. BMC Med Res Methodol. 2018;18(1):106.
2. Shariff SZ, Sontrop JM, Haynes RB, Iansavichus AV, McKibbon KA, Wilczynski NL, et al. Impact of PubMed search filters on the retrieval of evidence by physicians. Cmaj. 2012;184(3):E184-90.
3. Glanville J, Dooley G, Wisniewski S, Foxlee R, Noel-Storr A. Development of a search filter to identify reports of controlled clinical trials within CINAHL Plus. Health Info Libr J. 2019;36(1):73-90.
4. Glanville J, Kotas E, Featherstone R, Dooley G. Which are the most sensitive search filters to identify randomized controlled trials in MEDLINE? J Med Libr Assoc. 2020;108(4):556-63.
5. Tugwell P, Knottnerus JA. Current study design labels are confusing! Call for consensus on better terminology that clearly reflects specific features. J Clin Epidemiol. 2017;91:1-3.
6. Kastner M, Wilczynski NL, McKibbon AK, Garg AX, Haynes RB. Diagnostic test systematic reviews: bibliographic search filters ("Clinical Queries") for diagnostic accuracy studies perform well. J Clin Epidemiol. 2009;62(9):974-81.
7. Beynon R, Leeflang MM, McDonald S, Eisinga A, Mitchell RL, Whiting P, et al. Search strategies to identify diagnostic accuracy studies in MEDLINE and EMBASE. Cochrane Database Syst Rev. 2013(9):Mr000022.
8. Boulos L, Ogilvie R, Hayden JA. Search methods for prognostic factor systematic reviews: a methodologic investigation. J Med Libr Assoc. 2021;109(1):23-32.
9. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. Bmj. 2013;346:e5595.
10. Hingorani AD, Windt DA, Riley RD, Abrams K, Moons KG, Steyerberg EW, et al. Prognosis research strategy (PROGRESS) 4: stratified medicine research. Bmj. 2013;346:e5793.
11. Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. PLoS Med. 2013;10(2):e1001380.
12. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med. 2013;10(2):e1001381.
13. Geersing GJ, Bouwmeester W, Zuithoff P, Spijker R, Leeflang M, Moons KG. Search filters for finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews. PLoS One. 2012;7(2):e32844.
14. Chatterley T, Dennett L. Utilisation of search filters in systematic reviews of prognosis questions. Health Info Libr J. 2012;29(4):309-22.
15. López-Alcalde J, Stallings EC, Zamora J, Muriel A, Doorn S, Alvarez-Diaz N, et al. Sex as a prognostic factor for mortality in adults with acute symptomatic pulmonary embolism. Cochrane Database of Systematic Reviews. 2021(1).
16. Lopez-Alcalde J, Antequera Martín A, Stallings E, Muriel A, Fernández-Félix B, Solà I, et al. Evaluation of the role of sex as a prognostic factor in critically ill adults with sepsis: systematic review protocol. BMJ Open. 2020;10(5):e035927.
17. Glanville J, Bayliss S, Booth A, Dundar Y, Fernandes H, Fleeman ND, et al. So many filters, so little time: the development of a search filter appraisal checklist. J Med Libr Assoc. 2008;96(4):356-61.
18. Rietjens JA, Bramer WM, Geijteman EC, van der Heide A, Oldenmenger WH. Development and validation of search filters to find articles on palliative care in bibliographic databases. Palliat Med. 2019;33(4):470-4.

19. Sampson M, Zhang L, Morrison A, Barrowman NJ, Clifford TJ, Platt RW, et al. An alternative to the hand searching gold standard: validating methodological search filters using relative recall. BMC Med Res Methodol. 2006;6:33.
20. Jenkins M. Evaluation of methodological search filters--a review. Health Info Libr J. 2004;21(3):148-63.
21. Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR. The RAND/UCLA appropriateness method user's manual. Rand Corp Santa Monica CA; 2001.
22. van de Glind EM, van Munster BC, Spijker R, Scholten RJ, Hooft L. Search filters to identify geriatric medicine in Medline. J Am Med Inform Assoc. 2012;19(3):468-72.
23. Kok R, Verbeek JA, Faber B, van Dijk FJ, Hoving JL. A search strategy to identify studies on the prognosis of work disability: a diagnostic test framework. BMJ Open. 2015;5(5):e006315.
24. Cooper C, Varley-Campbell J, Booth A, Britten N, Garside R. Systematic review identifies six metrics and one method for assessing literature search effectiveness but no consensus on appropriate use. J Clin Epidemiol. 2018;99:53-63.
25. Westby MJ, Dumville JC, Stubbs N, Norman G, Wong JK, Cullum N, et al. Protease activity as a prognostic factor for wound healing in venous leg ulcers. Cochrane Database Syst Rev. 2018;9(9):Cd012841.
26. Aldin A, Umlauff L, Estcourt LJ, Collins G, Moons KG, Engert A, et al. Interim PET-results for prognosis in adults with Hodgkin lymphoma: a systematic review and meta-analysis of prognostic factor studies. Cochrane Database Syst Rev. 2019;9(9):Cd012643.
27. Takagi K, Domagala P, Polak WG, Buettner S, Ijzermans JNM. Prognostic significance of the controlling nutritional status (CONUT) score in patients undergoing hepatectomy for hepatocellular carcinoma: a systematic review and meta-analysis. BMC Gastroenterol. 2019;19(1):211.
28. Yang M, Shen Y, Tan L, Li W. Prognostic Value of Sarcopenia in Lung Cancer: A Systematic Review and Meta-analysis. Chest. 2019;156(1):101-11.
29. Kamiya H, Panlaqui OM. Systematic review and meta-analysis of the prognosis and prognostic factors of interstitial pneumonia with autoimmune features. BMJ Open. 2019;9(12):e031444.
30. Pinheiro MB, Ferreira ML, Refshauge K, Maher CG, Ordoñana JR, Andrade TB, et al. Symptoms of depression as a prognostic factor for low back pain: a systematic review. Spine J. 2016;16(1):105-16.
31. Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: an analytic survey. BMC Med. 2004;2:23.

5. 3rd Article: Sex as a prognostic factor in systematic reviews: Challenges and lessons learned

During the process of carrying out a SR on sex as a PF in patients with sepsis (65) and writing the protocol for the Cochrane PF review, I gained significant experience in PF reviews and dealing with sex as the PF. I wrote this article to give insights into the challenges we encountered when conducting SRs of sex as a PF and to describe how we overcame these obstacles. For the methods of these reviews we modified various sections to facilitate sex as a PF, such as the PF section of CHARMS-PF (used for data extraction), certain sections of QUIPS (for risk of bias), and we extracted data on the sex and gender terms used throughout the studies.

I hope to see more systematic reviews of sex as a PF being published in the future and hope this article will be useful to authors in adjusting their methods appropriately. Realizing the importance of studying sex as a PF is a critical step in the right direction towards precision medicine that will help reduce healthcare inequities and benefit both males and females alike.

Journal of *Personalized Medicine*

MDPI

Article

Sex as a Prognostic Factor in Systematic Reviews: Challenges and Lessons Learned

Elena Stallings [1,2,*], Alba Antequera [3], Jesús López-Alcalde [1,2,4,5], Miguel García-Martín [2,6], Gerard Urrútia [2,3] and Javier Zamora [1,2,7]

1 Clinical Biostatistics Unit, Instituto Ramón y Cajal de Investigación Sanitaria, 28034 Madrid, Spain; cochrane.madrid@ufv.es (J.L.-A.); javierza@gmail.com (J.Z.)
2 CIBER Epidemiología y Salud Pública (CIBERESP), 28029 Madrid, Spain; mgar@ugr.es (M.G.-M.); GUrrutia@santpau.cat (G.U.)
3 Sant Pau Institute for Biomedical Research (IIB Sant Pau), 08041 Barcelona, Spain; AAntequera@santpau.cat
4 Faculty of Health Sciences, Universidad Francisco de Vitoria, 28223 Madrid, Spain
5 Institute for Complementary and Integrative Medicine, University Hospital Zurich, University of Zurich, CH-8091 Zurich, Switzerland
6 Department of Preventive Medicine and Public Health, University of Granada, 18010 Granada, Spain
7 WHO Collaborating Centre for Global Women's Health, Institute of Metabolism and Systems Research, University of Birmingham, Birmingham B15 2TT, UK
* Correspondence: elenastallings.pubhealth@gmail.com

Abstract: Sex is a common baseline factor collected in studies that has the potential to be a prognostic factor (PF) in several clinical areas. In recent years, research on sex as a PF has increased; however, this influx of new studies frequently shows conflicting results across the same treatment or disease state. Thus, systematic reviews (SRs) addressing sex as a PF may help us to better understand diseases and further personalize healthcare. We wrote this article to offer insights into the challenges we encountered when conducting SRs on sex as a PF and suggestions on how to overcome these obstacles, regardless of the clinical domain. When carrying out a PF SR with sex as the index factor, it is important to keep in mind the modifications that must be made in various SR stages, such as modifying the PF section of CHARMS-PF, adjusting certain sections of QUIPS and extracting data on the sex and gender terms used throughout the studies. In this paper, we provide an overview of the lessons learned from carrying out our reviews on sex as a PF in different disciplines and now call on researchers, funding agencies and journals to realize the importance of studying sex as a PF.

Keywords: sex; gender; prognosis; prognostic factor; systematic review; methods

Citation: Stallings, E.; Antequera, A.; López-Alcalde, J.; García-Martín, M.; Urrútia, G.; Zamora, J. Sex as a Prognostic Factor in Systematic Reviews: Challenges and Lessons Learned. *J. Pers. Med.* **2021**, *11*, 441. https://doi.org/10.3390/jpm11060441

Academic Editor: Mary V. Seeman

Received: 28 March 2021
Accepted: 18 May 2021
Published: 21 May 2021

1. Introduction

People are living longer, with one or more health problems; prognosis research is thus vital for explaining and predicting future clinical outcomes in people with existing health conditions. Prognosis research aims to summarize and predict relevant outcomes such as death, recovery, recurrence, disability, or quality of life. In the past 10 years, research on prognosis has rapidly increased [1–4] along with many novel studies and new methods being developed. However, results from different studies are often contradictory, making it difficult to assess a specific prognostic factor (PF). This is where systematic reviews come into play. Nevertheless systematic reviews of PFs have received little attention by scientists to date. In clinical medicine, we are starting to see a transition from a universal medicine that has a one-size-fits-all approach to personalized medicine. Personalized medicine is a unique individualized approach to treatment based on a patient's diagnosis and prognosis [5]. This intertwinement has led to theragnostics, which is the connection of diagnosis and therapeutics addressed to people on an individual basis [6]. This novel connection can provide better prognoses relying on specific features, i.e., PFs. Genetic information plays an important role in theragnostics and pharmacogenetics—which is the

study of how people respond differently to drug therapies based on their genes—and helps to individually tailor treatments [7]. There are underlying genetic mechanisms to most sex differences in disease [8], which suggests that sex is an excellent candidate as a PF.

Research on PFs is becoming more widespread, and its importance in clinical practice is gradually being recognized. A PF "is any measure that, among people with a given health condition (that is, a start point), is associated with a subsequent clinical outcome (an endpoint)" [9]. Therefore, PFs can distinguish groups of people with a different average prognosis. PF research has a wide variety of applications related to both clinical and public health research [9]. For example, for many cancer patients, tumour grade at the time of diagnosis is a prognostic factor, as each group of patients with the same tumour grade should have broadly similar outcomes [3]. Also, a high body mass index (BMI) is a PF for worse outcomes in patients diagnosed with COVID-19 [10]. Similarly, male sex is a poor PF in non-small-cell lung cancer [11,12] and in gastric cancer, as females experience a better survival rate [13]. On the other hand, female sex is a poor PF for mortality in acute myocardial infarction [14,15]. Thus, minimal clinically relevant differences associated with patients' sex may have a great impact on the understanding of disease processes, the applicability of the findings to specific patient groups, and the planning of future research [16]. Sex refers to the biological, genetic, and physiological processes that generally distinguish females from males, while gender refers to the roles, relationships, behaviours, and other traits that societies typically attribute to women, men, and people of diverse gender identities (e.g., transgender people) [17]. Sex is, with age, the most common baseline factor collected in the context of randomized and non-randomized studies, regardless of whether a study addresses a therapeutic, etiologic, diagnostic, or prognostic topic. Therefore, sex has clearly the potential to be evaluated as a PF in almost all clinical areas [8,10,11]. This issue is typically assessed in primary studies but is generally not a considered topic in systematic reviews.

Systematic reviews (SRs) are the cornerstone of evidence-based medicine as they play a major role in summarizing the available body of evidence and in identifying knowledge gaps [18]. Accordingly, addressing sex-related findings in systematic reviews is important to better guide clinical practice and tailor patient care to provide the optimal treatment for different sexes. In contrast, in primary studies or systematic reviews, not considering how meaningful such differences between sexes are can lead to poorer healthcare quality, limiting the generalizability of study results and promoting inequities. In recent years, research on sex as a PF has increased [1–4]; however, this influx of new studies frequently shows conflicting results across the same treatment or disease state.

Sex differs from other PFs in that it is difficult to verify through tests and it generally does not change over time. We have written this article to offer insights into the challenges we encountered when conducting SRs on sex as a PF and to provide suggestions on how to overcome these obstacles, regardless of the outcome or clinical domain. We carried out two systematic reviews on sex as a PF. The first review studied the prognostic role of sex on mortality outcomes in sepsis [19], while the second one looked at the role of sex in the prognosis of patients with acute pulmonary embolism [20]. In the current paper, we will comment on the lessons learned from these two reviews and, in the methodological challenges section, we will present three SRs studying sex as a PF, providing examples and making comparisons. Our objective is to discuss the methodological challenges we encountered and reflect on the lessons learned in carrying out these reviews. Therefore, future systematic reviewers will be able to learn from our experiences and use the same framework whilst investigating sex as a prognostic factor. Primary researchers will also be able to benefit from this paper, as they will understand the difficulties that reviewers encounter when trying to synthesize this type of studies. For example, authors may become aware of what terms should be used in abstracts and titles to maximize the likelihood of their study being captured in review searches.

2. Importance of Sex as a PF and Its Demarcation from Gender

Sex is a biological attribute that is associated with physical and physiological features including gene expression, hormone function and reproductive and sexual anatomy [17,21,22]. Sex, typically assigned at birth based on the appearance of external genitalia, is defined as female, male, intersex, etc. However, it is often mislabelled as gender. Sex and gender are interconnected but vastly different. In comparison to sex, gender refers to the socially constructed roles, behaviours and identities of female, male and gender-diverse people [22] and to the terms men and women or boys and girls. Both sex and gender play roles as prognostic factors in various illnesses (cardiovascular disease, sepsis, cancer) [15,19,23]. Therefore, it is important to distinguish between them when studying sex or gender as a prognostic factor.

Many differences exist between the sexes, mostly due to genetics and hormones (different levels of androgens and oestrogens). Many illnesses are characterized by a higher incidence in one sex versus the other. For example, 99% of people diagnosed with breast cancer are females [24]. In the same manner, four times more females than males are diagnosed with osteoporosis [25]. Other illnesses will occur at the same rate in both sexes, but they can manifest differently according to sex. For example, in schizophrenia, the disease usually starts at an earlier age and with severer symptoms in males [26]. In myocardial infarctions, again the first myocardial infarction is experienced at a younger age by males than by females, and females tend to present with symptoms of nausea and shortness of breath instead of the usual chest pain. These sex differences in disease incidence and in diagnosing illness also predict differences in prognosis. Just as differences have been found in the manifestation of myocardial infarction in females, it has also been found that females tend to have poorer outcomes [27–29]. Similarly, in our review, we found an independent prognostic impact of sex on mortality, although in this case the certainty of evidence was very low [19].

Prognosis research has increased in the past decade, and the same can be said of sex and gender research [1–4]. However, in general, there are not many SRs on PFs. For example, when we searched the Cochrane database of systematic reviews, we found three completed PF SRs [30–32]. In comparison to intervention reviews, which is the more traditional review type with thousands of reviews completed, this is a novel area of research and evidence synthesis. Thus, taking into consideration that so few reviews have been published on PFs and sex separately and that both areas of research are rapidly developing, it is understandable that there is a lack of SRs on sex as a PF. However, we did find a few SRs evaluating the role of sex, such as Bougouin et al., Giuliano et al. and Kim et al., in various illnesses, and thus we were able to compare and contrast these reviews and review the methods that they used [33–35].

3. Methodological Challenges

3.1. Search and Selection

A search of reviews on sex as a PF retrieves thousands of references; therefore, it is important to use a search filter. We added a sex and gender filter ("sex factors" OR "sex distribution" OR "Sex characteristics" OR "Sex ratio" OR sex OR "women's health" OR "men's health") OR TITLE: (boy* OR male* OR girl* OR female* OR gender OR women OR men OR sex) to the search strategy in an attempt to narrow down the search field. By adding the filter, the number of studies retrieved from the searches was reduced by 20%. In combination with searching electronic databases, we also hand-searched conferences. For conferences on sex and gender, we found the congress "Organization for the study of sex differences". We hand-screened 8 years of abstracts from this conference and did not retrieve any study that met all inclusion criteria. While retrieving unpublished studies from conference abstracts is considered good practice in systematic review development, it is important to consider the expected large number of results from the electronic database searches. Thus, we encourage review authors to choose to extend their search to conference proceedings based on their resource availability.

When screening studies, both sexes must be included for a study to be eligible for inclusion, as it is impossible to measure the prognostic significance of sex while only looking at one sex. As mentioned above, sex refers to genotypic, phenotypic and physiological characteristics, including chromosomes, gene expression, hormone levels and reproductive and/or sexual anatomy [17]. In our reviews, we accepted any assessment of sex and evaluated the appropriate use of the terms sex and gender when applicable. We were aware that the terms 'sex' and 'gender' are poorly described and defined in the majority of published articles. Thus, when no additional information was provided, if it was clear that the authors were referring to sex but mistakenly used the terms for gender, we assumed that the study was considering sex. If the authors explicitly stated that they evaluated the social aspect, then we considered that they were evaluating gender and not sex.

3.2. Data Extraction

For data extraction, we used the CHARMS-PF (critical appraisal and data extraction for systematic reviews of prediction modelling studies for prognostic factors) template [9]. CHARMS-PF is a checklist of key data to be extracted from primary PF studies. It is based on additions and modifications of the original CHARMS data extraction sheet for prediction modelling [36]. In the CHARMS-PF extraction sheet, there is a section created for PFs (index and comparator factors). In this section, we extracted the PF definition and method of measurement of the PF. We accepted any definition of sex (our PF of interest) and any method of sex measurement given by the authors. The timing of PF measurement does not matter when studying sex in primary studies or reviews, as it is not normally a temporal variable that may change. We extracted information on the use of the terms sex and gender in each study to evaluate if the terms were being used adequately in the primary studies. These data are important to extract and take note of, as the lack of literacy surrounding the terms for sex and gender should be highlighted in SRs.

Bougouin et al. did not use CHARMS for data extraction [33]. However, CHARMS, is relatively new (2014) and was only published a year before the publication of this review [36]. Thus, the authors may had previously planned their data extraction methods in a protocol and did not change them. However, they did extract the adjusted data, though they did not extract many data on gender, their prognostic factor of interest. Giuliano et al. did not mention the use of CHARMS but created their own data extraction template [35]. Kim et al. also did not use CHARMS, as it was not yet published [34].

3.3. Risk of Bias Assessment

A critical step in carrying out a systematic review is assessing the risk of bias of the included studies. Tools used to measure quality are ROB and ROB2 for randomized trials and PROBAST for prediction model studies [37–40]. To assess the risk of bias in PF studies, the "Quality in Prognosis Studies" (QUIPS) tool was created [41,42]. The tool consists of several prompting questions within six different domains, each domain being judged on a three-grade scale. Hayden et al. determined six key domains for the risk of bias appraisal included in PF studies: study participation, study attrition, PF measurement, other prognostic factor adjustment, outcome measurement and analysis and reporting [42].

We used an amendment to the QUIPS tool proposed by Aldin and colleagues [30] using four categories (low, moderate, high and unclear risk) instead of the initial three categories (low, moderate and high). In SRs of sex as a PF, the unclear category may be especially relevant, since some signalling items of QUIPS, such as those related to PF domains with a high likelihood of lack of sex definition, have a limited value for the assessment and rating. Therefore, rating as unclear risk may be the fairest alternative. Following on from the rating amendment, we also made some slight modifications to the QUIPS sections to adapt it for sex as a PF, which are highlighted in Table 1. Some items were particularly hard to differentiate, and a learning phase was required to increase the interrater agreement.

Table 1. QUIPS modifications for studying sex as a prognostic factor.

Domains	QUIPS	QUIPS Modified for Sex as PF	Comments
1. Study participation	Description of the baseline study sample	Baseline number and characteristics of participants by sex are clearly described and reported separately for males and females	The regular QUIPS refers to a description of the baseline sample in general (both sexes combined); however, we specified that it was necessary to have the participants characteristics described by sex. Example: Females (N): race of females (N), obesity in females (N). Males (N): race of males (N), obesity in males (N).
2. Study attrition	Adequate description of participants lost to follow-up	Key characteristics of participants lost to follow-up are provided separately for males and females	The key characteristics of the lost-to-follow-up participants must be recorded by sex. N of females and N of males per characteristic. However, this was never reported.
3. Prognostic factor measurement	Clear definition or description of the PF	Clear definition or description of sex	The authors must provide an adequate definition for the prognostic factor, in this case sex [1].
	Adequately valid and reliable method of measurement	Not applicable	We do not anticipate specific sex measurement for this type of research question.
	Continuous variables reported or appropriate cut points used	Not applicable	Sex measurement is not a continuous variable.
	Same method and setting of measurement used in all study participants	Not applicable	We do not anticipate method and setting measurement for this type of research question.
	Adequate proportion of the study sample had complete data	Not applicable	We do not anticipate missing data of sex measurement for this type of research question.
	Appropriate methods of imputation were used for missing data	Not applicable	We do not anticipate missing data of sex measurement for this type of research question.
4. Outcome measurement		No differences in this domain.	
5. Adjustment for other prognostic factors		No differences in this domain.	
6. Statistical analysis and reporting		No differences in this domain.	

[1] We considered an adequate definition as listing any of the following: sex for biological characteristics; gender for socially constructed roles, behaviours, and identities; females or males for sex; women or men for gender.

In contrast, Bougouin et al. did not use QUIPS for quality assessment, but again this could be due to the short time frame between QUIPS and the review being published [33]. Giuliano et al. used QUIPS to assess the risk of bias of the included studies; however, they did not mention any modifications being made to the tool [35]. Kim et al. did not use QUIPS in their assessment of risk of bias, as it was not yet published [34].

3.4. Data Analysis

The studies incorporated in a systematic review of sex as a PF need to be similar in terms of population (ages, ethnicity, etc.), index factor measurement (sex) and outcome measurement (how the outcomes are measured, for example, mortality in 30 days, 90 days, etc.): a meta-analysis must combine the results from sufficiently homogenous individual studies to provide a meaningful pooled prognostic effect estimate. If the studies are not homogenous in design, we may carry out a subgroup analysis or meta-regression. Studies may report the crude association between PF, sex, and the outcome or the adjusted association, where one adjusts for the contribution of other PFs compared to the index factor (here, sex). However, we did not require the consideration of the complete core

set of additional PFs. In our experience, if the researchers adjusted for at least one of our pre-defined confounding factors, then the study was valid for inclusion in the review and the meta-analysis. Deciding the pre-defined confounding factors was complicated, as sex is a factor that is present from birth; therefore, it is complex to define what is a confounding factor of sex. To make a list of the most important confounding factors, we created a Delphi panel of reviewers and clinicians to decide on which factors to include. Our data extraction and risk of bias assessments considered the confounding factors that were measured, controlled for (by the study design) and adjusted for (in the analysis).

An additional part of the analysis in reviews studying sex as a PF is to analyse the sex/gender terminology used. To judge if the terms sex and gender were used correctly in the primary studies, we conducted a frequency analysis of the results. In our SR on sex as PF in patients with sepsis, we included 13 studies [19]. No primary study included in our review defined sex correctly. Twelve of the included studies in our review had an inadequate use of sex and gender terms, using all the terms interchangeably throughout the study. The correct usage of terms was unclear in the remaining study, as it used the term gender and all the related terminology for gender; however, from the study context, it could be presumed that the study authors were in fact referring to sex [43].

In some SRs, there are discrepancies and interchangeability of the sex and gender terms. The correct terms for sex are male and female, and those for gender are boy, girl, man or woman. Many published reviews use the word gender when referring to sex, thus making it confusing for readers. Authors feel that they are being repetitive and do not realise that sex and gender are two distinct terms. For example, in Bougouin et al. and Kim et al., the authors use the terms gender, men and women consistently but in reality, they are discussing topics related to sex, not gender [33,34]. In Kim et al., the authors state "women tend to have smaller coronary arteries than men" [34]. This is a sex difference, not a gender difference and should instead read "female patients tend to have . . . than male...". Likewise, Giuliano et al. talk about sex differences whilst referring to men and women. In other parts of the paper, they also use the terms male and female, thus making their usage of the terminology incorrect and inconsistent [35].

4. Concluding Remarks

The role of sex in human health and medical research continues to be understudied, as sex-based medicine is often viewed as a specialist niche instead of being central to all medical research [16]. We must bring sex- and gender-based analysis to the forefront of research and base future research around it. Systematic reviews evaluating the role of sex as a PF fosters rigorous, reproducible, inclusive and responsible science.

When carrying out a PF SR with sex as the index factor, it is important to keep in mind the adaptations that must be made in various SR stages. This is outlined in Table 2 below and includes modifying the PF section of CHARMS-PF, adjusting certain sections of QUIPS and extracting data on the terms sex and gender used throughout the studies. The lack of literacy regarding the sex and gender terms needs to be addressed, as this is a widespread problem among researchers. It is especially important that researchers wishing to study and publish sex and gender research understand the differences between these concepts and use the correct terminology. This lack of understanding can have serious implications in prognosis research, such as creating confusion among investigators and the general public.

There are methods available to rigorously synthesize the role of sex as a PF. We hope to see more systematic reviews of this kind in the future. In this paper, we have provided an overview of the lessons we learned from carrying out our reviews on sex as a PF in different disciplines and we now call on researchers, funding agencies, journals and research institutions to acknowledge the importance of studying sex as a PF. Realizing this importance is a critical step in the right direction towards precision medicine that will help reduce health inequities and benefit both males and females alike.

Table 2. Summary of challenges and solutions in a systematic review evaluating the role of sex as a prognostic factor.

	Challenge	Solution
1. Search and selection	Too many references retrieved in the search.	Add a sex and gender search filter to the search
2. Data extraction	Sections of CHARMS-PF not totally compatible with sex as a prognostic factor (PF)	Take the following into consideration: • Accept any definition of sex and any method of sex measurement given by authors • Timing of PF measurement is not important, as sex is not normally a temporal variable that may change • Extract data on the use of the terms sex and gender
3. Risk of Bias	Sections of QUIPS not compatible with sex as a PF	Specific modifications in QUIPS tool as listed in Table 1
4. Data analysis	Deciding the confounding factors for sex as a PF	Delphi panel (expert input) to aid in this decision-making process
	Sex and gender terms used inadequately and interchangeably in many primary studies	Analyse the sex and gender terminology used in primary studies

Author Contributions: Author contributions: Conceptualization, E.S. and J.Z.; Writing—Original Draft Preparation, E.S.; Writing—Review & Editing, E.S., A.A., J.L.-A., G.U., M.G.-M., J.Z.; All authors have read and agreed to the published version of the manuscript.

Funding: The SEXCOMPLEX project was supported by Instituto de Salud Carlos III (Plan Estatal de I + D + i 2013–2016) and co-financed by the European Development Regional Fund. "A way to achieve Europe" (ERDF) grant number PIE16/00050. This funding source had no role in the conception, design, execution or decision to submit this work.

Acknowledgments: We would like to thank Karl Moons of the Cochrane Prognosis Methods Group for his advice on the writing of this article. We would also like to thank all researchers who were involved in the carrying out of our two prognostic factor systematic reviews on sex.

Conflicts of Interest: The authors declare no conflict of interest.

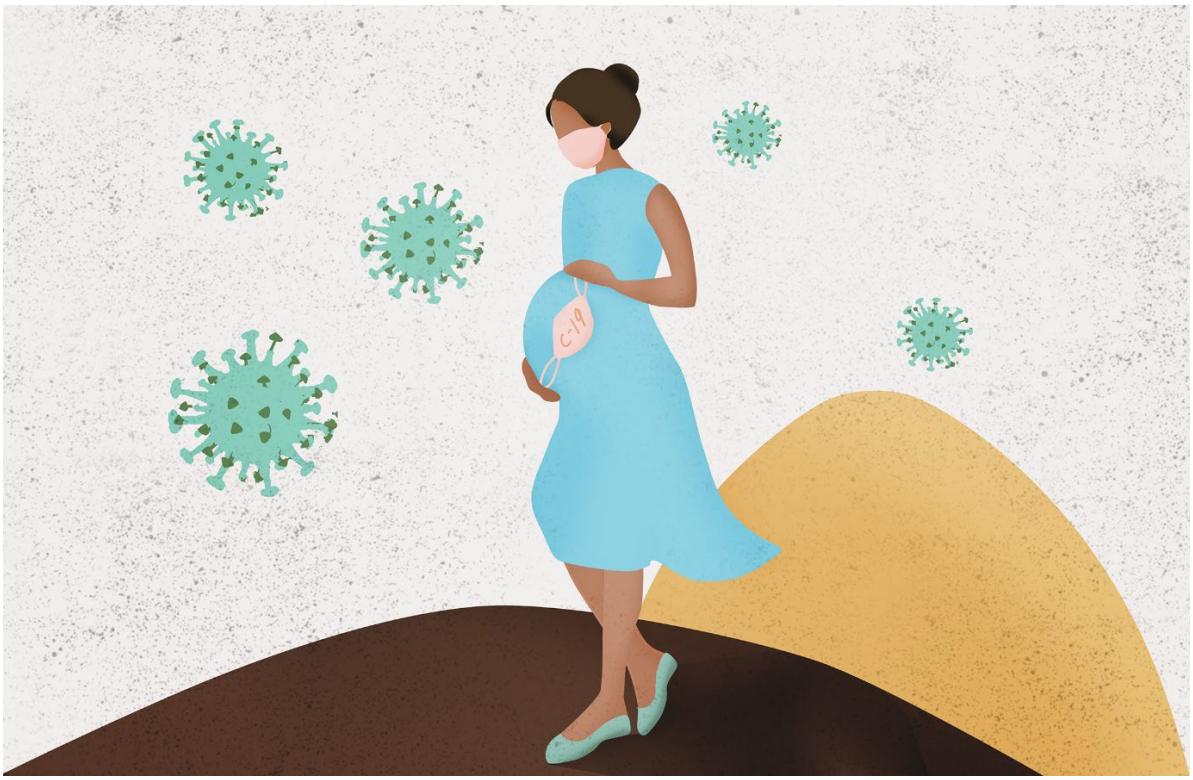
References

1. Hemingway, H.; Croft, P.; Perel, P.; Hayden, J.A.; Abrams, K.; Timmis, A.; Briggs, A.; Udumyan, R.; Moons, K.G.M.; Steyerberg, E.W.; et al. Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. *BMJ* **2013**, *346*, e5595. [CrossRef]
2. Hingorani, A.D.; van der Windt, D.A.; Riley, R.D.; Abrams, K.; Moons, K.G.M.; Steyerberg, E.W.; Schroter, S.; Sauerbrei, W.; Altman, D.G.; Hemingway, H.; et al. Prognosis research strategy (PROGRESS) 4: Stratified medicine research. *BMJ* **2013**, *346*, e5793. [CrossRef]
3. Riley, R.D.; Hayden, J.A.; Steyerberg, E.W.; Moons, K.G.M.; Abrams, K.; Kyzas, P.A.; Malats, N.; Briggs, A.; Schroter, S.; Altman, D.G.; et al. Prognosis research strategy (PROGRESS) 2: Prognostic factor research. *PLoS Med.* **2013**, *10*, e1001380. [CrossRef]
4. Steyerberg, E.W.; Moons, K.G.M.; van der Windt, D.A.; Hayden, J.A.; Perel, P.; Schroter, S.; Riley, R.D.; Hemingway, H.; Altman, D.G. Prognosis Research Strategy (PROGRESS) 3: Prognostic model research. *PLoS Med.* **2013**, *10*, e1001381. [CrossRef] [PubMed]
5. Goetz, L.H.; Schork, N.J. Personalized medicine: Motivation, challenges, and progress. *Fertil. Steril.* **2018**, *109*, 952–963. [CrossRef]

6. Pene, F.; Courtine, E.; Cariou, A.; Mira, J.P. Toward theragnostics. *Crit. Care Med.* **2009**, *37* (Suppl. 1), S50–S58. [CrossRef] [PubMed]
7. Lanfear, D.E.; McLeod, H.L. Pharmacogenetics: Using DNA to optimize drug therapy. *Am. Fam. Physician* **2007**, *76*, 1179–1182.
8. Khramtsova, E.A.; Davis, L.K.; Stranger, B.E. The role of sex in the genomics of human complex traits. *Nat. Rev. Genet.* **2019**, *20*, 173–190. [CrossRef] [PubMed]
9. Riley, R.D.; Moons, K.G.M.; Snell, K.I.E.; Ensor, J.; Hooft, L.; Altman, D.G.; Hayden, J.; Collins, G.S.; Debray, T.P.A. A guide to systematic review and meta-analysis of prognostic factor studies. *BMJ* **2019**, *364*, k4597. [CrossRef]
10. Peres, K.C.; Riera, R.; Martimbianco, A.L.C.; Ward, L.S.; Cunha, L.L. Body mass index and prognosis of COVID-19 infection. A systematic review. *Front. Endocrinol.* **2020**, *11*, 562. [CrossRef] [PubMed]
11. Pitz, M.W.; Musto, G.; Navaratnam, S. Sex as an independent prognostic factor in a population-based non-small cell lung cancer cohort. *Can. Respir. J.* **2013**, *20*, 30–34. [CrossRef]
12. Hsu, L.H.; Chu, N.M.; Liu, C.C.; Tsai, S.Y.C.; You, D.L.; Ko, J.S.; Lu, M.C.; Feng, C.A. Sex-associated differences in non-small cell lung cancer in the new era: Is gender an independent prognostic factor? *Lung Cancer* **2009**, *66*, 262–267. [CrossRef]
13. Maguire, A.; Porta, M.; Sanz-Anquela, J.M.; Ruano, I.; Malats, N.; Piñol, J.L. Sex as a prognostic factor in gastric cancer. *Eur. J. Cancer* **1996**, *32*, 1303–1309. [CrossRef]
14. Ceniceros Rozalén, I.; Gastaldo Simeón, R.; Cabadés O'Callaghan, A.; Cebrián Doménech, J. The female sex is an independent prognostic factor for mortality in the acute phase of myocardial infarction. *Med. Clin.* **1997**, *109*, 171–174.
15. Trappolini, M.; Chillotti, F.M.; Rinaldi, R.; Trappolini, F.; Coclite, D.; Napoletano, A.M.; Matteoli, S. Sex differences in incidence of mortality after acute myocardial infarction. *Ital. Heart J. Suppl.* **2002**, *3*, 759–766. [PubMed]
16. Mauvais-Jarvis, F.; Merz, N.B.; Barnes, P.J.; Brinton, R.D.; Carrero, J.J.; DeMeo, D.L.; de Vries, G.J.; Epperson, C.N.; Govindan, R.; Klein, S.L.; et al. Sex and gender: Modifiers of health, disease, and medicine. *Lancet* **2020**, *396*, 565–582. [CrossRef]
17. Clayton, J.A.; Tannenbaum, C. Reporting sex, gender, or both in clinical research? *JAMA* **2016**, *316*, 1863–1864. [CrossRef]
18. Higgins, J.P.T.; Thomas, J.; Chandler, J.; Cumpston, M.; Li, T.; Page, M.J.; Welch, V.A. (Eds.) Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (Updated February 2021). Cochrane. 2021. Available online: www.training.cochrane.org/handbook (accessed on 2 March 2021).
19. Lopez-Alcalde, J.; Martín, A.A.; Stallings, E.; Muriel, A.; Félix, B.F.; Solà, I.; del Campo, R.; Alonso, M.P.; Gordo, F.; Fidalgo, P.; et al. Evaluation of the role of sex as a prognostic factor in critically ill adults with sepsis: Systematic review protocol. *BMJ Open* **2020**, *10*, e035927. [CrossRef]
20. López-Alcalde, J.; Stallings, E.C.; Zamora, J.; Muriel, A.; van Doorn, S.; Alvarez-Diaz, N.; Fernandez-Felix, B.M.; Quezada, L.; Carlos, A.; Perez, R.; et al. Sex as a prognostic factor for mortality in adults with acute symptomatic pulmonary embolism. *Cochrane Database Syst. Rev.* **2021**, *1*, 1–29.
21. Arnegard, M.E.; Whitten, L.A.; Hunter, C.; Clayton, J.A. Sex as a Biological Variable: A 5-Year Progress Report and Call to Action. *J Womens Health (Larchmt).* **2020**, *29*, 858–864. [CrossRef]
22. Heidari, S.; Thomas, F.B.; Paola, D.C.; Sera, T.; Mirjam, C. Sex and gender equity in research: Rationale for the SAGER guidelines and recommended use. *Res. Integr. Peer Rev.* **2016**, *1*, 2. [CrossRef] [PubMed]
23. Gittleman, H.; Ostrom, Q.T.; Stetson, L.C.; Waite, K.; Hodges, T.R.; Wright, C.H.; Wright, J.; Rubin, J.B.; Berens, M.E.; Lathia, J.; et al. Sex is an important prognostic factor for glioblastoma but not for nonglioblastoma. *Neurooncol. Pract.* **2019**, *6*, 451–462. [CrossRef] [PubMed]
24. Wang, F.; Shu, X.; Meszoely, I.; Pal, T.; Mayer, I.A.; Yu, Z.; Zheng, W.; Bailey, C.E.; Shu, X. Overall mortality after diagnosis of breast cancer in men vs women. *JAMA Oncol.* **2019**, *5*, 1589–1596. [CrossRef] [PubMed]
25. Alswat, K.A. Gender disparities in osteoporosis. *J. Clin. Med. Res.* **2017**, *9*, 382–387. [CrossRef]
26. Picchioni, M.M.; Murray, R.M. Schizophrenia. *BMJ* **2007**, *335*, 91–95. [CrossRef]
27. Neumann, J.T.; Goßling, A.; Sörensen, N.A.; Blankenberg, S.; Magnussen, C.; Westermann, D. Sex-specific outcomes in patients with acute coronary syndrome. *J. Clin. Med.* **2020**, *9*, 2124. [CrossRef]
28. Mehta, L.S.; Beckie, T.M.; DeVon, H.A.; Grines, C.L.; Krumholz, H.M.; Johnson, M.N.; Lindley, K.J.; Vaccarino, V.; Wang, T.Y.; Watson, K.E.; et al. Acute myocardial infarction in women: A scientific statement from the american heart association. *Circulation* **2016**, *133*, 916–947. [CrossRef]
29. Liakos, M.; Parikh, P.B. Gender disparities in presentation, management, and outcomes of acute myocardial infarction. *Curr. Cardiol. Rep.* **2018**, *20*, 64. [CrossRef] [PubMed]
30. Aldin, A.; Umlauff, L.; Estcourt, L.J.; Collins, G.; Moons, K.G.; Engert, A.; Kobe, C.; von Tresckow, B.; Haque, M.; Foroutan, F.; et al. Interim PET-results for prognosis in adults with Hodgkin lymphoma: A systematic review and meta-analysis of prognostic factor studies. *Cochrane Database Syst. Rev.* **2019**, *9*, Cd012643. [CrossRef]
31. Hayden, J.A.; Wilson, M.N.; Riley, R.D.; Iles, R.; Pincus, T.; Ogilvie, R. Individual recovery expectations and prognosis of outcomes in non-specific low back pain: Prognostic factor review. *Cochrane Database Syst. Rev.* **2019**, *2019*, CD011284. [CrossRef]
32. Westby, M.J.; Dumville, J.C.; Stubbs, N.; Norman, G.; Wong, J.K.; Cullum, N.; Riley, R.D. Protease activity as a prognostic factor for wound healing in venous leg ulcers. *Cochrane Database Syst. Rev.* **2018**, *9*, Cd012841. [CrossRef] [PubMed]
33. Bougouin, W.; Mustafic, H.; Marijon, E.; Murad, M.H.; Dumas, F.; Barbouttis, A.; Jabre, P.; Beganton, F.; Empana, J.-P.; Celermajer, D.S.; et al. Gender and survival after sudden cardiac arrest: A systematic review and meta-analysis. *Resuscitation* **2015**, *94*, 55–60. [CrossRef]

34. Kim, C.; Redberg, R.F.; Pavlic, T.; Eagle, K.A. A systematic review of gender differences in mortality after coronary artery bypass graft surgery and percutaneous coronary interventions. *Clin. Cardiol.* **2007**, *30*, 491–495. [CrossRef]
35. Giuliano, L.; Mainieri, G.; Aguglia, U.; Bilo, L.; Durante, V.; Ermio, C.; Galimberti, C.A.; Neve, A.; Monti, G.; Ranzato, F.; et al. Long-term prognosis of juvenile myoclonic epilepsy: A systematic review searching for sex differences. *Seizure* **2021**, *86*, 41–48. [CrossRef]
36. Moons, K.G.; de Groot, J.A.H.; Bouwmeester, W.; Vergouwe, Y.; Mallett, S.; Altman, D.G.; Reitsma, J.B.; Collins, G.S. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: The CHARMS checklist. *PLoS Med.* **2014**, *11*, e1001744. [CrossRef] [PubMed]
37. Moons, K.G.M.; Wolff, R.F.; Riley, R.D.; Whiting, P.F.; Westwood, M.; Collins, G.S.; Reitsma, J.B.; Kleijnen, J.; Mallett, S. PROBAST: A tool to assess risk of bias and applicability of prediction model studies: Explanation and elaboration. *Ann. Intern. Med.* **2019**, *170*, W1–W33. [CrossRef] [PubMed]
38. Higgins, J.P.T.; Altman, D.G.; Sterne, J.A.C. Chapter 8: Assessing risk of bias in included studies. In *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*; (Updated March 2011); Higgins, J.P.T., Green, S., Eds.; The Cochrane Collaboration; 2011; Available online: www.cochrane-handbook.org (accessed on 2 March 2021).
39. Higgins, J.P.T.; Altman, D.G.; Gøtzsche, P.C.; Jüni, P.; Moher, D.; Oxman, A.D.; Savović, J.; Schulz, K.F.; Weeks, L.; Sterne, J.A.C.; et al. The cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ* **2011**, *343*, d5928. [CrossRef]
40. Sterne, J.A.C.; Savović, J.; Page, M.J.; Elbers, R.G.; Blencowe, N.S.; Boutron, I.; Cates, C.J.; Cheng, H.Y.; Corbett, M.S.; Eldridge, S.M.; et al. RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ* **2019**, *366*, l4898. [CrossRef]
41. Hayden, J.A.; Cote, P.; Bombardier, C. Evaluation of the quality of prognosis studies in systematic reviews. *Ann. Intern. Med.* **2006**, *144*, 427–437. [CrossRef]
42. Hayden, J.A.; van der Windt, D.A.; Cartwright, J.L.; Côté, P. Assessing bias in studies of prognostic factors. *Ann. Intern. Med.* **2013**, *158*, 280–286. [CrossRef]
43. Mahmood, K.; Eldeirawi, K.; Wahidi, M.M. Association of gender with outcomes in critically ill patients. *Crit. Care* **2012**, *16*, R92. [CrossRef] [PubMed]

6. 4th Article: Clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy: living systematic review and meta-analysis



In March of 2020 most of the world went into lockdown to control the transmission of coronavirus (Covid-19) which was spreading rapidly around the world. By April 2020, projects were paused due to the urgency of the pandemic, allowing the PregCOV19 project (https://www.birmingham.ac.uk/research/who-collaborating-centre/pregcov/index.aspx) priority. In the PregCOV-19 project, we are conducting a series of living systematic reviews (LSR) involving pregnant and postnatal women at risk, suspected, and diagnosed to have COVID-19 according to recommended methods. The PregCov19 team is continuously updating the findings, by incorporating relevant new evidence as it becomes available. In September 2020 we published the LSR on clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy including 77 studies in total. In March 2021 the BMJ published our first update of the LSR which now includes 192 studies. We will continue to publish our LSR updates every 5-6 months as new evidence becomes available for inclusion. I have included the original BMJ manuscript as an appendix and the updated manuscript within the thesis body for comparison purposes.

Clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy: living systematic review and meta-analysis

John Allotey,[1,2] Elena Stallings,[3,4] Mercedes Bonet,[5] Magnus Yap,[6] Shaunak Chatterjee,[6] Tania Kew,[6] Luke Debenham,[6] Anna Clavé Llavall,[6] Anushka Dixit,[6] Dengyi Zhou,[6] Rishab Balaji,[6] Siang Ing Lee,[1] Xiu Qiu,[7,8,9] Mingyang Yuan,[1,7] Dyuti Coomar,[1] Jameela Sheikh,[6] Heidi Lawson,[6] Kehkashan Ansari,[2] Madelon van Wely,[10] Elizabeth van Leeuwen,[11] Elena Kostova,[10] Heinke Kunst,[12,13] Asma Khalil,[14] Simon Tiberi,[12,13] Vanessa Brizuela,[5] Nathalie Broutet,[5] Edna Kara,[3] Caron Rahn Kim,[5] Anna Thorson,[5] Ramón Escuriet,[15] Olufemi T Oladapo,[5] Lynne Mofenson,[16] Javier Zamora,[2,3,4] Shakila Thangaratinam,[2,18] on behalf of the PregCOV-19 Living Systematic Review Consortium

For numbered affiliations see end of the article.

Correspondence to: S Thangaratinam s.thangaratinam.1@bham.ac.uk (or @thangaratinam on Twitter: ORCID 0000-0002-4254-460X)

Additional material is published online only. To view please visit the journal online.

cite this as: *BMJ* 2020;370:m3320
http://dx.doi.org/10.1136/bmj.m3320

Originally accepted: 23 August 2020

Final version accepted: 2 February 2021

AbstrAct

Objective

To determine the clinical manifestations, risk factors, and maternal and perinatal outcomes in pregnant and recently pregnant women with suspected or confirmed coronavirus disease 2019 (covid-19).

Design

Living systematic review and meta-analysis.

Data sOurces

Medline, Embase, Cochrane database, WHO COVID-19 database, China National Knowledge Infrastructure (CNKI), and Wanfang databases from 1 December 2019 to 6 October 2020, along with preprint servers, social media, and reference lists.

stuDy selectiOn

Cohort studies reporting the rates, clinical manifestations (symptoms, laboratory and radiological findings), risk factors, and maternal and perinatal outcomes in pregnant and recently pregnant women with suspected or confirmed covid-19.

Data extractiOn

At least two researchers independently extracted the data and assessed study quality. Random effects meta-analysis was performed, with estimates pooled as odds ratios and proportions with 95% confidence intervals. All analyses will be updated regularly.

results

192 studies were included. Overall, 10% (95% confidence interval 7% to 12%; 73 studies, 67 271 women) of pregnant and recently pregnant women attending or admitted to hospital for any reason were diagnosed as having suspected or confirmed covid-19. The most common clinical manifestations of covid-19 in pregnancy were fever (40%) and cough (41%). Compared with non-pregnant women of reproductive age, pregnant and recently pregnant women with covid-19 were less likely to have symptoms (odds ratio 0.28, 95% confidence interval 0.13 to 0.62; I^2=42.9%) or report symptoms of fever (0.49, 0.38 to 0.63; I^2=40.8%), dyspnoea (0.76, 0.67 to 0.85; I^2=4.4%) and myalgia (0.53, 0.36 to 0.78; I^2=59.4%). The odds of admission to an intensive care unit (odds ratio 2.13, 1.53 to 2.95; I^2=71.2%), invasive ventilation (2.59, 2.28 to 2.94; I^2=0%) and need for extra corporeal membrane oxygenation (2.02, 1.22 to 3.34; I^2=0%) were higher in pregnant and recently pregnant than non-pregnant reproductive aged women. Overall, 339 pregnant women (0.02%, 59 studies, 41 664 women) with confirmed covid-19 died from any cause. Increased maternal age (odds ratio 1.83, 1.27 to 2.63; I^2=43.4%), high body mass index (2.37, 1.83 to 3.07; I^2=0%), any pre-existing maternal comorbidity (1.81, 1.49 to 2.20; I^2=0%), chronic hypertension (2.0, 1.14 to 3.48; I^2=0%), pre-existing diabetes (2.12, 1.62 to 2.78; I^2=0%), and pre-eclampsia (4.21, 1.27 to 14.0; I^2=0%) were associated with severe covid-19 in pregnancy. In pregnant women with covid-19, increased maternal age, high body mass index, non-white ethnicity, any pre-existing maternal comorbidity including chronic hypertension and diabetes, and pre-eclampsia were associated with serious complications such as admission to an intensive care unit, invasive

WhAt is AlreAdy knoWn on this topic

Pregnant women are considered to be a high risk group for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, and the potential adverse effects of the virus on maternal and perinatal outcomes are of concern

In non-pregnant populations admitted to hospital with coronavirus disease 2019 (covid-19) the most common symptoms are fever, cough, and dyspnoea, reported in more than two thirds of individuals

Advancing age, high body mass index, non-white ethnicity, and pre-existing comorbidities are risk factors for severe covid-19 in the general population

WhAt this study Adds

Pregnant and recently pregnant women with covid-19 diagnosed in hospital are less likely to have or manifest symptoms of fever, dyspnoea, and myalgia than non-pregnant women of reproductive age

Pregnant and recently pregnant women are at increased risk of admission to an intensive care unit, receiving invasive ventilation and extra corporeal membrane oxygenation treatment, compared with non-pregnant women of reproductive age

Risk factors for severe covid-19 in pregnancy include increasing maternal age, high body mass index, non-white ethnicity, pre-existing comorbidities, and pregnancy specific disorders such as gestational diabetes and pre-eclampsia

Pregnant women with covid-19 are more likely to experience preterm birth and their neonates are more likely to be admitted to a neonatal unit

ventilation and maternal death. Compared to pregnant women without covid-19, those with the disease had increased odds of maternal death (odds ratio 2.85, 1.08 to 7.52; I^2=0%), of needing admission to the intensive care unit (18.58, 7.53 to 45.82; I^2=0%), and of preterm birth (1.47, 1.14 to 1.91; I^2=18.6%). The odds of admission to the neonatal intensive care unit (4.89, 1.87 to 12.81, I^2=96.2%) were higher in babies born to mothers with covid-19 versus those without covid-19.

cOnclusiOn

Pregnant and recently pregnant women with covid-19 attending or admitted to the hospitals for any reason are less likely to manifest symptoms such as fever, dyspnoea, and myalgia, and are more likely to be admitted to the intensive care unit or needing invasive ventilation than non-pregnant women of reproductive age. Pre-existing comorbidities, non-white ethnicity, chronic hypertension, pre-existing diabetes, high maternal age, and high body mass index are risk factors for severe covid-19 in pregnancy. Pregnant women with covid-19 versus without covid-19 are more likely to deliver preterm and could have an increased risk of maternal death and of being admitted to the intensive care unit. Their babies are more likely to be admitted to the neonatal unit.

systematic review registratiOn

PROSPERO CRD42020178076.

reaDers' nOte

This article is a living systematic review that will be updated to reflect emerging evidence. Updates may occur for up to two years from the date of original publication. This version is update 1 of the original article published on 1 September 2020 (*BMJ* 2020;370:m3320), and previous updates can be found as data supplements (https://www.bmj.com/content/370/bmj.m3320/related#datasupp). When citing this paper please consider adding the update number and date of access for clarity.

introduction

Since the first report (December 2019) of the novel coronavirus disease 2019 (covid-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the number of confirmed cases and associated mortality and morbidity have increased rapidly.[1 2] Pregnant women are considered a high risk group because of concerns about the effect of covid-19 on them during and after pregnancy, and on their neonates.[3] Quantification of the rates of covid-19, its risk factors, clinical manifestations, and outcomes is key to planning clinical maternal care and management in an evolving pandemic scenario.[4]

Publications on covid-19 in pregnancy have risen steeply through individual case reports, case series, observational studies, and systematic reviews. Since the publication of our first version of the living systematic review on covid-19 in pregnancy,[5] over 150 reviews have been published in this area,[6-11] with many more registered in PROSPERO.[9 12] Early reviews mostly included case reports and case series that were often inappropriately meta-analysed providing biased estimates.[13] Subsequent reviews differed little from each other, often including similar primary studies, many with duplicate data. These reviews became quickly outdated as new evidence emerged. Moreover, the sampling frames in primary studies have varied, ranging from universal SARS-CoV-2 testing for all pregnant women admitted to hospital[14 15] to symptom based testing.[16 17] Testing strategies have also differed within and between countries, with diagnosis in many early studies based on epidemiological risk assessment and clinical features without confirmed SARS-CoV-2 infection, which need to be considered in the analysis.[18] Limitations in the external and internal validity of studies make it challenging for guideline developers and policy makers to make evidence based recommendations for the management of pregnant and recently pregnant women with covid-19.

We started this living systematic review in April 2020 to determine the clinical manifestations of covid-19 in pregnant and recently pregnant women, identify the risk factors for complications, and quantify maternal and perinatal outcomes. The systematic review is being updated on a regular basis.

Methods

Our systematic review is based on a prospectively registered protocol (PROSPERO CRD42020178076; registered 22 April 2020)[19] to evaluate a series of research questions on covid-19 during and after pregnancy. We report our findings on the rates, clinical manifestations, risk factors, and maternal and perinatal outcomes in women with covid-19 in line with the preferred reporting items for systematic reviews and meta-analyses (PRISMA) recommendations (see appendix 1). As more relevant data become available, we shall address the research questions in our published protocol.[20] Each cycle of our living systematic review involves weekly search updates (rounds), with analysis performed every 2-4 months for reporting through a dedicated website, with early analysis if new definitive evidence emerges. We are regularly reviewing the planned frequency of updates.

literature search

For the first publication of the review, we performed a systematic search of major databases: Medline, Embase, Cochrane database, WHO (World Health Organization) COVID-19 database, China National Knowledge Infrastructure (CNKI), and Wanfang databases from 1 December 2019 to 26 June 2020 for relevant studies on covid-19 in pregnant and recently pregnant women.[5] For this first update of the review, we searched databases up to 6 October 2020. To identify potential studies, we coordinated our search efforts with the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre), the WHO Library, and the Cochrane Gynaecology and Fertility group. Additional searches were conducted of preprint servers, blogs, websites that serve as repositories

for covid-19 studies, social media, guidelines, and reference lists of included studies and unpublished data. We also searched the Living Overview of the Evidence (LOVE) platform from June 2020.[21] We contacted established groups that were coordinating or conducting surveillance and studies in pregnant women with covid-19, such as the WHO Maternal, Newborn, Child and Adolescent health (MNCAH) covid-19 research network, the International Network of Obstetric Survey Systems (INOSS), the United States Centers for Disease Control and Prevention (CDC), and the European Centre for Disease Prevention and Control for information on published and upcoming data. No language restrictions were applied. Appendix 2 provides details of the search strategies and databases searched.

study selection

Two reviewers independently selected studies using a two stage process: they first screened the titles and abstracts of studies and then assessed the full text of the selected studies in detail for eligibility. A total of 10 reviewers contributed to study selection. Disagreements were resolved through discussion with a third reviewer (ST or JA). We excluded studies if the duplicated data for all outcomes of interest were published elsewhere, as reported by the study authors, or when the characteristics of the mother or neonate matched the setting, characteristics, and duration of another study from the same geographical location. When we suspected an overlap of data between studies, the study that provided comparative data was included. If there was an overlap of data or suspicion of duplicates of participants in studies between the previous and current versions of the living systematic review, we included studies based on their study design (prioritising comparative cohorts), and sample size (larger study prioritised). When there was uncertainty about duplicate data, we contacted the authors of primary studies.

We defined women as having confirmed covid-19 if they had laboratory confirmation of SARS-CoV-2 infection irrespective of clinical signs and symptoms.[22] Women with a diagnosis based only on clinical or radiological findings were defined as having suspected covid-19. The recently pregnant group comprised women in the postpartum and post-abortion period. We included studies that compared covid-19 rates, clinical manifestations (symptoms, laboratory and radiological results), risk factors, and associated mortality and morbidity between pregnant and recently pregnant and non-pregnant women of reproductive age, and those that compared maternal and perinatal outcomes in pregnant women with and without covid-19. In studies comparing maternal and perinatal outcomes of pregnant women with covid-19 to those without, we classified the comparative controls as being historical if the cohort of pregnant women without covid-19 were pregnant before December 2019. Studies on non-comparative cohorts with a minimum of 10 participants were included if they reported on the rates and clinical manifestations of covid-19 and relevant outcomes in pregnant and recently pregnant women. We defined cohort studies as those that sampled participants on the basis of exposure, followed-up participants over time, and ascertained the outcomes.[23] The PROSPERO protocol provides a full list of the risk factors, clinical features, and outcomes evaluated.[19]

The sampling frames for detecting covid-19 included universal screening and testing, when all women were assessed for covid-19 using reverse transcriptase polymerase chain reaction (RT-PCR) for SARS-CoV-2 or chest computed tomography; risk based testing on the basis of epidemiological history and clinical manifestations by National Health Commission of China (NHCC) guidelines[18]; and symptom based when testing was performed on women with symptoms and those with a history of contact with affected individuals. We defined the population as being selected when only specific groups of women were included, such as those undergoing caesarean section or in the third trimester. We categorised studies as a high risk group if only women with any pre-existing medical or obstetric risk factors were included, low risk if women did not have any risk factors, and any risk if all women were included.

study quality assessment and data extraction

The quality of the comparative cohort studies was assessed for selection, comparability, and outcome ascertainment bias using the Newcastle Ottawa scale.[24] Studies achieving four stars for selection, two for comparability, and three for ascertainment of the outcome were considered to have a low risk of bias. Studies achieving two or three stars for selection, one for comparability, and two for outcome ascertainment were considered to have a medium risk of bias, and any study achieving one star for selection or outcome ascertainment, or zero for any of the three domains, was regarded as having a high risk of bias. We assessed the quality of studies reporting on the prevalence of clinical manifestations or outcomes for internal and external validity using an existing tool.[25] The following were considered as low risk of bias for external validity: representative of national population for relevant variables (population), representative of target population (sampling frame), random selection (selection bias), and more than 75% response rate in individuals with and without the outcome (non-response bias).[25] Two independent reviewers extracted data using a pre-piloted form.

statistical analysis

We pooled the comparative dichotomous data using random effects meta-analysis and summarised the findings as odds ratios with 95% confidence intervals. To combine comparative continuous data with dichotomous data we transformed standardised mean differences to logarithm odds ratios, assuming a normal underlying distribution.[26] We pooled the dichotomous non-comparative data for rates of clinical manifestations and maternal and perinatal outcomes

as proportions with 95% confidence intervals using Dersimonian and Laird random effects meta-analysis after transforming data using Freeman-Tukey double arcsin transformation. Heterogeneity was reported as I^2 statistics. We undertook subgroup analysis by country status (high income *v* low and middle income), sampling frame (universal, risk based, and symptom based testing, including not reported), and risk status of women in the studies (high, low, any). Sensitivity analysis was performed by restricting the analysis to women with confirmed covid-19, study quality (high, low), and population (unselected, selected). All analyses were done with Stata (version 16).

Patient and public involvement

The study was supported by Katie's Team, a dedicated patients and public involvement group in Women's Health. The team was involved in the conduct, interpretation, and reporting of this living systematic review through participation in virtual meetings.

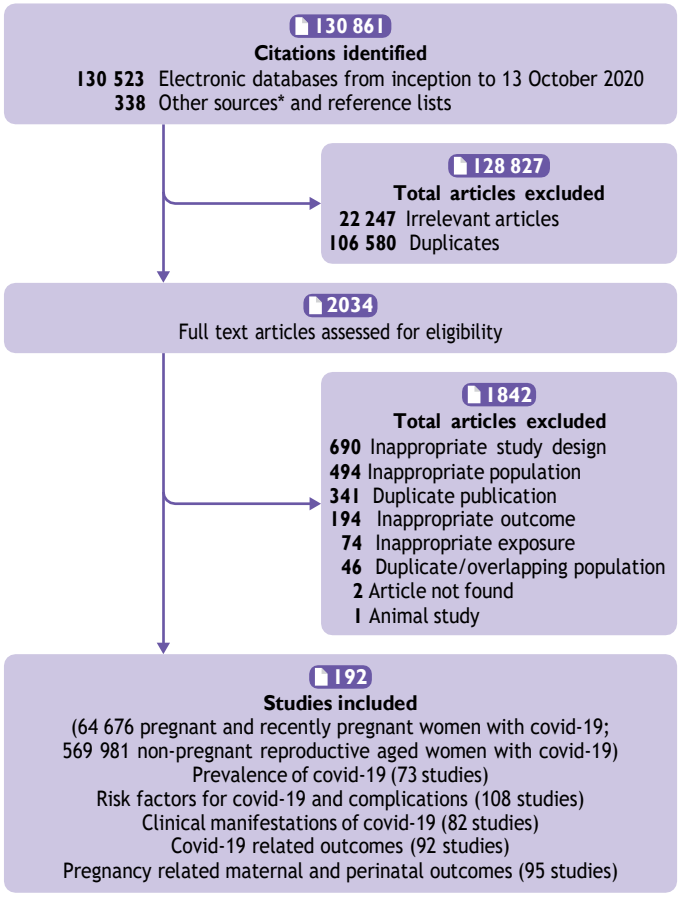


Fig 1 | study selection process. *twitter, national reports, blog by j thornton, Obg Project, cOviD-19 and Pregnancy cases, www.obgproject.com/2020/04/07/covid-19-research-watch-with-dr-jim-thornton/; ePPi-centre, cOviD-19: a living systematic map of evidence, http://eppi.ioe.ac.uk/cms/Projects/DepartmentofHealthandsocialcare/Publishedreviews/cOviD-19livingsystematicmapoftheevidence/tabid/3765/Default.aspx; norwegian institute of Public Health, niPH systematic and living map on cOviD-19 evidence, www.nornesk.no/forskningskart/niPH_mainmap.html; johns Hopkins university center for Humanitarian Health; cOviD-19, maternal and child Health, nutrition, http://hopkinshumanitarianhealth.org/empower/advocacy/covid-19/covid-19-children-and-nutrition/; researchgate, cOviD-19 research community, www.researchgate.net/community/cOviD-19; and living Overview of the evidence, coronavirus disease (cOviD-19), https://app.iloveevidence.com/loves/5e6fdb9669c00e4ac072701d?population=5d062d5fc80dd41e58ba8459

results

In the original review, 20 625 unique citations were identified after removing duplicates from 49 684 citations, with 77 cohort studies included in the review.[5] After removing duplicates from 130 861 citations, 24 281 unique citations were identified and 192 cohort studies (131 comparative, 61 non-comparative) were included in this update of the systematic review (fig 1). Two studies included in the original systematic review were excluded from the update because the information reported in those studies were reported in more recent and larger studies.[27 28]

characteristics of included studies

Of 192 studies, 58 (30%) were from the United States; 31 from China (16%); 17 from Italy; 15 from Spain; eight from Turkey; seven each from the United Kingdom and India; five each from Brazil, France, and Mexico; three each from Iran and Portugal; two each from Belgium, Denmark, the Netherlands, Peru, and Sweden; and one each from Bangladesh, Chile, Estonia, Israel, Japan, Germany, Ireland, Kuwait, Pakistan, Qatar, Romania, Russia, and Switzerland. Most studies tested respiratory samples using RT-PCR to confirm the presence of SARS-CoV-2 (97%, 187/192); five studies tested for SARS-CoV-2 antibodies to confirm the diagnosis of covid-19; 43 studies additionally diagnosed covid-19 based only on clinical suspicion. Fourteen studies (602 565 women) compared pregnant populations with non-pregnant populations,[29-42] and 47 studies (26 017 women) compared pregnant women with covid-19 versus pregnant women without covid-19.[43-89] Eighty two cohort studies reported on clinical manifestations (41 396 pregnant, 434 348 non-pregnant women), 92 studies reported on covid-19 related maternal outcomes (49 443 pregnant, 568 386 non-pregnant women), and 95 studies reported on pregnancy related maternal (54 943 women) and perinatal outcomes (9466 neonates) (see appendix 3). The sampling frames included universal testing (89 studies), risk based NHCC guidelines (25 studies), and symptom based (32 studies) strategies. Forty six studies did not report the sampling strategy.

Quality of included studies

Overall, 56% (73/131) of the comparative cohort studies evaluated using the Newcastle Ottawa scale had an overall low risk of bias (see appendix 4a). Most (93%, 122/131) had a low risk of bias for study selection and nine (7%) had a medium risk. The risk of bias for comparability of cohorts was low in 59 of the studies (45%), medium in 71 (54%), and high in one (1%). For outcome assessment of the cohorts, 47 (36%) studies had a low risk of bias, 82 (63%) a medium risk, and two (2%) a high risk. Quality assessment of the prevalence studies for external validity showed a low risk of bias for representativeness in 15% (28/192) of

the studies, sampling in 30% (57/192), selection in 82% (157/192), and non-response in 99% (191/192). For internal validity, there was low risk of bias for data collection in 96% (184/192) of the studies, case definition in 56% (108/192), measurement in 98% (189/192), differential verification in 95% (182/192), adequate follow-up in 35% (67/192), and appropriate numerator and denominator in 92% (177/192) (see appendix 4b).

rates of covid-19 in pregnant and recently pregnant women

The overall rate of covid-19 diagnosis in pregnant and recently pregnant women attending or admitted to hospital for any reason was 10% (95% confidence interval 7% to 12%; 73 studies, 67 271 women; fig 2 and fig 3). Rates varied by sampling strategy: of the women sampled by universal screening, 7% (5% to 8%; 60 studies, 57 144 women) were diagnosed as having covid-19 compared with 28% (15% to 43%; 11 studies, 2436 women) of women sampled on the basis of symptoms. Most studies with a prevalence rate for covid-19 greater than 15% were from the US, except for two studies from the UK, and one each from Mexico, Turkey, France, and Iran.[90-95] One in 20 asymptomatic women (4%, 3% to 7%; 26 studies) attending or admitted to hospital had a diagnosis of covid-19 (see appendix 5a). Three quarters (73%, 62% to 82%; 38 studies) of the 906 pregnant women with covid-19 in the universal screening population were asymptomatic (see appendix 5b). Non-white ethnicity was associated with a diagnosis of covid-19 in pregnancy (odds ratio 1.66, 95% confidence interval 1.01 to 2.72; 11 studies; 8691 women); none of the other maternal factors assessed were associated with a diagnosis of covid-19 in pregnant women (see appendix 6a).

clinical manifestations of covid-19 during pregnancy and after delivery

The most common symptoms reported by pregnant and recently pregnant women with suspected or confirmed covid-19 were fever (40%) and cough (41%); raised white cell count (26%), lymphopaenia (33%) and raised C reactive protein levels (49%) were the most common laboratory findings (fig 4). Compared with non-pregnant women of reproductive age with covid-19, pregnant and recently pregnant women with the disease were less likely to have symptoms (odds ratio 0.28, 95% confidence interval 0.13 to 0.62; 4 studies; 462 051 women), or manifest symptoms of fever (0.49, 0.38 to 0.63; 11 studies, 240 324 women), dyspnoea (0.76, 0.67 to 0.85; 11 studies; 240 324 women) and myalgia (0.53, 0.36 to 0.78; 8 studies, 240 105 women) (fig 5). Pregnant women with covid-19 had increased body mass index compared to non-pregnant women with the disease (1.98, 1.74 to 2.26; 2 studies, 461 897 women), and were more likely to have pre-existing diabetes (1.35, 1.24 to 1.46; 5 studies, 462 262 women) (see appendix 6b). Sensitivity analysis restricted to various sampling frames showed lower estimates of reported symptoms in the universal screening population and higher estimates of fever, cough, and dyspnoea in the symptom-based population (see appendix 7). The rates of clinical manifestations varied when the analysis was restricted to only women with RT-PCR confirmed covid-19, unselected populations, and women with any risk (see appendix 7).

Outcomes related to covid-19 in pregnant and recently pregnant women

Overall, 339 pregnant women (59 studies, 41 664 women) with confirmed covid-19 died from any cause (0.02%, 95% confidence interval 0.00% to 0.42%). Severe covid-19 infection as defined by the authors, was diagnosed in 10% (6% to 15%; 39 studies, 5621 women) of pregnant and recently pregnant women with suspected or confirmed covid-19; 4% (2% to 7%; 50 studies, 41 288 women) of pregnant women with covid-19 were admitted to an intensive care unit, 3% (1% to 5%; 31 studies, 42 026 women) required invasive ventilation, and 0.2% (0.0% to 0.7%; 13 studies, 33 521 women) required extracorporeal membrane oxygenation (ECMO) (fig 4). Appendix 8 provides the rates of complications by sampling strategy. Compared with non-pregnant women of reproductive age with covid-19, the odds of admission to the intensive care unit (odds ratio 2.13, 95% confidence interval 1.53 to 2.95; seven studies, 601 108 women) and need for invasive ventilation (2.59, 2.28 to 2.94; six studies, 601 044 women) and ECMO (2.02, 1.22 to 3.34; two studies, 461 936 women) were higher in pregnant and recently pregnant women (table 1).

Maternal risk factors associated with severe covid-19 were increasing age (odds ratio 1.83, 95% confidence interval 1.27 to 2.63; seven studies, 3561 women), high body mass index (2.37, 1.83 to 3.07; five studies, 3367 women), any pre-existing maternal comorbidity (1.81, 1.49 to 2.20; 3 studies; 2634 women), chronic hypertension (2.0, 1.14 to 3.48; two studies, 858 women), pre-eclampsia (4.21, 1.27 to 14.0; 4 studies; 274 women), and pre-existing diabetes (2.12, 1.62 to 2.78; 3 studies, 3333 women) (fig 6). Increasing maternal age (2.11, 1.69 to 2.63; 7 studies, 31 710 women), high body mass index (2.71, 1.10 to 6.63; 4 studies, 31 456 women), non-white ethnicity (1.66, 1.20 to 2.29; 4 studies, 31 543 women), pre-existing maternal comorbidity (1.70, 1.34 to 2.15; 5 studies, 31 512 women), chronic hypertension (4.72, 2.37 to 9.41; 5 studies, 31 433 women), pre-existing diabetes (4.67, 1.94 to 11.22; 6 studies, 31 473 women), and gestational diabetes (3.27, 1.55 to 6.89; 2 studies, 777 women), were associated with admission to an intensive care unit. Risk factors associated with maternal death and the need for invasive ventilation included: non-white ethnicity (1.61, 1.05 to 2.47; 3 studies, 31 469 women; 2.23, 1.25 to 3.97; 1 study, 669 women; respectively), and high body mass index (2.27, 1.20 to 4.31; 3 studies, 31 085 women; 6.61, 1.98 to 22.02; 2 studies, 485 women; respectively; table 2).

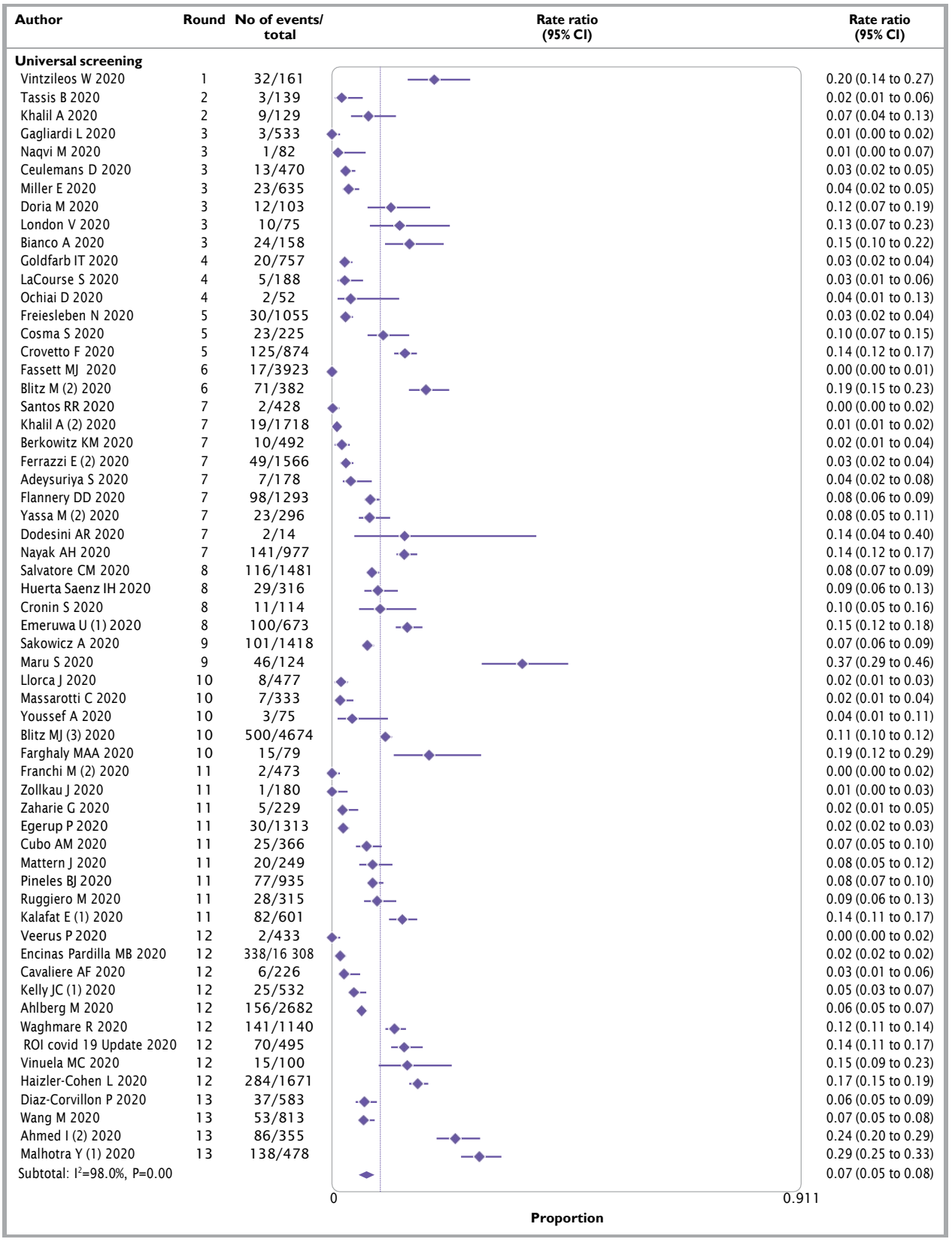


Fig 2 | Prevalence of severe acute respiratory syndrome coronavirus 2 in pregnant and recently pregnant women identified by universal screening. meta-analysis includes one study (liao 2020)[46] screened using national Health commission china criteria with no events. round number represents search strategy updates in the living systematic review. Overall estimate for sampling strategies can be found in figure 3

maternal and perinatal outcomes in pregnant and recently pregnant women with covid-19

In pregnant and recently pregnant women with covid-19 compared with pregnant and recently pregnant women without the disease, the odds of all cause mortality (odds ratio 2.85, 95% confidence interval 1.08 to 7.51; 8 studies, 4820 women), and admission to the intensive care unit (18.58, 95% confidence interval 7.53 to 45.82; 7 studies, 4990 women) were higher (table 1). In pregnant and recently pregnant women with covid-19, the overall rate of preterm birth was 17% (95% confidence interval 14% to 19%; 70 studies, 9369 women) and of spontaneous preterm birth was 6% (4% to 9%; 17 studies, 1629 women) (fig 4). Seventy two stillbirths (47 studies; 9020 offspring) and 41 neonatal deaths (51 studies; 8263 neonates) occurred among these women (fig 4). Compared to pregnant and recently pregnant women without the disease, pregnant women with covid-19 were at higher risk of any preterm birth (odds ratio 1.47, 95% confidence interval 1.14 to 1.91; 18 studies, 8549 women) and stillbirth (2.84, 95% confidence interval 1.25 to 6.45; 9 studies, 5794 women), although the overall number of stillbirth was small (only nine events in the covid-19 group).

Overall, 33% (95% confidence interval 24% to 43%; 41 studies, 3323 women) of neonates born to women with covid-19 were admitted to the neonatal intensive care unit (NICU) (fig 4), with a higher risk of NICU admission (odds ratio 4.89, 95% confidence interval 1.87 to 12.81; 10 studies, 5873 neonates) than neonates born to women without the disease. No differences were observed for other perinatal outcomes. Appendix 9 provides the rates of covid-19 related and pregnancy related outcomes for the individual studies.

discussion

Compared with the original version of our living systematic review, the findings in this update remain consistent for prevalence of covid-19, rates of clinical manifestations, and outcomes in pregnant and recently pregnant women. One in 10 pregnant or recently pregnant women who are attending or admitted to hospital for any reason were diagnosed as having suspected or confirmed covid-19, although the rates vary by sampling strategy. Pregnant and recently pregnant women were more likely to be asymptomatic than non-pregnant women of reproductive age, and showed covid-19 related symptoms of fever, dyspnoea, and myalgia less often than non-pregnant women with covid-19. Whereas testing for SARS-CoV-2 in non-pregnant women is based on symptoms or contact history, testing in pregnant women is usually done when they are in hospital for reasons that might not be related to covid-19. Pregnant or recently pregnant women with covid-19 seem to be at increased risk of requiring admission to an intensive care unit, invasive ventilation, and extra corporeal membrane oxygenation compared to non-pregnant, reproductive

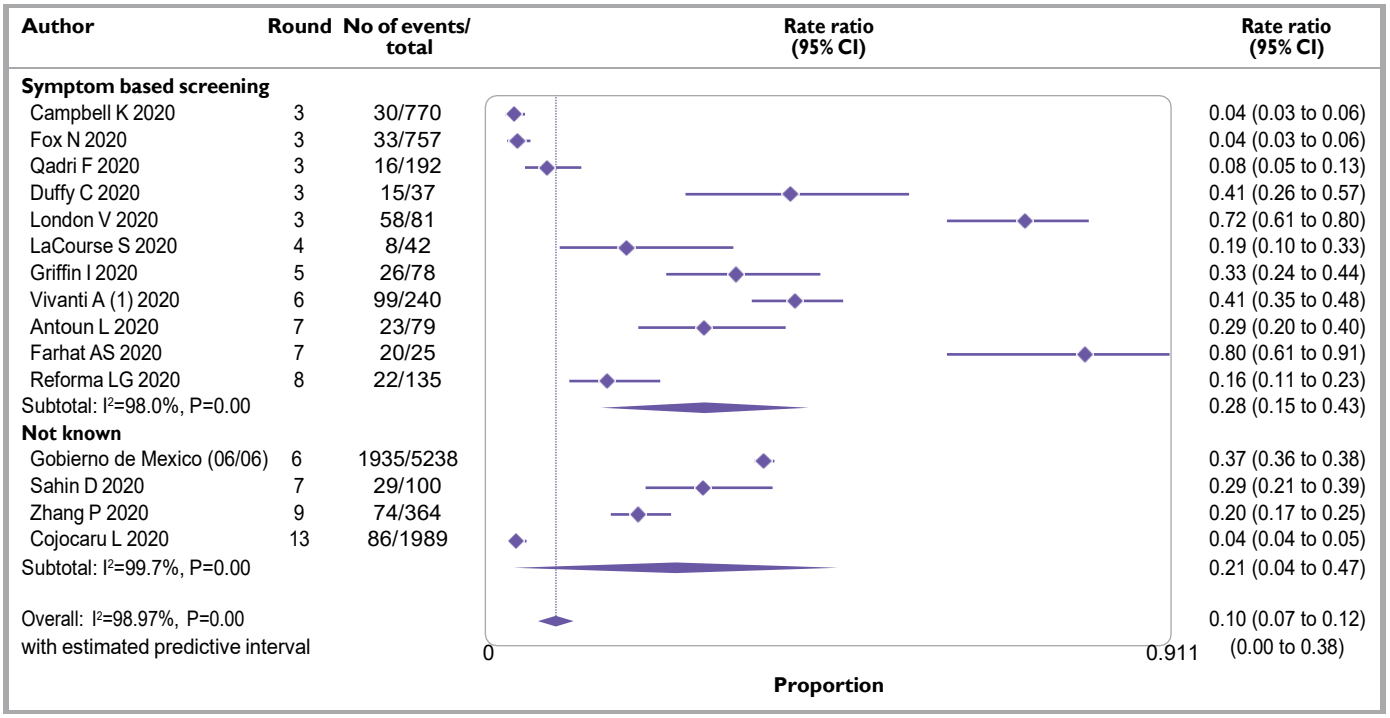


Fig 3 | Prevalence of severe acute respiratory syndrome coronavirus 2 in pregnant and recently pregnant women identified by symptom based screening and unknown sampling strategies. meta-analysis includes one study (liao 2020)[46] screened using national Health commission china criteria with no events. symptom based screening includes screening based on symptoms or history of contact with individuals with covid-19. round number represents search strategy updates in the living systematic review. Overall estimate for sampling strategies also includes prevalence data identified by universal screening, which are shown in figure 2

aged women with covid-19. Increased maternal age, high body mass index, non-white ethnicity, and pre-existing comorbidities are associated with severe disease. Compared to pregnant women without covid-19, pregnant women with covid-19 are at increased risk of death, admission to the intensive care unit, delivering preterm, and their babies being admitted to the neonatal unit. The overall rates of stillbirth and neonatal death are low in women with suspected or confirmed covid-19. Substantial heterogeneity was observed in the estimates for rates of clinical manifestations and outcomes, which varied by sampling frames, participant selection, and risk status of the participants.

This update of the living systematic review includes more than double the number of studies included in the original version, and five times more pregnant women with covid-19. In addition to an increase in precision of the estimates for previously identified risk factors (age, body mass index, and comorbidities such as diabetes and chronic hypertension) for serious complications in pregnant and recently pregnant women with covid-19,

Clinical manifestation	Studies	No of events/ total	Proportion (95% CI)	I^2 (%) (P value)	Range
Clinical manifestations					
Symptoms					
Fever	53	8033/39 429	0.40 (0.31 to 0.49)	99.2 (0.00)	(0.05-0.78)
Cough	53	10 379/39 641	0.41 (0.33 to 0.50)	99.1 (0.00)	(0.03-0.83)
Dyspnoea	42	5408/39 014	0.21 (0.15 to 0.28)	98.7 (0.00)	(0.00-0.62)
Myalgia	22	5196/34 663	0.19 (0.12 to 0.27)	98.4 (0.00)	(0.00-0.67)
Ageusia	10	83/776	0.14 (0.06 to 0.24)	89.6 (0.00)	(0.03-0.55)
Diarrhoea	29	2236/38 206	0.08 (0.06 to 0.10)	93.4 (0.00)	(0.00-0.53)
Laboratory findings					
Raised WCC	13	159/580	0.26 (0.14 to 0.40)	90.9 (0.00)	(0.00-0.65)
Lymphopaenia	27	659/1833	0.33 (0.25 to 0.41)	90.4 (0.00)	(0.00-0.90)
Thrombocytopaenia	13	91/1383	0.06 (0.02 to 0.10)	86.3 (0.00)	(0.00-0.35)
Abnormal LFTs	12	99/641	0.13 (0.06 to 0.21)	84.8 (0.00)	(0.00-0.36)
Raised PCT	5	60/261	0.21 (0.00 to 0.59)	96.6 (0.00)	(0.00-0.97)
Raised CRP	10	298/637	0.49 (0.36 to 0.62)	89.5 (0.00)	(0.10-0.71)
Radiological findings					
Ground glass appearance	14	338/569	0.69 (0.46 to 0.87)	96.3 (0.00)	(0.09-1.00)
Any CT abnormality	24	694/2120	0.64 (0.47 to 0.80)	98.2 (0.00)	(0.02-1.00)
Maternal and perinatal outcomes					
Clinical outcomes					
Covid-19 related outcomes					
All cause mortality	59	339/41 664	0.00 (0.00 to 0.00)	91.8 (0.00)	(0.00-0.08)
Admission to ICU	50	1373/41 288	0.04 (0.02 to 0.07)	97.5 (0.00)	(0.00-0.29)
Severe covid-19	39	633/5621	0.10 (0.06 to 0.15)	94.4 (0.00)	(0.00-1.00)
Invasive ventilation	31	668/42 026	0.03 (0.01 to 0.05)	97.5 (0.00)	(0.00-0.13)
ECMO	13	37/33 521	0.00 (0.00 to 0.00)	76.0 (0.00)	(0.00-0.01)
Oxygen/cannula	17	261/1522	0.22 (0.12 to 0.36)	96.2 (0.00)	(0.02-1.00)
ARDS	15	315/2348	0.07 (0.01 to 0.17)	97.8 (0.00)	(0.00-0.51)
Pneumonia	36	1257/7198	0.35 (0.26 to 0.45)	97.9 (0.00)	(0.00-1.00)
Cardiac/liver/renal failure	13	15/2046	0.00 (0.00 to 0.00)	28.3 (0.16)	(0.00-0.13)
Pregnancy related outcomes					
Preterm birth <37 weeks	70	1406/9396	0.17 (0.14 to 0.19)	79.6 (0.00)	(0.00-0.57)
Spontaneous preterm birth	17	104/1629	0.06 (0.04 to 0.09)	67.2 (0.00)	(0.00-0.31)
PPROM <37 weeks	18	58/993	0.05 (0.03 to 0.07)	29.7 (0.11)	(0.00-0.21)
Caesarean section	75	3760/9725	0.54 (0.49 to 0.58)	93.4 (0.00)	(0.00-1.00)
Vaginal delivery	74	5410/9708	0.46 (0.42 to 0.50)	91.8 (0.00)	(0.00-1.00)
Postpartum haemorrhage	15	91/908	0.08 (0.03 to 0.14)	86.7 (0.00)	(0.00-0.30)
Offspring outcomes					
Stillbirth	47	72/9020	0.00 (0.00 to 0.00)	31.2 (0.02)	(0.00-0.24)
Neonatal death	51	41/8263	0.00 (0.00 to 0.00)	33.0 (0.01)	(0.00-0.13)
Admission to neonatal unit	41	934/3323	0.33 (0.24 to 0.43)	96.8 (0.00)	(0.00-1.00)
Neonatal sepsis	6	9/499	0.01 (0.00 to 0.03)	22.6 (0.26)	(0.01-0.06)
Abnormal Apgar score	31	42/1479	0.01 (0.00 to 0.02)	34.8 (0.03)	(0.00-0.26)
Fetal distress	17	65/553	0.11 (0.08 to 0.15)	40.2 (0.04)	(0.04-0.50)

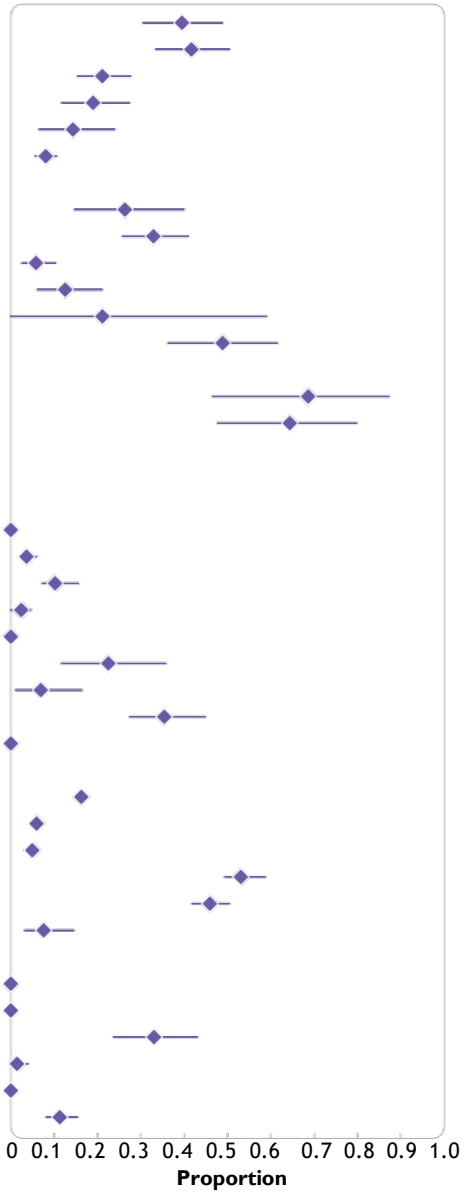


Fig 4 | rates of clinical manifestations of coronavirus disease 2019 (covid-19) in pregnant women and recently pregnant women with suspected or confirmed covid-19 and associated maternal and perinatal outcomes. ecmO=extracorporeal membrane oxygenation; arDs=acute respiratory distress syndrome; PPrOm=preterm premature rupture of membranes; wcc=white cell count; lFt=liver function test; Pct=procalcitonin; crP=c reactive protein; ct=computed tomography; icu=intensive care unit

in this update, we identified additional risk factors such as non-white ethnicity, and potential association with pregnancy specific conditions such as gestational diabetes and pre-eclampsia, and increased risk of adverse outcomes in pregnant women with covid-19 than without the disease.

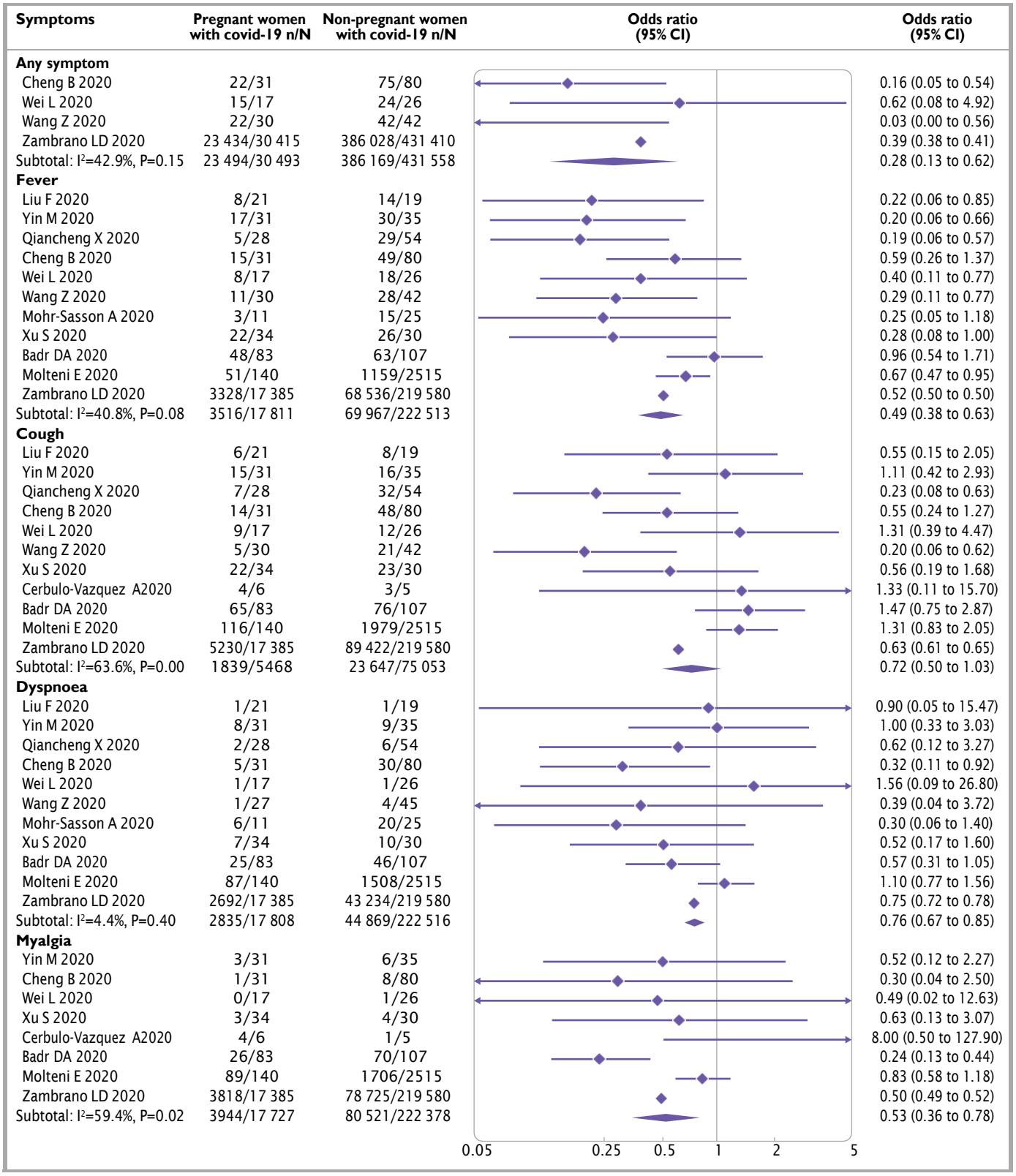


Fig 5 | clinical manifestations of coronavirus disease 2019 (covid-19) in pregnant and recently pregnant women compared with non-pregnant women of reproductive age with covid-19

strengths and limitations of this review

In this unprecedented pandemic situation, where evidence is rapidly produced and published in various formats, our living systematic review underpinned by robust methods and continually updated at regular intervals is relevant for several reasons. Firstly, it addresses important research questions relevant to clinical decision making and policies. Secondly, uncertainties remain for key outcomes that require further evidence. Thirdly, the rapid turnover of evidence in various formats requires assessments of study quality and regular updating of the findings. Finally, our living systematic review is producing strong evidence base for living guidelines on covid-19 and pregnancy.

We undertook a comprehensive search and coordinated our efforts with key organisations and research groups, such as WHO, the Cochrane Centre, and EPPI-Centre. To minimise risk of bias we restricted our meta-analysis to cohort studies, and we reported the quality of the included studies. By contacting the authors and obtaining reports not published in PubMed, we minimised the risk of missing relevant studies. Our systematic review has a large sample size and it is continuously increasing. Our living meta-analyses framework will enable us to rapidly update the findings as new data emerge. We undertook extensive work to ensure that duplicate data are not included. Our various comparative analyses allowed us to comprehensively assess the association between pregnancy and covid-19 related outcomes, covid-19 and pregnancy outcomes, risk factors for SARS-CoV-2 infection, and complications. Our review helps to understand the variations in estimates through sensitivity analyses by sampling strategies, population characteristics, and risk factors, and it provides confidence in the rates of reported outcomes. The update has allowed us to seamlessly incorporate new evidence from 115 studies and more than half a million women, published since our original review in June 2020.

Our systematic review also has limitations. The primary studies used varied sampling frames to identify women with covid-19, comprised women with suspected and confirmed covid-19, and primarily reported on pregnant women who required visits to hospital, including for childbirth, thereby affecting the generalisability of the estimates. Although our sensitivity analyses aimed to tackle some of these problems, the numbers and sample sizes of the individual studies were too small to identify differences between the subgroups. The timing of assessment of the clinical manifestations of disease was generally not available. The definitions of symptoms, tests, and outcomes were heterogeneous. Furthermore, poor reporting of the criteria for caesarean section, admissions to the neonatal unit, and the causes of preterm birth, made it difficult to disentangle iatrogenic effect from the true impact of the disease. There continues to be a paucity of comparative data to assess the risk of pregnancy complications in women with and without covid-19. Studies comparing maternal and perinatal outcomes in pregnant women with covid-19 against historical cohorts of pregnant women, could be biased owing to differences in the environment in which deliveries occur. During the pandemic, healthcare systems have faced increased pressure and strain on services, with resulting effects on service delivery and quality of care.[96 97] Lockdown measures, social distancing, and changes to livelihood have led to increased depression and anxiety, and reduction in physical activity and access or attendance

table 1 | Outcomes in pregnant and recently pregnant women with coronavirus disease 2019 (covid-19)

Outcomes	no of studies	women (no with event/no in group (%)) Pregnant women with covid-19	comparison group	Odds ratio (95% ci)	i² (%)
comparison group: non-pregnant women of reproductive age with covid-19					
All cause mortality	8	103/34 047 (0.3)	3388/567 075 (0.6)	0.96 (0.79 to 1.18)	0
ICU admission	7	616/34 035 (1.8)	9568/567 073 (1.7)	2.13 (1.54 to 2.95)	71.2
Invasive ventilation	6	270/34 001 (0.8)	3280/567 043 (0.6)	2.59 (2.28 to 2.94)	0
ECMO	2	17/30 446 (0.1)	120/431 490 (0.0)	2.02 (1.22 to 3.34)	0
Oxygen through nasal cannula	2	8/48 (16.7)	49/106 (46.2)	0.21 (0.04 to 1.13)	65.7
ARDS	1	0/17 (0)	0/26 (0)	1.51 (0.03 to 79.93)	NE
Major organ failure	1	0/17 (0)	0/26 (0)	1.51 (0.03 to 79.93)	NE
comparison group: pregnant women without covid-19					
Maternal outcomes:					
All cause mortality	8*	8/1195 (0.7)	8/3625 (0.2)	2.85 (1.08 to 7.52)	0
ICU admission	7*	64/1508 (4.2)	4/3482 (0.1)	18.58 (7.53 to 45.82)	0
Preterm birth <37 weeks	18	147/1184 (12.4)	572/7365 (7.8)	1.47 (1.14 to 1.91)	18.6
Caesarean section	21*†	669/1854 (36.1)	4221/11842 (35.6)	1.12 (0.91 to 1.38)	57.6
Perinatal outcomes:					
Stillbirth	9*	9/1039 (0.9)	26/4755 (0.5)	2.84 (1.25 to 6.45)	0
Neonatal death	8*	4/970 (0.4)	5/3316 (0.2)	2.77 (0.92 to 8.37)	0
Admission to neonatal unit	10*	329/1285 (25.6)	519/4588 (11.3)	4.89 (1.87 to 12.81)	96.2
Abnormal Apgar score at 5 minutes	6	13/662 (2.0)	46/2823 (1.6)	1.38 (0.71 to 2.70)	0
Fetal distress	2	11/77 (14.3)	13/263 (4.9)	2.37 (0.77 to 7.31)	0

ICU=intensive care unit; ECMO=extracorporeal membrane oxygenation; ARDS=acute respiratory distress syndrome; NE=not estimable.
The denominator is number of pregnancies for all outcomes.
*Includes UK Obstetric Surveillance System[44] study with historical comparative cohort (694 women).
†Includes Gulersen et al 2020[60] with historical comparative cohort (50 women).

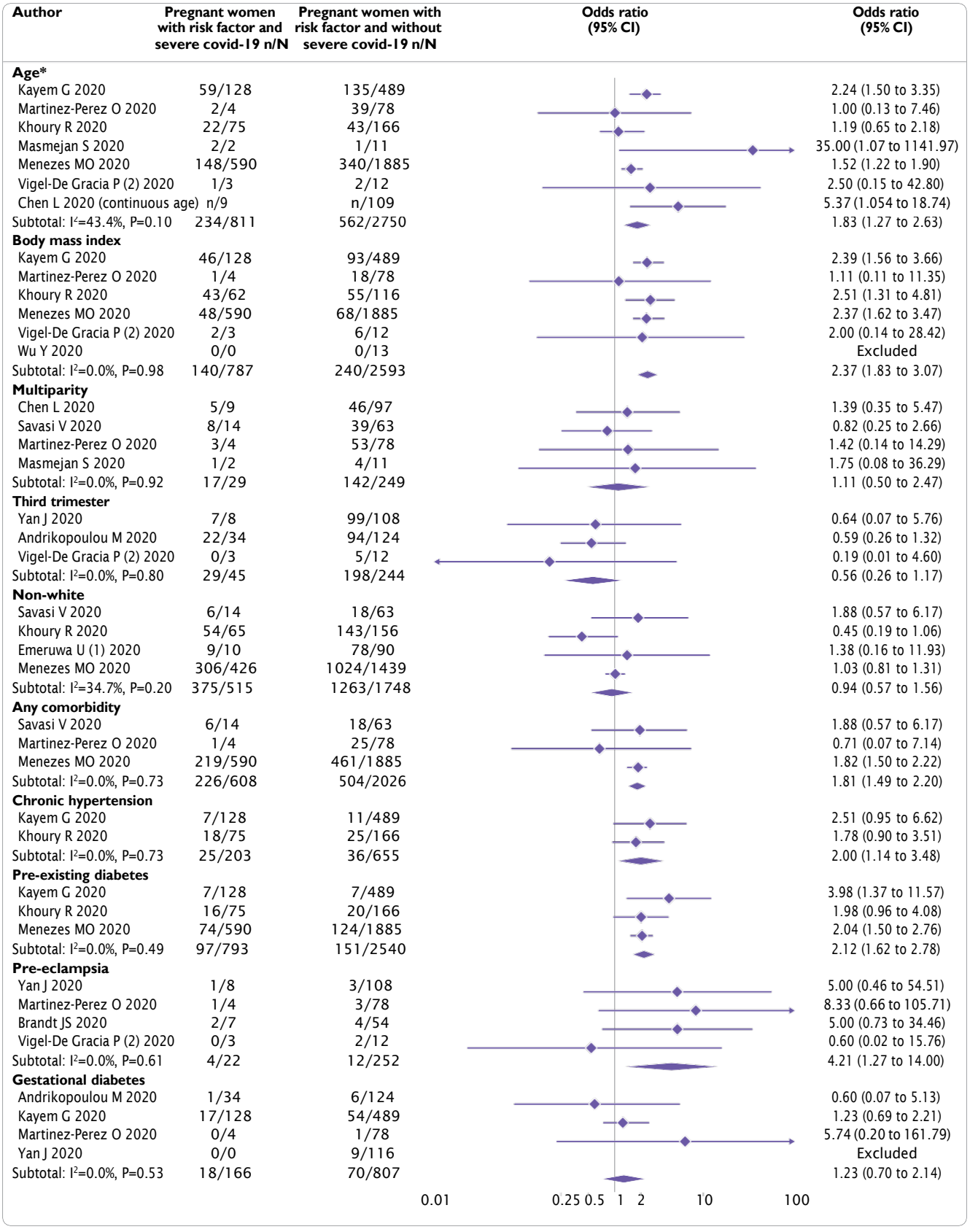


Fig 6 | risk factors associated with severe coronavirus disease 2019 (covid-19) in pregnant and recently pregnant women. symptom based screening: savasi v, Kayem g; nHcc (national Health commission china). criteria based screening: chen, wu, yan. all other studies used universal screening. cut-off for age is 35 years or more, and for body mass index is 30 or more. *includes one study with continuous measurement of risk factor

to healthcare facilities, which could increase the risk of maternal and perinatal complications.[98]

Not many studies reported outcomes by trimester for symptom onset, making it difficult to assess the rates of miscarriage and postpartum complications. For some outcomes, the findings were influenced by a single large study.[42] Many studies had to be excluded as we could not rule out potential overlap in the study populations.

Areas of uncertainty in some of our review findings will still need to be resolved in the next updates of the living systematic review. In seeking an efficient balance between resource consumption and the value the review provides to end users, we will make

table 2 | maternal characteristics associated with severe coronavirus disease 2019 (covid-19) and all cause death in pregnant and recently pregnant women with a diagnosis of covid-19

maternal risk factors and outcomes	no of studies	total no of women	Pregnant women (no with risk factor/no in group (%))		Odds ratio (95% ci)	i² (%)
			with outcome	without outcome		
Age ≥35 years:						
Severe disease	7	3561	811*	2750*	1.83 (1.27 to 2.63)	43
ICU admission	7	31710	348*	31362*	2.11 (1.69 to 2.63)	0
Invasive ventilation	3	718	18*	700*	1.72 (0.60 to 4.97)	17
Maternal death	3	31710	176*	31525*	0.91 (0.22 to 3.72)	93
Multiparity:						
Severe disease	4	278	17/159 (10.7)	12/119 (10.1)	1.11 (0.50 to 2.46)	0
ICU admission	3	815	34/501(6.8)	17/314 (5.4)	1.34 (0.72 to 2.50)	0
Invasive ventilation	1	350	1/216 (0.5)	0/134 (0)	1.87 (0.08 to 46.30)	NE
Body mass index ≥30:						
Severe disease	5	3367	787*	2580*	2.37 (1.83 to 3.07)	0
ICU admission	4	31456	339*	31117*	2.71 (1.10 to 6.63)	63
Invasive ventilation	2	485	12*	4473*	6.61 (1.98 to 22.02)	0
Maternal death	3	31085	113*	30972*	2.27 (1.20 to 4.31)	0
Non-white ethnicity:						
Severe disease	4	2263	375/1638 (22.9)	140/625 (22.4)	0.94 (0.57 to 1.57)	35
ICU admission	4	31543	306/23996 (1.3)	158/7547 (2.1)	1.66 (1.20 to 2.29)	26
Invasive ventilation	1	669	20/134 (14.9)	39/535 (7.3)	2.23 (1.25 to 3.97)	NE
Maternal death	3	31469	110/24 124 (0.5)	36/7345 (0.5)	1.61 (1.05 to 2.47)	0
Any comorbidity:						
Severe disease	3	2634	226/730 (31.0)	382/1904 (20.1)	1.81 (1.49 to 2.20)	0
ICU admission	5	31512	106/6639 (1.6)	226/24 873.9)	1.70 (1.34 to 2.15)	0
Invasive ventilation	3	715	7/71 (9.9)	11/644(1.7)	5.26 (1.76 to 15.68)	0
Maternal death	2	30639	19/6493 (0.3)	33/24 146 (0.1)	2.53 (0.78 to 8.17)	50
Chronic hypertension:						
Severe disease	2	858	25/61 (41.0)	178/797 (22.3)	2.00 (1.14 to 3.48)	0
ICU admission	5	31433	15/262 (5.7)	319/31 171 (1.0)	4.72 (2.37 to 9.41)	13
Invasive ventilation	2	484	5/24 (20.8)	7/460 (1.5)	63.82 (9.69 to 420.45)	0
Maternal death	3	31011	7/249 (2.8)	81/30 762 (0.3)	4.25 (1.82 to 9.95)	0
Pre-existing diabetes:						
Severe disease	3	3333	97/248 (39.1)	696/3085 (22.6)	2.12 (1.62 to 2.78)	0
ICU admission	6	31473	36/638 (5.6)	306/30 835 (1.0)	4.67 (1.94 to 11.22)	38
Invasive ventilation	2	482	2/12 (16.7)	9/470 (1.9)	18.61 (0.26 to 1324.16)	78
Maternal death	2	30723	11/620 (1.8)	41/30 103 (0.1)	14.88 (4.19 to 52.81)	53
Asthma:						
Severe disease	4	3332	39/148 (26.4)	717/3184 (22.5)	1.43 (0.85 to 2.38)	28
ICU admission	1	100	2/9 (22.2)	8/91 (8.8)	2.96 (0.53 to 16.74)	NE
Maternal death	3	889	5/39 (12.8)	63/850 (7.4)	1.68 (0.66 to 4.24)	0
Smoking:						
Severe disease	3	776	5/23 (21.7)	141/753 (18.7)	1.67 (0.64 to 4.40)	0
ICU admission	2	142	1/4 (25.0)	17/138 (12.3)	2.92 (0.35 to 24.23)	0
Maternal death	1	308	0/10 (0)	7/298 (2.3)	1.85 (0.10 to 34.60)	NE
Gestation ≥28 weeks:						
Severe disease	3	289	29/227 (12.8)	16/62 (25.8)	0.56 (0.27 to 1.17)	0
Maternal death	1	721	46/495 (9.3)	23/226 (10.2)	0.90 (0.53 to 1.53)	NE
Gestational diabetes:						
Severe disease	4	973	18/88 (20.5)	148/885 (16.7)	1.23 (0.70 to 2.14)	0
ICU admission	2	777	11/81 (13.6)	31/696 (4.5)	3.27 (1.55 to 6.89)	0
Invasive ventilation	1	350	0/32 (0)	0/318 (0)	—	NE
Pre-eclampsia:						
Severe disease	4	274	4/16 (25.0)	18/258 (7.0)	4.21 (1.27 to 14.00)	0
ICU admission	1	42	6/6 (100.0)	2/36 (5.6)	179.40 (7.69 to 4186.05)	NE

ICU=intensive care unit; NE=not estimable.
*Includes one or more studies with continuous measurement of risk factor.

decisions about the pacing of the updates of our living systematic review using a formal framework for decision making. We will use a mixed approach based on the Ottawa method to identify quantitative or qualitative signals for the need of an update,[99] and a more complex statistical prediction tool to estimate the probability that new studies identified would change the review conclusions.[100]

comparison with existing evidence

Between the publication of the original living systematic review and this update, estimates for the prevalence of covid-19, and rates of clinical manifestations and outcomes of pregnant and recently pregnant women with covid-19 have remained similar, with improved precision in the findings. The rates for postpartum haemorrhage and admission to the neonatal unit appear to be slightly increased from the first version, while the rate of maternal pneumonia appears to be lower. High heterogeneity remains in the estimates for rates of clinical manifestations and outcomes.

We found that the same risk factors for severe covid-19 identified in the original version of the living systematic review remained associated with severe covid-19 with increased precision. Additional risk factors for severe disease, such as non-white ethnicity identified in this update, were also identified from large cohort studies such as the UK Obstetric Surveillance System and the US CDC surveillance report.[42 101] Our findings are consistent with the reports of disproportionately high rates of severe covid-19 in non-pregnant ethnic minority populations,[102] and in other areas of maternity care.[103 104] The observed disparity could be attributed to associated comorbidities, socioeconomic characteristics, and factors related to access to and quality of care in the preconception, pregnancy, and postpartum periods.[105] The multifaceted contributors to ethnic disparities need to be investigated to reduce mortality and morbidity related to both covid-19 and pregnancy.

Our review update also identified an increased risk for maternal death, need for maternal admission to the intensive care unit, and stillbirth in pregnant women with covid-19 compared to pregnant women without the disease. However, our confidence in these estimates is not high, owing to the small numbers of events in both groups. Further data are still needed to robustly assess these outcomes, along with the emerging data on increased risk of severe outcomes such as the need for ECMO.[42]

Alongside the spread of the pandemic, a shift has occurred in the types of studies published, with initial studies involving pregnant women from epidemic regions in China, followed by reports of large regional and national datasets from the US, UK, Netherlands, Spain, and, more recently, Latin American countries. The study design has also changed from initial small case series and case reports to large observational data, with recent studies also providing comparative data.

The prevalence of covid-19 varied widely between studies, particularly when sampling was done based on symptoms or history of contact, highlighting the variations in criteria for testing. The current update includes 50 new studies from 11 additional countries on the prevalence of covid-19 in pregnancy. Despite the addition of five times more studies between the original version of our living systematic review and this update, from diverse populations globally, the prevalence of covid-19 in pregnant and recently pregnant women remains unchanged. Unlike the general population who are mostly tested for SARS-CoV-2 on the basis of symptoms or contact history, universal screening of all pregnant women attending the hospital for any reason could contribute to the consistency in the findings. However, the true prevalence of covid-19 in pregnancy is likely to be lower than the current estimate if all pregnant women, including those not attending the hospital are included.

In the recent cohort study of all individuals admitted with covid-19 in the UK, the cluster of respiratory symptoms of cough, fever, and breathlessness were observed in more than two thirds of individuals,[106] similar to reported rates in the US and China.[107-109] But in our review, fewer pregnant and recently pregnant women with covid-19 manifested these symptoms than the non-pregnant population, indicating possible high rates of asymptomatic presentation in this population. This is likely because of the strategy of universal screening for covid-19 in pregnancy and the low thresholds for testing in pregnant women than in non-pregnant women. Despite the potential higher possibility of universal screening to detect pregnant women with mild disease, we observed an increase in admissions to the intensive care unit and need for invasive ventilation compared with non-pregnant women of reproductive age with covid-19. The findings were mainly influenced by the recently updated large Centers for Disease Control and Prevention report from the US,[42] and a report from the Mexican General Directorate of Epidemiology registry.[41]

By accessing the unpublished data from our collaborators, we were able to include both women with and without symptoms from the US CDC surveillance data, in addition to the women with symptoms only who were included in the published report.[42] Pregnancy status was not ascertained in a large proportion of women of reproductive age in the CDC report, which could affect the estimates. Furthermore, the outcomes for which the data were missing from the report were considered to be absent, potentially leading to bias. The report from the Mexican General Directorate of Epidemiology registry, available only as a preprint, included only women with symptoms who might be at high risk of complications. We recommend that studies comparing covid-19 related outcomes in pregnant versus non-pregnant women report the relevant estimates for both women with and without symptoms to avoid overestimation of the risk of complications due to selective reporting. The pooled estimates for severe covid-19 and admission to an intensive care

unit were, however, still relatively high in the non-comparative data, indicative of a potential high risk in pregnancy. This is supported by the recent analysis in a Swedish study suggesting a high risk of admission to an intensive care unit and invasive ventilation in pregnant women compared to non-pregnant women.[110]

Similar to the general population, high body mass index and pre-existing comorbidity seemed to be risk factors for severity of covid-19 in pregnancy, including admission to an intensive care unit and invasive ventilation.[106] Complications related to covid-19 did not seem to be increased in women presenting in the third trimester versus earlier in pregnancy or in multiparous versus primiparous women—but existing sample sizes are not large. Both chronic hypertension and pre-existing diabetes were associated with maternal death in pregnant women with covid-19, which are known risk factors in the general population. But it is not known if covid-19 was the direct cause of death for these women, and the numbers of studies are small. We observed an increase in rates of preterm birth in pregnant women with covid-19 compared with pregnant women without the disease. These preterm births could have been medically indicated, as the overall rates of spontaneous preterm births in pregnant women with covid-19 was broadly similar to those observed in the pre-pandemic period. Although about 50% of pregnant women underwent caesarean section in the non-comparative studies, we did not find a statistically significant difference in comparative studies of pregnant women with and without covid-19. The precision of the estimates is expected to improve with the publication of more data in the future. The overall rates of stillbirths and neonatal deaths do not seem to be higher than the background rates. The indications for admissions to the neonatal unit, observed in about a third of neonates delivered to mothers with covid-19, have not been reported. Local policies on observation and quarantine of infants with exposure to SARS-CoV-2 might have influenced these rates.

relevance for clinical practice and research

Based on existing data, healthcare professionals should be aware that pregnant and recently pregnant women with covid-19 might manifest fewer symptoms than the general population, with the overall pattern similar to that of the general population. Pregnant women should be informed of the increase in severity of covid-19 including admission to intensive care units, need for ECMO and invasive ventilation compared with non-pregnant women, and encouraged to undertake safety measures to reduce the risk of infection. Pregnant women with pre-existing comorbidities will need to be considered as a high risk group for covid-19, along with those who are obese and of older maternal age. Healthcare professionals need to be aware of the increased risk of severe covid-19 in pregnant and recently pregnant women of non-white ethnic origin, to plan close monitoring and have a low threshold for escalation of care. Clinicians will need to balance the need for regular multidisciplinary antenatal care to manage women with pre-existing comorbidities against unnecessary exposure to the virus, through virtual clinic appointments when possible. Pregnant women with covid-19 before term gestation might need to be managed in a unit with facilities to care for preterm neonates.

Further data are still needed to assess robustly if pregnancy related maternal and neonatal complications are increased in women with covid-19 compared to pregnant women without the disease. Similarly, the association between pregnancy specific risk factors such as pre-eclampsia and gestational diabetes on covid-19 related outcomes needs further evaluation. Pre-eclampsia was reported to be associated with severe covid-19 in small studies, but this requires further assessment as the clinical and laboratory presentation of severe pre-eclampsia could mimic worsening covid-19.[111] Robust collection of maternal data by trimester of exposure, including the periconception period, is required to determine the effects of covid-19 on early pregnancy outcomes, fetal growth, and risk of miscarriage or stillbirth. We need detailed reporting of outcomes by ethnicity to quantify the risk of severe covid-19 in women from different ethnicities. Qualitative studies on behaviour and attitude to the pandemic can disentangle the relative importance of factors behind the ethnic disparities observed in the severity of covid-19.

Systematic reviews are considered to be the highest quality evidence informing guidelines, and poor quality reviews will have a direct impact on clinical care. Despite the urgent need for evidence on the impact of covid-19 in pregnant women, systematic reviews and meta-analyses still need to adhere to the reporting guidelines on search criteria, quality assessment, and analysis. This is particularly important as large numbers of non-peer reviewed scientific papers and reports are currently available in the public domain in multiple versions. Primary studies need to explicitly state if duplicate data have been included to avoid double counting of participants in evidence synthesis. Individual participant data meta-analysis of the emerging cohorts is critical to assess both differential presentation and outcomes by underlying risk factors, and to determine the differential effects of interventions to reduce the rates of complications. With the establishment of several national and global prospective cohorts, we expect the sample size of our meta-analysis to increase further in the coming months. Our living systematic review and meta-analysis with its regular search and analyses updates is ideally placed to assess the impact of new findings on the rapidly growing evidence base.

autHOr aFFiliatiOns

[1]Institute of Applied Health Research, University of Birmingham, Birmingham, UK

[2]WHO Collaborating Centre for Global Women's Health, Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK

[3]Clinical Biostatistics Unit, Hospital Universitario Ramón y Cajal (IRYCIS), Madrid, Spain

[4]CIBER Epidemiology and Public Health (CIBERESP), Madrid, Spain
[5]UNDP/UNFPA/UNICEF/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP), Department of Sexual and Reproductive Health and Research, World Health Organization, Geneva, Switzerland
[6]Birmingham Medical School, University of Birmingham, Birmingham, UK
[7]Division of Birth Cohort Study, Guangzhou Women and Children's Medical Centre, Guangzhou Medical University, Guangzhou, China
[8]Department of Woman and Child Health Care, Guangzhou Women and Children's Medical Centre, Guangzhou Medical University, Guangzhou, China
[9]Department of Obstetrics and Gynaecology, Guangzhou Women and Children's Medical Centre, Guangzhou Medical University, Guangzhou, China
[10]Netherlands Satellite of the Cochrane Gynaecology and Fertility Group, Amsterdam University Medical Centre, Amsterdam, Netherlands
[11]Department of Obstetrics and Gynaecology, Amsterdam University Medical Centre, Amsterdam, Netherlands
[12]Blizard Institute, Queen Mary University of London, London, UK
[13]Barts Health NHS Trust, London, UK
[14]St George's, University of London, London, UK
[15]Sexual and Reproductive Health care, Catalan Health Service, Barcelona, Catalonia, Spain
[16]Elizabeth Glaser Paediatric AIDS Foundation, Washington, DC, USA
[17]Women's Health Research Unit, Queen Mary University of London, London, UK
[18]Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK
Other members of the PregCOV-19 Living Systematic Review Consortium are Van T Tong, Sascha Ellington, Gianfranco Spiteri, Julien Beaute, Uma Ram, Ajith S Nair, Pura Rayco-Solon, and Hector Pardo-Hernandez. The Cochrane Gynaecology and Fertility Group thank Maxime Verschuuren, Marijke Strikwerda, and Bethany Clark for help with searches and data extraction. The PregCOV-19 Living Systematic Review Group would also like to thank Katie's Team for its contribution towards the development and reporting of this work, James Thomas from the EPPI-Centre for helping with search updates, and the PregCOV-19 Living Systematic Review steering committee members, Pisake Lumbiganon, Carolina Carvalho Ribeiro do Valle, Samantha Lissauer, Clare Whitehead, David Lissauer, Joao Paulo Souza, and Marian Knight, who provided guidance throughout.

Contributors: ST, MB, and JA conceptualised the study. MY, SC, LD, TK, ACL, AD, DZ, RB, SL, XQ, MYuan, JS, HL, and KA selected the studies. JA, ES, MY, LD, DZ, XQ, and MYuan extracted the data. JZ conducted the analyses. JA and ES are joint first authors. All coauthors contributed to the writing of the manuscript and approved the final version. ST, JA, ES, and JZ are the guarantors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted

Funding: The project was partially funded by the World Health Organization and UNDP-UNFPA-UNICEF-WHO-World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP), a cosponsored programme executed by the World Health Organization (WHO). The authors alone are responsible for the views expressed in this article and they do not necessarily represent the views, decisions or policies of the institutions with which they are affiliated.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: partial funding by the World Health Organization and HRP; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required.

Data sharing: No additional data available.

The corresponding author (ST) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been disclosed.

Dissemination to participants and related patient and public communities: The PregCov-19 LSR Group will disseminate the findings through a dedicated website (www.birmingham.ac.uk/research/who-collaborating-centre/pregcov/index.aspx) and social media.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;395:497-506. doi:10.1016/S0140-6736(20)30183-5
2 World Health Organization (WHO). Coronavirus disease. (COVID-19) Pandemic, https://www.who.int/emergencies/diseases/novel-coronavirus-2019 (accessed 7 May 2020)
3 Cabinet Office. Guidance. Staying alert and safe (social distancing). Coronavirus (COVID-19) Guidance and support. Updated 22 May 2020 https://www.gov.uk/government/publications/staying-alert-and-safe-social-distancing/staying-alert-and-safe-social-distancing (accessed 24 May 2020).
4 RCOG. (COVID-19) Infection in Pregnancy, https://www.rcog.org.uk/en/guidelines-research-services/guidelines/coronavirus-pregnancy/
5 Allotey J, Stallings E, Bonet M, et al, for PregCOV-19 Living Systematic Review Consortium. Clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy: living systematic review and meta-analysis. *BMJ* 2020;370:m3320. doi:10.1136/bmj.m3320
6 Zaigham M, Andersson O. Maternal and perinatal outcomes with COVID-19: A systematic review of 108 pregnancies. *Acta Obstet Gynecol Scand* 2020;99:823-9. doi:10.1111/aogs.13867
7 Parazzini F, Bortolus R, Mauri PA, Favilli A, Gerli S, Ferrazzi E. Delivery in pregnant women infected with SARS-CoV-2: A fast review. *Int J Gynaecol Obstet* 2020;150:41-6. doi:10.1002/ijgo.13166
8 Elshafeey F, Magdi R, Hindi N, et al. A systematic scoping review of COVID-19 during pregnancy and childbirth. *Int J Gynaecol Obstet* 2020;150:47-52. doi:10.1002/ijgo.13182
9 Di Mascio D, Khalil A, Saccone G, et al. Outcome of coronavirus spectrum infections (SARS, MERS, COVID-19) during pregnancy: a systematic review and meta-analysis. *Am J Obstet Gynecol MFM* 2020;2:100107.
10 Della Gatta AN, Rizzo R, Pilu G, Simonazzi G. Coronavirus disease 2019 during pregnancy: a systematic review of reported cases. *Am J Obstet Gynecol* 2020;223:36-41. doi:10.1016/j.ajog.2020.04.013
11 Cheruiyot I, Henry BM, Lippi G. Is there evidence of intra-uterine vertical transmission potential of COVID-19 infection in samples tested by quantitative RT-PCR?*Eur J Obstet Gynecol Reprod Biol* 2020;249:100-1. doi:10.1016/j.ejogrb.2020.04.034
12 Gajbhiye R, Modi D, Mahale S. Pregnancy outcomes, Newborn complications and Maternal-Fetal Transmission of SARS-CoV-2 in women with COVID-19: A systematic review of 441 cases. medRxiv [Preprint]. 2020. doi.org/10.1101/2020.04.11.20062356
13 Murad MH, Sultan S, Haffar S, Bazerbachi F. Methodological quality and synthesis of case series and case reports. *BMJ Evid Based Med* 2018;23:60-3. doi:10.1136/bmjebm-2017-110853
14 Breslin N, Baptiste C, Gyamfi-Bannerman C, et al. Coronavirus disease 2019 infection among asymptomatic and symptomatic pregnant women: Two weeks of confirmed presentations to an affiliated pair of New York City hospitals. *Am J Obstet Gynecol MFM* 2020;2:100118.
15 Vintzileos WS, Muscat J, Hoffmann E, et al. Screening all pregnant women admitted to labor and delivery for the virus responsible for coronavirus disease 2019. *Am J Obstet Gynecol* 2020;223:284-6.
16 Xu L, Yang Q, Shi H, et al. Clinical presentations and outcomes of SARS-CoV-2 infected pneumonia in pregnant women and health status of their neonates. *Sci Bull (Beijing)* 2020;65:1537-42.
17 Blitz MJ, Grunebaum A, Tekbali A, et al. Intensive care unit admissions for pregnant and nonpregnant women with coronavirus disease 2019. *Am J Obstet Gynecol* 2020;223:290-1.
18 Released by National Health Commission & National Administration of Traditional Chinese Medicine on 3 March 2020). Trial Version 7. *Chin Med J (Engl)* 2020;133:1087-95.
19 Allotey J, Bonet M, Zamora J, et al. COVID-19 in pregnant women: a Living Systematic Review on prevalence, presentation, prognosis and treatment. PROSPERO 2020 CRD42020178076. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020178076.
20 Yap M, Debenham L, Kew T, et al, PregCOV-19 Consortium. Clinical manifestations, prevalence, risk factors, outcomes, transmission, diagnosis and treatment of COVID-19 in pregnancy and postpartum: a living systematic review protocol. *BMJ Open* 2020;10:e041868. doi:10.1136/bmjopen-2020-041868
21 Living Overview of the Evidence. (LOVE) Platform. app.iloveevidence.com. (accessed 8 July 2020)

22 World Health Organization. Global surveillance for COVID-19 caused by human infection with COVID-19 virus: interim guidance, 20 March 2020. World Health Organization. https://apps.who.int/iris/handle/10665/331506. 2020.
23 Dekkers OM, Egger M, Altman DG, Vandenbroucke JP. Distinguishing case series from cohort studies. *Ann Intern Med* 2012;156:37-40. doi:10.7326/0003-4819-156-1-201201030-00006
24 Wells G. Proceedings or the Third Symposium on Systematic Reviews beyond the Basics. SBOD. Improving Quality and Impact; The Newcastle–Ottawa Scale (NOS) for Assessing the Quality of non-randomised Studies in Meta-analysis. July 3-5 2000, Oxford.
25 Hoy D, Brooks P, Woolf A, et al. Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement. *J Clin Epidemiol* 2012;65:934-9. doi:10.1016/j.jclinepi.2011.11.014
26 Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Stat Med* 2000;19:3127-31. doi:10.1002/1097-0258(20001130)19:22<3127::AID-SIM784>3.0.CO;2-M
27 Ellington S, Strid P, Tong VT, et al. Characteristics of Women of Reproductive Age with Laboratory-Confirmed SARS-CoV-2 Infection by Pregnancy Status - United States, January 22-June 7, 2020. *MMWR Morb Mortal Wkly Rep* 2020;69:769-75. doi:10.15585/mmwr.mm6925a1
28 Cohen J, Vignaux O, Jacquemard F. Covid-19 in pregnant women: General data from a French National Survey. *Eur J Obstet Gynecol Reprod Biol* 2020;251:267-8. doi:10.1016/j.ejogrb.2020.06.002
29 Cheng B, Jiang T, Zhang L, et al. Clinical Characteristics of Pregnant Women with Coronavirus Disease 2019 in Wuhan, China. *SSRN* 2020; https://ssrn.com/abstract=3555240
30 Liu F, Liu H, Li J, Hou L, Lan W, Wang D. Clinico-Radiological Features and Outcomes in Pregnant Women with COVID-19: Compared with Age-Matched Non-Pregnant Women. *SSRN* 2020. https://ssrn.com/abstract=3556647
31 Mohr-Sasson A, Chayo J, Bart Y, et al. Laboratory characteristics of pregnant compared to non-pregnant women infected with SARS-CoV-2. *Arch Gynecol Obstet* 2020; published online 22 June. doi:10.1007/s00404-020-05655-7
32 Qiancheng X, Jian S, Lingling P, et al, sixth batch of Anhui medical team aiding Wuhan for COVID-19. Coronavirus disease 2019 in pregnancy. *Int J Infect Dis* 2020;95:376-83. doi:10.1016/j.ijid.2020.04.065
33 Wang Z, Wang Z, Xiong G. Clinical characteristics and laboratory results of pregnant women with COVID-19 in Wuhan, China. *Int J Gynaecol Obstet* 2020. doi:10.1002/ijgo.13265
34 Wei L, Gao X, Chen S, et al. Clinical Characteristics and Outcomes between Pregnant and Non-Pregnant Women with Coronavirus Disease 2019: A Retrospective Cohort Study. SSRN2020. https://ssrn.com/abstract=3569858
35 Yin M, Zhang L, Deng G, et al. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infection During Pregnancy In China: A Retrospective Cohort Study. medRxiv [Preprint]. 2020. doi.org/10.1101/2020.04.07.20053744
36 Blitz MJ, Grünebaum A, Tekbali A, et al. Intensive care unit admissions for pregnant and nonpregnant women with coronavirus disease 2019. *Am J Obstet Gynecol* 2020;223:290-1. doi:10.1016/j.ajog.2020.05.004
37 Xu S, Shao F, Bao B, et al. Clinical Manifestation and Neonatal Outcomes of Pregnant Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. *Open Forum Infect Dis* 2020;7:a283. doi:10.1093/ofid/ofaa283
38 Cérbulo-Vázquez A, Zavala-Barrios B, Briones-Garduño JC, et al. Serological Cytokine and chemokine profile in pregnant women with COVID19 in Mexico City.*medRxiv* 2020. doi:10.1101/2020.07.14.20153585
39 Badr DA, Mattern J, Carlin A, et al. Are clinical outcomes worse for pregnant women at ≥20 weeks' gestation infected with coronavirus disease 2019? A multicenter case-control study with propensity score matching. *Am J Obstet Gynecol* 2020;223:764-8. doi:10.1016/j.ajog.2020.07.045
40 Molteni E, Astley CM, Ma W, et al. SARS-CoV-2 (COVID-19) infection in pregnant women: characterization of symptoms and syndromes predictive of disease and severity through real-time, remote participatory epidemiology.*medRxiv* 2020. doi:10.1101/2020.08.17.20161760
41 Leon-Abarca JA, Pena-Gallardo MT, Soliz J, Accinelli RA. Clinical evolution of COVID-19 during pregnancy at different altitudes: a population-based study.*medRxiv* 2020. doi:10.1101/2020.09.14.20193177
42 Zambrano LD, Ellington S, Strid P, et al, CDC COVID-19 Response Pregnancy and Infant Linked Outcomes Team. Update: Characteristics of Symptomatic Women of Reproductive Age with Laboratory-Confirmed SARS-CoV-2 Infection by Pregnancy Status - United States, January 22-October 3, 2020. *MMWR Morb Mortal Wkly Rep* 2020;69:1641-7. doi:10.15585/mmwr.mm6944e3
43 Campbell KH, Tornatore JM, Lawrence KE, et al. Prevalence of SARS-CoV-2 Among Patients Admitted for Childbirth in Southern Connecticut. *JAMA* 2020. doi:10.1001/jama.2020.8904
44 Knight M, Bunch K, Vousden N, et al. Characteristics and outcomes of pregnant women hospitalised with confirmed SARS-CoV-2 infection in the UK: a national cohort study using the UK Obstetric Surveillance System (UKOSS). 2020.doi.org/10.1101/2020.05.08.20089268
45 Li N, Han L, Peng M, et al. Maternal and neonatal outcomes of pregnant women with COVID-19 pneumonia: a case-control study. *Clin Infect Dis* 2020;ciaa352. doi:10.1093/cid/ciaa352
46 Liao J, He X, Gong Q, Yang L, Zhou C, Li J. Analysis of vaginal delivery outcomes among pregnant women in Wuhan, China during the COVID-19 pandemic. *Int J Gynaecol Obstet* 2020;150:53-7. doi:10.1002/ijgo.13188
47 Ahlberg M, Neovius M, Saltvedt S, et al. Association of SARS-CoV-2 Test Status and Pregnancy Outcomes. *JAMA* 2020;324:1782. doi:10.1001/jama.2020.19124
48 Bender WR, Srinivas S, Coutifaris P, Acker A, Hirshberg A. The Psychological Experience of Obstetric Patients and Health Care Workers after Implementation of Universal SARS-CoV-2 Testing. *Am J Perinatol* 2020;37:1271-9. doi:10.1055/s-0040-1715505
49 Bianco A, Buckley AB, Overbey J, et al. Testing of patients and support persons for coronavirus disease 2019 (COVID-19) infection before scheduled deliveries. *Obstet Gynecol* 2020;136:283-7. doi:10.1097/AOG.0000000000003985
50 Brandt JS, Hill J, Reddy A, et al. Epidemiology of coronavirus disease 2019 in pregnancy: risk factors and associations with adverse maternal and neonatal outcomes. *Am J Obstet Gynecol* 2020;S0002-9378(20)31134-0.
51 Buckley A, Bianco A, Stone J. Universal testing of patients and their support persons for severe acute respiratory syndrome coronavirus 2 when presenting for admission to labor and delivery at Mount Sinai Health System. *Am J Obstet Gynecol MFM* 2020;2:100147. doi:10.1016/j.ajogmf.2020.100147
52 Cosma S, Borella F, Carosso A, et al. The "scar" of a pandemic: Cumulative incidence of COVID-19 during the first trimester of pregnancy. *J Med Virol* 2021;93:537-40. doi:10.1002/jmv.26267
53 Egerup P, Fich Olsen L, Hellerung Christiansen A-M, et al. Impact of SARS-CoV-2 antibodies at delivery in women, partners and newborns. *medRxiv* 2020. doi:10.1101/2020.09.14.20191106
54 Emeruwa UN, Ona S, Shaman JL, et al. Associations Between Built Environment, Neighborhood Socioeconomic Status, and SARS-CoV-2 Infection Among Pregnant Women in New York City. *JAMA* 2020;324:390-2. doi:10.1001/jama.2020.11370
55 Facchetti F, Bugatti M, Drera E, et al. SARS-CoV2 vertical transmission with adverse effects on the newborn revealed through integrated immunohistochemical, electron microscopy and molecular analyses of Placenta. *EBioMedicine* 2020;59:102951. doi:10.1016/j.ebiom.2020.102951
56 Farghaly MAA, Kupferman F, Castillo F, Kim RM. Characteristics of Newborns Born to SARS-CoV-2-Positive Mothers: A Retrospective Cohort Study. *Am J Perinatol* 2020;37:1310-6. doi:10.1055/s-0040-1715862
57 Fassett MJ, Lurvey LD, Yasumura L, et al. Universal SARS-Cov-2 Screening in Women Admitted for Delivery in a Large Managed Care Organization. *Am J Perinatol* 2020;37:1110-4. doi:10.1055/s-0040-1714060
58 Flaherman VJ, Afshar Y, Boscardin J, et al. Infant Outcomes Following Maternal Infection with SARS-CoV-2: First Report from the PRIORITY Study. *Clin Infect Dis* 2020;ciaa1411. doi:10.1093/cid/ciaa1411
59 Goldfarb IT, Clapp MA, Soffer MD, et al. Prevalence and Severity of Coronavirus Disease 2019 (COVID-19) Illness in Symptomatic Pregnant and Postpartum Women Stratified by Hispanic Ethnicity. *Obstet Gynecol* 2020;136:300-2. doi:10.1097/AOG.0000000000004005
60 Gulersen M, Prasannan L, Tam Tam H, et al. Histopathologic evaluation of placentas after diagnosis of maternal severe acute respiratory syndrome coronavirus 2 infection. *Am J Obstet Gynecol MFM* 2020;2:100211. doi:10.1016/j.ajogmf.2020.100211
61 Kalafat E, Yassa M, Koc A, Tug N, TULIP collaboration. Utility of lung ultrasound assessment for probable SARS-CoV-2 infection during pregnancy and universal screening of asymptomatic individuals. *Ultrasound Obstet Gynecol* 2020;56:624-6. doi:10.1002/uog.23099
62 Khalil A, Hill R, Ladhani S, Pattisson K, O'Brien P. Severe acute respiratory syndrome coronavirus 2 in pregnancy: symptomatic pregnant women are only the tip of the iceberg. *Am J Obstet Gynecol* 2020;223:296-7. doi:10.1016/j.ajog.2020.05.005
63 LaCourse SM, Kachikis A, Blain M, et al. Low prevalence of SARS-CoV-2 among pregnant and postpartum patients with universal

screening in Seattle, Washington. *Clin Infect Dis* 2020;ciaa675. doi:10.1093/cid/ciaa675
64 Li M, Yin H, Jin Z, et al. Impact of Wuhan lockdown on the indications of cesarean delivery and newborn weights during the epidemic period of COVID-19. *PLoS One* 2020;15:e0237420. doi:10.1371/journal.pone.0237420
65 Liu W, Cheng H, Wang J, et al. Clinical Analysis of Neonates Born to Mothers with or without COVID-19: A Retrospective Analysis of 48 Cases from Two Neonatal Intensive Care Units in Hubei Province. *Am J Perinatol* 2020;37:1317-23. doi:10.1055/s-0040-1716505
66 Martinez-Perez O, Rodriguez PP, Hernandez MM, et al. The association between COVID-19 and preterm delivery: A cohort study with a multivariate analysis.*medRxiv* 2020.
67 Maru S, Patil U, Carroll-Bennett R, et al. Universal screening for SARS-CoV-2 infection among pregnant women at Elmhurst 2 Hospital Center, Queens, New York.*medRxiv* 2020. doi:10.1101/2020.08.12.20171694
68 Miller ES, Grobman WA, Sakowicz A, Rosati J, Peaceman AM. Clinical Implications of Universal Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Testing in Pregnancy. *Obstet Gynecol* 2020;136:232-4. doi:10.1097/AOG.0000000000003983
69 Naqvi M, Burwick RM, Ozimek JA, Greene NH, Kilpatrick SJ, Wong MS. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Universal Testing Experience on a Los Angeles Labor and Delivery Unit. *Obstet Gynecol* 2020;136:235-6. doi:10.1097/AOG.0000000000003987
70 Nayak AH, Kapote DS, Fonseca M, et al. Impact of the Coronavirus Infection in Pregnancy: A Preliminary Study of 141 Patients. *J Obstet Gynaecol India* 2020;70:256-61. doi:10.1007/s13224-020- 01335-3
71 Norooznezhad AH, Eskandarion S, Akbari R, et al. Changes of leukocytes, neutrophils, and lymphocytes count and dependent variables in pregnant women with coronavirus disease 2019 before and after cesarean delivery. *J Med Virol* 2021;93:664-6.
72 Ochiai D, Kasuga Y, Iida M, Ikenoue S, Tanaka M. Universal screening for SARS-CoV-2 in asymptomatic obstetric patients in Tokyo, Japan. *Int J Gynaecol Obstet* 2020;150:268-9. doi:10.1002/ijgo.13252
73 Peng S, Zhu H, Yang L, et al. A Study of Breastfeeding Practices, SARS-CoV-2 and Its Antibodies in the Breast Milk of Mothers Confirmed with COVID-19. *Lancet Regional Health Western Pacific* 2020. https://doi.org/10.1016/j.lanwpc.2020.100045
74 Pineles BL, Alamo IC, Farooq N, et al. Racial-ethnic disparities and pregnancy outcomes in SARS-CoV-2 infection in a universally-tested cohort in Houston, Texas. *Eur J Obstet Gynecol Reprod Biol* 2020;254:329-30. doi:10.1016/j.ejogrb.2020.09.012
75 Pirjani R, Hosseini R, Soori T, et al. Maternal and neonatal outcomes in COVID-19 infected pregnancies: a prospective cohort study. *J Travel Med* 2020;27:taaa158. doi:10.1093/jtm/taaa158
76 Prabhu M, Cagino K, Matthews KC, et al. Pregnancy and postpartum outcomes in a universally tested population for SARS-CoV-2 in New York City: a prospective cohort study. *BJOG* 2020;127:1548-56. doi:10.1111/1471-0528.16403
77 Ruggiero M, Somigliana E, Tassis B, et al. Covid-19 in the second half of pregnancy:prevalence and clinical relevance.*Research Square* 2020.
78 Sakowicz A, Ayala AE, Ukeje CC, Witting CS, Grobman WA, Miller ES. Risk factors for severe acute respiratory syndrome coronavirus 2 infection in pregnant women. *Am J Obstet Gynecol MFM* 2020;2:100198. doi:10.1016/j.ajogmf.2020.100198
79 Smithgall MC, Liu-Jarin X, Hamele-Bena D, et al. Third-trimester placentas of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)-positive women: histomorphology, including viral immunohistochemistry and in-situ hybridization. *Histopathology* 2020;77:994-9. doi:10.1111/his.14215
80 Sutton D, Fuchs K, D'Alton M, Goffman D. Universal Screening for SARS-CoV-2 in Women Admitted for Delivery. *N Engl J Med* 2020;382:2163-4. doi:10.1056/NEJMc2009316
81 van Keulen BJ, Romijn M, Bondt A, et al. Breastmilk; a source of SARS-CoV-2 specific IgA antibodies.*medRxiv* 2020.
82 Vintzileos WS, Muscat J, Hoffmann E, et al. Screening all pregnant women admitted to labor and delivery for the virus responsible for coronavirus disease 2019. *Am J Obstet Gynecol* 2020;223:284-6. doi:10.1016/j.ajog.2020.04.024
83 Zhang P, Salafia C, Heyman T, Salafia C, Lederman S, Dygulska B. Detection of severe acute respiratory syndrome coronavirus 2 in placentas with pathology and vertical transmission. *Am J Obstet Gynecol MFM* 2020;2:100197. doi:10.1016/j.ajogmf.2020.100197
84 Crovetto F, Crispi F, Llurba E, Figueras F, Gomez-Roig MD, Gratacos E. Seroprevalence and clinical spectrum of SARS-Cov-2 infection in the first versus third trimester of pregnancy. *medRxiv* 2020.
85 Yang H, Xie F, Cheng Y, et al. Research on Pregnant Women Suspected of Having COVID-19 in the Epidemic Outbreak Area.*SSRN* 2020.
86 Wang MJ, Schapero M, Iverson R, Yarrington CD. Obstetric Hemorrhage Risk Associated with Novel COVID-19 Diagnosis from a Single-Institution Cohort in the United States. *Am J Perinatol* 2020;37:1411-6. doi:10.1055/s-0040-1718403
87 Mattern J, Vauloup-Fellous C, Zakaria H, et al. Post lockdown COVID-19 seroprevalence and circulation at the time of delivery, France. *PLoS One* 2020;15:e0240782. doi:10.1371/journal.pone.0240782
88 Díaz-Corvillón P, Mönckeberg M, Barros A, et al. Routine screening for SARS CoV-2 in unselected pregnant women at delivery. *PLoS One* 2020;15:e0239887. doi:10.1371/journal.pone.0239887
89 Afshar Y, Gaw SL, Flaherman VJ, et al, Pregnancy CoRonavIrus Outcomes RegIsTrY (PRIORITY) Study. Clinical Presentation of Coronavirus Disease 2019 (COVID-19) in Pregnant and Recently Pregnant People. *Obstet Gynecol* 2020;136:1117-25. doi:10.1097/AOG.0000000000004178
90 Antoun L, Taweel NE, Ahmed I, Patni S, Honest H. Maternal COVID-19 infection, clinical characteristics, pregnancy, and neonatal outcome: A prospective cohort study. *Eur J Obstet Gynecol Reprod Biol* 2020;252:559-62. doi:10.1016/j.ejogrb.2020.07.008
91 Ahmed I, Quinn L, Tan BK. COVID-19 and the ABO blood group in pregnancy: A tale of two multiethnic cities. *Int J Lab Hematol* 2021;43:e45-7. doi:10.1111/ijlh.13355
92 Sahin D, Tanacan A, Erol SA, et al. A pandemic center's experience of managing pregnant women with COVID19 infection in Turkey: A prospective cohort study. *Int J Gynecol Obstet* 2020:ijgo.13318-ijgo.
93 Farhat AS, Sayedi SJ, Akhlaghi F, Hamedi A, Ghodsi A. Coronavirus (COVID-19) Infection in Newborns. *Int J Pediatr* 2020;8:11513-7.
94 Vivanti AJ, Mattern J, Vauloup-Fellous C, et al. Retrospective Description of Pregnant Women Infected with Severe Acute Respiratory Syndrome Coronavirus 2, France. *Emerg Infect Dis* 2020;26:2069-76. doi:10.3201/eid2609.202144
95 Secretaría de Salud. Informes epidemiológicos de embarazadas y puérperas estudiadas, ante sospecha de COVID 19 Dirección General de Epidemiología. Semana Epidemiológica 28 https://www.gob.mx/cms/uploads/attachment/file/594825/sem28.pdf.
96 Charlesworth A. Shock to the system: COVID-19's long-term impact on the NHS. 2020. https://www.health.org.uk/news- and-comment/blogs/shock-to-the-system-covid-19s-long-term-impact-on-the-nhs
97 Papoutsi E, Giannakoulis VG, Ntella V, Pappa S, Katsaounou P. Global burden of COVID-19 pandemic on healthcare workers. *ERJ Open Res* 2020;6:00195-2020. doi:10.1183/23120541.00195-2020
98 Davenport MH, Meyer S, Meah VL, Strynadka MC, Khurana R. Moms Are Not OK: COVID-19 and Maternal Mental Health. *Frontiers in Global Women's Health* 2020. https://www.frontiersin.org/articles/10.3389/fgwh.2020.00001/full
99 Moher D, Tsertsvadze A, Tricco AC, et al. When and how to update systematic reviews. *Cochrane Database Syst Rev* 2008;(1):MR000023. doi:10.1002/14651858.MR000023.pub3
100 Takwoingi Y, Hopewell S, Tovey D, Sutton AJ. A multicomponent decision tool for prioritising the updating of systematic reviews. *BMJ* 2013;347:f7191. doi:10.1136/bmj.f7191
101 Knight M, Bunch K, Vousden N, et al, UK Obstetric Surveillance System SARS-CoV-2 Infection in Pregnancy Collaborative Group. Characteristics and outcomes of pregnant women admitted to hospital with confirmed SARS-CoV-2 infection in UK: national population based cohort study. *BMJ* 2020;369:m2107. doi:10.1136/bmj.m2107
102 Public Health England. Disparities in the risk and outcomes of COVID-19. 2020 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/908434/Disparities_in_the_risk_and_outcomes_of_COVID_August_2020_update.pdf.
103 Petersen EE, Davis NL, Goodman D, et al. Racial/Ethnic Disparities in Pregnancy-Related Deaths - United States, 2007-2016. *MMWR Morb Mortal Wkly Rep* 2019;68:762-5. doi:10.15585/mmwr.mm6835a3
104 MBRRACE-UK. Saving lives, improving mothers' care. Lessons learned to inform maternity care from the UK and Ireland Confidential Enquiries into Maternal Deaths and Morbidity 2015-17. 2019 https://www.npeu.ox.ac.uk/assets/downloads/mbrrace-uk/reports/MBRRACE-UK%20Maternal%20Report%202019%20-%20WEB%20VERSION.pdf.
105 Pareek M, Bangash MN, Pareek N, et al. Ethnicity and COVID-19: an urgent public health research priority. *Lancet* 2020;395:1421-2. doi:10.1016/S0140-6736(20)30922-3
106 Docherty AB, Harrison EM, Green CA, et al, ISARIC4C investigators. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ* 2020;369:m1985. doi:10.1136/bmj.m1985
107 Guan WJ, Ni ZY, Hu Y, et al, China Medical Treatment Expert Group for Covid-19. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med* 2020;382:1708-20. doi:10.1056/NEJMoa2002032

108 Richardson S, Hirsch JS, Narasimhan M, et al, and the Northwell COVID-19 Research Consortium. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. *JAMA* 2020. doi:10.1001/jama.2020.6775
109 Garg S, Kim L, Whitaker M, et al. Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019 - COVID-NET, 14 States, March 1-30, 2020. *MMWR Morb Mortal Wkly Rep* 2020;69:458-64. doi:10.15585/mmwr.mm6915e3
110 Collin J, Byström E, Carnahan A, Ahrne M. Public Health Agency of Sweden's Brief Report: Pregnant and postpartum women with severe acute respiratory syndrome coronavirus 2 infection in intensive care in Sweden. *Acta Obstet Gynecol Scand* 2020;99:819-22. doi:10.1111/aogs.13901
111 Mendoza M, Garcia-Ruiz I, Maiz N, et al. Pre-eclampsia-like syndrome induced by severe COVID-19: a prospective observational study. *BJOG* 2020; published online 1 June.

Supplementary information: Appendices 1-9

7. 5th Article: Collaboration in times of Coronavirus: Reflections on a living systematic review of COVID-19 in pregnancy

After our LSR was published by the BMJ in September 2020, I had the opportunity to write an article for inclusion in a special Covid-19 issue in the Cochrane database of systematic reviews. I took this opportunity to discuss how we collaborated on the PregCOV19 project and the methods that we used which could be useful for other LSRs in the future. LSR´s are a very novel type of review, however in times of pandemic/ epidemic they are extremely useful since the information available is constantly being updated.

Cochrane also turned our article for the special issue into a case story to showcase on their website and this has been included in the discussion.

COVID-19 SHORT REPORT

Collaboration in times of Coronavirus: reflections on a living systematic review of COVID-19 in pregnancy

Authors: Elena Stallings[1,2,3], John Allotey[4,5], Madelon van Wely6, Shakita Thangaratinam[5,7], Javier Zamora[1,2,3,5] for the PregCOV collaboration

Participating groups: Cochrane Gynaecology and Fertility: Netherlands Satellite, WHO Collaborating Centre for Global Women's Health, University of Birmingham, Cochrane Madrid, CIBER Epidemiology and Public Health (CIBERESP), Guangzhou Women and Children's Medical Centre, WHO Human Reproduction Programme, Elizabeth Glaser Paediatric AIDS Foundation, Queen Mary University of London, St Georges University, UK

Corresponding Author: Javier *Zamora:javierza@gmail.com*

Introduction and background

Coronavirus disease 2019 (COVID-19) was officially declared a pandemic by the World Health Organization (WHO) in March 2020. COVID-19 is especially dangerous for at-risk populations, such as pregnant women,[!] thus making it critical to determine how COVID-19 affects them and their babies . A regular systematic review methodology would not be sufficient to synthesize the overwhelming amount of evidence produced daily worldwide. We needed to carry out a living systematic review (LSR), meaning that the review would be continually updated, incorporating new studies as they become available.[2]

The PregCOV-19 living systematic review working group was quickly established through an international collaboration, which included researchers and medical students at the University of Birmingham, UK; the World Health Organization; the Cochrane Gynaecology and Fertility Netherlands Satellite; and researchers in other parts of the UK, Spain (Cochrane Madrid), China, and the USA. This project commenced at the beginning of April 2020, just when the pandemic was gathering full force around Europe. At that point we were in a full lockdown, working from home, and online meetings had quickly become the new normal.

Key activities and strategies

We developed a protocol for the project that encompassed numerous clinical questions. For the LSR that we published in September 2020, we took the usual systematic review steps. We carried out rigorous searches on a weekly basis in the major medical databases for studies relating to COVID-19 in pregnant women . We screened thousands of studies (49,684) for inclusion in our review. [3] We also extracted data and carried out quality assessment on a weekly basis. At first we carried out statistical analysis every two weeks, however once we saw the results were not varying greatly, the analysis moved to monthly, and then bi-monthly.

In order to co-ordinate this process, we had weekly team meetings, which have now moved to biweekly. Throughout the review process we were also constantly adding new members to the team. The new team members go through a training process, in which they shadow other team members, for one to two weeks before carrying out the work independently.

We created a website (birmingham.ac.uk/research/who-collaborating-centre/pregcov), linked to the University of Birmingham, to highlight the project and make it easily accessible to pregnant women, researchers, and clinicians worldwide. We are currently updating the results on the website every two months.

Outcomes and impact of activities

We published the PregCOV19 project protocol on PROSPERO in April 2020.[4] The aim of this project is to assess the impact of COVID-19 in pregnancy. From this protocol we have several different studies planned. The first study was a LSR, fast-tracked by the *BMJ,* entitled 'Clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy : living systematic review and meta-analysis'. The review took only

'Hospital Universitario Ramon y Cajal (IRYCIS), Spain; ' CIBER Epidemiology and Public Health (CIBERESP), Spain; [3]Cochrane Associate Centre of Ma drid, Spain; ' University of Birmingham, UK; [5]WHO Collaborating Centre for Global Women's Health, University of Birmingham , UK; ' Nethe rlan ds Satellite of Cochrane Gynaecology and Fertility, Amsterdam University Medical Center, The Netherlands; ' Birmingham Women's and Children's NHS Foundation Trust, UK

In: Collaborating in response to COVID-19: editorial and methods initiatives across Cochrane
https://doi.org/10.1002/14651858.CD202002

five months from initiation to publication in the *BMJ,* peer-review process included, while most systematic reviews take years to complete and then be published . During public health crises, the best evidence needs to be peer-reviewed and made available to medical professionals and the public immediately. That is why a living systematic review format is the best option during times of crisis such as a pandemic .[S] Future updates with additional studies will allow us to confirm our results with a higher level of certainty.

One of our significant findings so far is that pregnant and recently pregnant women may be at increased risk of admission to an intensive care unit (ICU}.[3] This finding led to a lot of media interest, with the published LSR being reported by various news outlets worldwide, such as CNN, *The Guardian,* and Bloomberg. We hope that it will inform clinical guidelines and practice on the management of pregnant women with COVID-19.

Lessons for the future: sustainability and transferability

This project proved challenging on a few fronts. The first challenge was the pace we were working at . There were many late nights, which turned into early mornings endured by all the team. Just when we thought we were finished, the next influx of data would come rolling in and we would start the process again . It is difficult to maintain the pace that we have been working at for the past five months. During the lockdown we pushed aside other commitments and projects to dedicate nearly all our time to this project. This is not feasible moving forward. We must start dividing our time among our other various tasks again . We are also losing a lot of valuable team members as the medical students transition back to classes. So, going forward into the second wave of the virus, we must be conscious of these new limitations we face with the team .

The second challenge, albeit an opportunity, is that we were such a large group, collaborating from many different countries around all points of the globe. Frequent zoom meetings were a must, although sometimes a struggle co-ordinating different time zones . However, this also allowed us to have clinicians, statisticians, epidemiologists and students working in our team, which helped keep it diverse and see all angles of the problem.

We have created a basic framework for a large-scale LSR that can answer multiple research questions. This research framework can be deployed in the future for other public health crises or future pandemics . The way we have organized the team, each member carrying out their specific duties every week, makes all the systematic review stages tick on without fault. The data extraction sheets are organized in an Excel workbook for all the various review questions, varying from prevalence to risk factors . This format makes it easily transferable to other projects as an LSR framework in the future.

Links to additional resources

- The LSR published in the *BMJ:* doi.org/10.1136/bmj.m3320
- The PregCOV19 website: birmingham.ac.uk/research/who-collaborating-centre/pregcov/index.aspx
- LSR protocol: crd .york .ac.uk/ prospero/display_recor d. php?ID=CRD42020178076

Acknowledgements

We would like to thank all of the members of the PregCOV team for all of their hard work on this project, especially the medical students from the University of Birmingham.

Declarations of interest

Madelon Van Wely leads the Netherlands Satellite of Cochrane Gynaecology and Fertility . Elena Stallings and Javier Zamora are members of the Cochrane Associate Centre Madrid.

Funding

The project was partially funded by the World Health Organization .

References

1. Centers for Disease Control and Prevention. People at increased risk and other people who need extra precautions [updated 21July 2020]. Available at cdc.gov/coronavirus / 2019-ncov/need-extra-precautions / index.html

2. Elliott JH, Synnot A, Turner T, Simmonds M, Aki EA, McDonald S, et al. Living systematic review: 1. Introduction-the why, what, when, and how. *Journal of Clinical Epidemiology* 2017;91:23-30. htt ps:// doi.org/ 10.1016/j.jclinepi.2017.08.010

3. Allotey J, Stallings E, Bonet M, Yap M, Chatterjee S, Kew T, et al. Clinical manifestations, risk factors , and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy: living systematic review and meta-analysis. *BMJ* 2020;370:m3320. http s:// doi. org/10.1136/ bmj.m3320

4. Allotey J, Bonet M, Stallings E, Yap M, Chatterjee S, Kew T, el al. COVID-19 in pregnant women: a living systematic review on prevalence, presentatio n, prognosis and treat ment. PROSPERO 2020 CRD42020178076 Available from: https:// www.crd.york.ac.uk /prospero/ display_record.php?ID=CRD42020178076

5. Turner T, McDonald S. Cochrane COVID-19 living systematic reviews: why, how, and what has been done? In: Collaborating in response to COVID-19: Editorial and methods initiatives across Cochrane. *Cochrane Database of Systematic Reviews* 2020;(11 Suppl 1). http s:// doi. org/10.1002/14651858.CD202002

In: Collaborating in response to COVID-19: editorial and methods initiatives across Cochrane
https://doi.org/10.1002/14651858.CD202002

8. Discussion

8.1 Main findings

8.1.1 Systematic review of sex as a prognostic factor in patients with pulmonary embolism

There are not many systematic reviews of prognostic factors and even fewer Cochrane reviews (6 Cochrane PF reviews in total). Thus, it is a novel area of research with new methods still being developed and evaluated. Our Cochrane review of sex as a prognostic factor in patients with pulmonary embolism is still ongoing, so I have included preliminary results in this thesis. During the pilot stage of the screening on title and abstract, we calculated the Kappa score of each pair of reviewers to ensure consistency. The reviewers had a good agreement percentage that ranged between 83-100%, with most reviewers falling around 96%.

At the time of writing this thesis, we have identified 5 studies that can be included in the review: Agarwal 2015, Barrios 2017, Borrero 2007, Feng 2020, Rosovsky 2019 (31, 68-71). We included these 5 studies in a preliminary meta-analysis studying the outcome of all cause 30- day mortality. The five studies had adjusted for different covariates including age, race, smoking status, comorbidities and baseline characteristics. The most commonly adjusted for covariates were age and race. We found that overall, there was no association between sex and mortality at 30 days (OR 0.98; CI95% 0.81-1.16).

I have included the list of studies, tables of study characteristics, the Kappa results from the pilot and a preliminary PRISMA flowchart in Annex 4.

8.1.2 Development and evaluation of a search filter to identify prognostic factor studies in Ovid MEDLINE.

After carrying out two PF systematic reviews last year and having carried out the preliminary searches for the Cochrane PF review, I realised the need for the development of a PF search filter. I created a search filter that was highly sensitive in capturing PF studies. The overall sensitivity of the filter was calculated to be 95%, while the overall specificity was 41%. The precision of the filter varied considerably, ranging from 0.36 to 17%. The NNR (number

needed to read) value depends on the total number of hits in the search and varied largely from 6 to 278. We compared it with the Haynes clinical queries broad prognosis filter and our filter performed better in the relative recall of the studies. However, our filter had a low specificity and the Haynes filter performed better than ours in this domain.

Our aim of the project was to create a filter that was highly sensitive as we did not want to risk losing any relevant studies. At the same time, if we could reduce the NNR it would be very beneficial as PF systematic review searches generally retrieve a lot of varied studies due to the poor indexing of PF studies. Our filter had and acceptable performance measures and can be used in systematic reviews of PF studies. Using this filter could save as much as 40% of the title and abstract screening task. In the future, the specificity of the filter could be improved by defining additional terms to be included, although it is important to evaluate any modification to guarantee the filter is still highly sensitive.

8.1.3 Clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy: living systematic review and meta-analysis.

The research team started work on the review in April 2020. For the original review we collated data from published articles between December 2019 and June 2020. Each week, we searched for newly published data, and carried out the screening, data extraction and ROB assessment weekly. For the first update, we added new studies published until October 2020. The review now includes data from 192 studies and 29 different countries. Of these studies, 115 were new editions to the latest update. The findings remained consistent between the update and original review. However, the update helped us to narrow down the confidence intervals with more precise values for many of the outcomes.

Overall, 10% (73 studies, 67 271 women) of pregnant and recently pregnant women attending or admitted to hospital for any reason were diagnosed as having suspected or confirmed covid-19. Pregnant women are more likely than nonpregnant women to have an asymptomatic SARS-CoV-2 infection. However, the reason for this might be related to the testing strategy. Pregnant women regularly have Covid-19 testing when they attend their hospital appointments or when they arrive at the hospital to give birth, whilst non-pregnant

women of the same age are only likely to have testing if they experience symptoms. The true prevalence of Covid-19 in pregnant women could in fact be lower if all pregnant women, including those not attending a hospital are included.

In pregnant women with covid-19, increased maternal age, high body mass index, non-white ethnicity, any pre-existing maternal comorbidity including chronic hypertension and diabetes, and pre-eclampsia were associated with serious complications such as admission to an intensive care unit, invasive ventilation and maternal death. The odds of admission to an intensive care unit (odds ratio 2.13, 1.53 to 2.95; I2=71.2%) and invasive ventilation (2.59, 2.28 to 2.94; I2=0%) were higher in pregnant and recently pregnant than non-pregnant reproductive aged women. Overall, 339 pregnant women (0.02%, 59 studies, 41,664 women) with confirmed covid-19 died from any cause.

Compared to pregnant women without covid-19, those with the disease had increased odds of preterm birth (1.47, 1.14 to 1.91; I2=18.6%). However, these premature births are likely to be the result of medical decisions to induce early delivery in those with COVID-19.

Due to all of the above factors, based on our findings, pregnant women should be considered a high-risk group, particularly those identified to have risk factors for severe COVID-19. Therefore, pregnant women should be a high priority on the Covid-19 vaccine list as the outcomes observed from severe Covid-19 are potentially worse than vaccine side effects, that still require further study.

8.1.4 Cochrane collaboration PregCOV19: what's new, what we did, results

LSRs are a relatively new type of SR which help to summarise new evidence as it emerges, which is important in a pandemic. We published an article in Cochrane on how we collaborated internationally and conducted our LSR during the pandemic. The editors at Cochrane then decided to turn our editorial into a case story, publishing it on the Cochrane website with graphics, thus making it easily accessible and readable for all. The full case story is located in Annex 3.

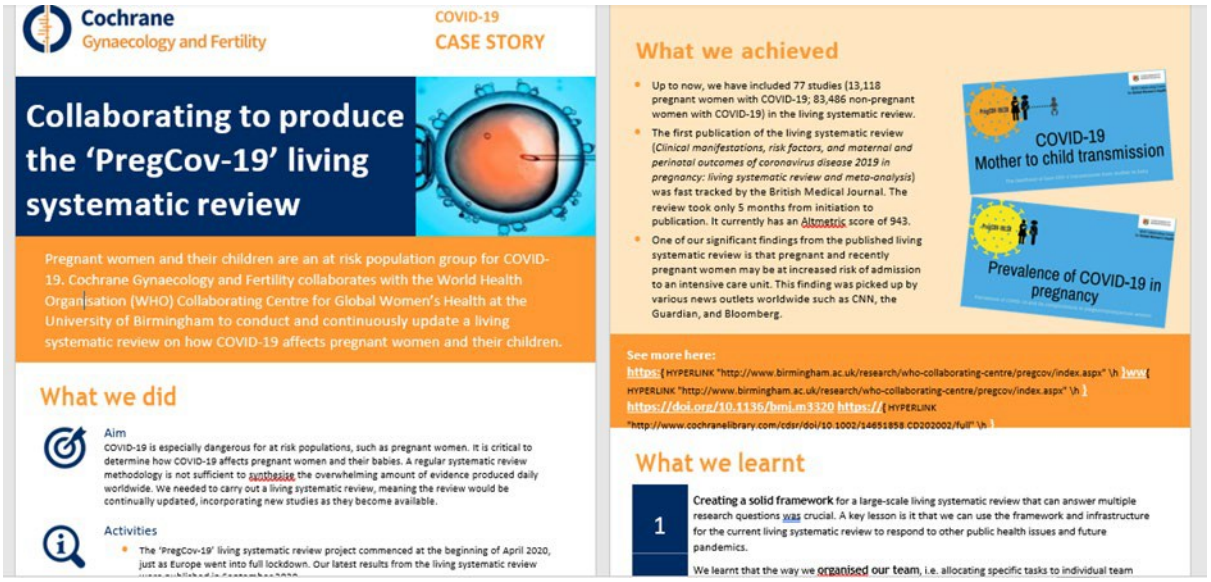


Figure 5: Case story published by Cochrane

8.2 Difficulties in personalizing medicine by sex

There are many difficulties to overcome when it comes to personalizing medicine by sex. These difficulties include the rapidly evolving vocabulary, social norms, financing of studies, and patient privacy issues.

Nowadays when we talk about sex and gender it's not just those two terms that we must be aware of, but also transgender, intersex, non-binary, asexual, genderqueer, among many more (72, 73). This is an area that is constantly evolving and difficult for clinicians and other healthcare staff to keep up to date on.

The next obstacle in personalizing medicine by sex is the financing of studies. Funding agencies have tried to instigate policies to include male and female animals and human participants in grant proposals but unfortunately they lack mechanisms to hold recipients accountable (74, 75). Journals have also implemented policies to encourage sex and gender-based analysis, but it is still often not included. Pharmaceutical companies are reluctant to carry out clinical trials on both sexes of animals as it increases costs. They also are very reluctant to carry out clinical trials on pregnant women as they are a high-risk population.

Finally, precision medicine carries a risk in relation to patient privacy, whereby patients could be disadvantaged with respect to medical insurance coverage. For example, if one sex has a poorer prognosis of pulmonary embolism and so are given specific drugs which cost more it could mean a higher cost of insurance. Likewise, it could be possible that with the rise of genetic testing insurance companies may not offer certain policies to those sexes with a genetic predisposition of a disease.

8.3 New trends in sex specific medicine

Within many fields of precision medicine, sex specific issues are only beginning to be explored (76). However, it has been acknowledged that in areas such as cardiovascular disease, Alzheimer's, and cancer there is an opportunity to use molecular data to manage male and female patients more accurately.

Research is now being carried out on the impact of sex hormones on pharmacological therapies (77). At the moment very few drugs have sex specific labelling or dosing recommendations, even though different doses could be more effective for males versus females and also could minimise adverse reactions. This will hopefully change in the coming years, with new strategies to improve precision medicine. One of these strategies is theragnostics which is an approach derived from combining therapeutics and diagnostics (97). It is a relatively new field that can be useful in personalizing medicine. It associates the most appropriate diagnostic test to identify patients most likely to be helped or harmed by an intervention with a targeted drug therapy (97).

Education reforms to include sex and gender teaching in medical school are on the way. In 2018 ´The Sex and Gender Health Education (SGHE)´ summit was held at the University of Utah. This summit was a national collaboration of educational leaders from various health professions that came together to improve health curricula by integrating sex and gender-based evidence into education (78). Due to this integration, the importance of sex and gender to healthcare workers knowledge and practice is starting to be recognized. More education on the utility of sex for prognosis will ultimately ensure more personalized healthcare and improve patient outcomes (78, 79).

Lastly, a growing number of funding agencies and journal publishers are now explicitly calling for and requiring that sex and gender are taken into consideration in research in funding applications and in the presentation of research findings. Thus, researchers are becoming obliged to consider patient/animal sex in their studies. They are also encouraged to use the Sex and Gender Equity in Research guidelines for the reporting of sex and gender information in study design, data analyses, results and interpretation.

8.4 Implications in practice

Covid-19 in pregnancy

Healthcare professionals need to be aware that pregnant women with covid-19 may show fewer symptoms that non-pregnant women of reproductive age or the rest of the population in general. Pregnant women should also be warned of the increased risks of severe covid-19 during pregnancy so that they can take the appropriate safety measures to reduce their risk of infection.

Our systematic review and meta-analysis has had a significant impact worldwide with 315 citations in only 9 months of publication. It has been cited in clinical studies, systematic reviews, and guidelines for dealing with covid-19 in pregnancy.

Sex as PF in pulmonary embolism

Depending on the final results of our systematic review and meta-analysis for sex as a PF in pulmonary embolism, there is the possibility to improve mortality outcomes for either males or females with pulmonary embolism (PE) by allowing sex to be included in models for prognosis stratification. If sex is determined to be a PF for PE it could also help in defining modifiable targets for interventions or treatments and to determine predictors of differential treatment response. Identifying PFs is a critical step in the pathway towards personalized medicine.

8.5 Implications in research

Cochrane collaboration

The framework that we set up to carry out the LSR was very robust, with data extraction templates created and protocols put in place for carrying each step of the review. Therefore, this framework could be very easily used with LSRs involving other topics in clinical medicine. We also incorporated medical students into the project, training them on all of the different sections of the LSR from screening of studies to data extraction and quality assessment. Thus, these students have learnt a lot about evidence based research and can now go on to work on other reviews.

Sex as a prognostic factor in pulmonary embolism

This review will provide essential information on whether sex differences are important in the development and outcome of PE. If true, this will guide pharmaceutical researchers in drug discovery and development, leading to more personalised treatment regimens. It could help to focus drugs specifically for one sex, depending on the likelihood of the poor outcomes (mortality) that we find in our review. Due to the large number of references retrieved from the review search strategy, we started a second project involving the development of a search filter for prognostic factor studies. This will help other reviewers in the future who are carrying out prognostic factor systematic reviews.

Filter for prognostic factor studies

Our PF filter reduced the number of studies retrieved from systematic review searches, without losing relevant studies, therefore it had a high sensitivity (95%) and was capable of reducing the number of studies requiring manual screening. It is expected that other systematic reviewers will be able to use and take advantage of this filter in their future research.

In general

Overall, I hope that this thesis highlights the need for better use of the sex and gender terminology in research. Too often, the terms sex and gender are used interchangeably and incorrectly which can in turn be misleading and confusing for researchers. I also hope that it encourages more research in the various areas of sex specific research, whether it be for

females or males, each group deserves to benefit from personalized medicine that will be realised through further research.

8.6 Strengths and limitations of the thesis research

8.6.1 Strengths

Covid-19 in pregnancy

During the Covid-19 pandemic there were many systematic reviews carried out of poor quality. Not all reviews followed reporting guidelines, such as PRISMA (49, 80). Systematic reviews included duplicate data, which in turn affected the validity of the findings. This is due to many primary studies and case series reporting on a subset of data without acknowledging that they were published elsewhere. Meta-analyses were performed in reviews inappropriately by pooling the data from case reports and case series, thereby biasing the estimates on prevalence, and rates of complications. These quickly published reviews might be due to the "Publish or perish" phenomenon. "Publish or perish" is an aphorism that describes the pressure to publish academic articles in order to succeed in an academic research career.

On the other hand, our SR followed the PRISMA guidelines, ensuring that it had the lowest risk of bias possible. In fact, in a review of reviews published in early 2021, our paper was the only review that had a low risk of bias (81). We only included cohort studies in our meta-analysis to minimize the ROB. In order to reduce the risk of missing relevant studies, we created a comprehensive search strategy, searching in 7 different databases as well as preprint servers and blogs. We also contacted study authors to obtain unpublished data.

Our PREGCov19 team is multidisciplinary with clinicians, medical students, methodologists, and statisticians coming together to form this project. We collaborated internationally with many different research groups such as the WHO, THE Cochrane Centre and the EPPI centre. We used robust methods and continually updated the review at regular intervals as new evidence emerged daily.

Sex as a prognostic factor in pulmonary embolism

In our systematic review and protocol, we employed a robust search strategy so we can be confident that we will not miss any relevant studies. After eliminating the duplicates, our search strategy retrieved approximately 78,000 references. With this many references it was necessary to coordinate a large team of 14 reviewers to screen through the titles and

abstracts. We also manually screened through conference papers searching for any relevant studies. As it is a Cochrane protocol, it has been peer reviewed both externally and by the Cochrane editors at least twice before finally being published. Due to this, the methods are as up to date and correct as possible.

Filter development

The development of the PF filter used various methods from different authors. In this way, we were able to choose the methods that were best suited for our limited resources and we can advise future filter developers to use similar methods. Our filter had an excellent sensitivity, meaning barely any studies escaped during the search. This is a critical consideration for filter development as it ensures the maximum capture of relevant studies. When we compared our filter to the Haynes clinical queries broad prognosis filter, our PF filter had a higher sensitivity, which should be expected as it is more specific to prognostic factor studies.

8.6.2 Limitations

Covid-19 in pregnancy

We faced several limitations in our review. Many studies included women with both suspected and confirmed covid-19 and the majority only reported on pregnant women in hospital settings, therefore affecting the generalisability of the studies. Similarly, the studies included in the analysis did not use the same methods to collect and gather data. Due to this, we tried to carry out subgroup analysis when possible, however this was not always the case.

Sex as a prognostic factor in pulmonary embolism

We encountered a difficult situation of retrieving a very large number of references from our searches that required manual screening (78,000 after duplicates were removed). To make the screening process more effective, we decided to do a pre-screening by two senior reviewers, where they eliminated specific irrelevant study types such as case reports, clinical trials, animal studies, laboratory studies or studies that were focused on only one sex. This process was carried out by the two reviewers independently and not in pairs, so it could be

criticised that a relevant study could remain undetected. However, it helped us rapidly exclude approximately 20,000 studies from the reference set that we needed to screen.

Filter development

Our filter produced a low specificity and precision; however, this is understandable given the relatively small number of prognostic factor studies in the reference sets. Also, our methods of developing the filter were restricted due to our limited resources available. A more robust method might be to use a true gold standard (created from hand searching for studies) rather than relative recall, however with limited resources it was the best option.

8.7 Future research

Much more research is clearly needed in the area of sex specific medicine. Within the scope of my thesis, there are new and exciting projects ongoing at the moment or planned for the future.

Covid-19 in pregnancy

We expect the sample size of our meta- analysis to continue increasing in the next few months and will shortly publish the second update to our original LSR in the BMJ as soon as we have collected sufficient data. A methodological problem that we have encountered in this review and something that we could look deeper into in the future is deciding when to stop updating a living systematic review. Research has been carried out on stopping updates in a review using clinical trial data, however there is no guidance yet on how and when to stop updating a review with cohorts of patients investigating risk factors.

At the moment, we have two other reviews planned, one on mother to child transmission of covid-19 in pregnancy and the other on maternal mortality in covid-19 and pregnancy. These projects are both already almost completed and on track to be published soon. As well as these reviews, it would also be of interest to start researching the new variants of covid-19 coming from Brazil, the UK, and South Africa to see if these variants affect outcomes in pregnancy differently. Lastly, as we start to see covid-19 vaccines being rolled out worldwide it would be beneficial to start collecting data on vaccine uptake, side effects etc in pregnant women to see if they are as effective for this population as the general population.

Sex as a prognostic factor in pulmonary embolism

Once we finish the systematic review on sex as a prognostic factor in patients with pulmonary embolism, if we identify any studies as having studied gender instead of sex, it would be of interest to carry out a review examining the role of gender as a PF in PE.

Enhancing the filter for prognostic studies

If more resources become available, I would like to start the filter development project again, but this time employ different methods such as using a true gold standard instead of relative recall. Using a true gold standard would require manual searching of journals to find

prognostic factor studies. Availability of human resources was the main restricting factor for not going by these methods before. It would be interesting to apply the filter (made for Ovid Medline) in PubMed and Embase to see if it performs the same in terms of sensitivity and specificity in these databases. I would also like to apply the filter to our “sex as prognostic factor in pulmonary embolism” review search strategy and compare results.

9. Conclusions

Sex should play a central role in personalized medical care. Although more studies providing sex specific evidence are being carried out, there still remains significant shortcomings in effectively implementing sex specific health care. This thesis aims to highlight the level of awareness and the role that sex plays in healthcare and health research. The main conclusions derived from the research are set out below:

- Pregnant women are considered to be a high-risk group for covid-19 infection and those with the infection need extra medical attention as they are more likely to experience preterm birth or be admitted to the intensive care unit.

- Preliminary results of the systematic review of sex as a prognostic factor in patients with pulmonary embolism show that there is no association between sex and mortality. However, we have still yet to rate the quality of the evidence, so more higher quality studies may be needed to confirm this.

- The development of a PF search filter that is both sensitive and specific is not an easy task due to the poor indexing of PF studies in databases. Our filter yielded a sensitivity of 95% and a specificity of 41%, so more research is needed to increase the filters specificity.

- When carrying out a systematic review investigating sex as a PF, researchers must keep in mind that various adaptations must be made to the review process. These adaptations include modifying the PF section of the data extraction template, adjusting certain sections of QUIPS (after ROB) and extracting data on the sex and gender terms used throughout the included studies.

10. Bibliography

1. National Research Council Committee on AFfDaNToD. The National Academies Collection: Reports funded by National Institutes of Health. Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease. Washington (DC): National Academies Press (US) Copyright © 2011, National Academy of Sciences.; 2011.

2. Carrasco-Ramiro F, Peiró-Pastor R, Aguado B. Human genomics projects and precision medicine. Gene Ther. 2017;24(9):551-61.
3. Weil AR. Precision Medicine. Health Aff (Millwood). 2018;37(5):687.
4. Cheng L, Lopez-Beltran A, Massari F, MacLennan GT, Montironi R. Molecular testing for BRAF mutations to inform melanoma treatment decisions: a move toward precision medicine. Mod Pathol. 2018;31(1):24-38.
5. Naito Y, Urasaki T. Precision medicine in breast cancer. Chin Clin Oncol. 2018;7(3):29.
6. Deluche E, Onesti E, Andre F. Precision medicine for metastatic breast cancer. Am Soc Clin Oncol Educ Book. 2015:e2-7.
7. Kraya AA, Maxwell KN, Wubbenhorst B, Wenz BM, Pluta J, Rech AJ, et al. Genomic Signatures Predict the Immunogenicity of BRCA-Deficient Breast Cancer. Clin Cancer Res. 2019;25(14):4363-74.
8. Horton RH, Lucassen AM. Recent developments in genetic/genomic medicine. Clin Sci (Lond). 2019;133(5):697-708.
9. Ghosh R, Tabrizi SJ. Huntington disease. Handb Clin Neurol. 2018;147:255-78.
10. Bartz D, Chitnis T, Kaiser UB, Rich-Edwards JW, Rexrode KM, Pennell PB, et al. Clinical Advances in Sex- and Gender-Informed Medicine to Improve the Health of All: A Review. JAMA Intern Med. 2020;180(4):574-83.
11. McCarthy M. Social determinants and inequalities in urban health. Rev Environ Health. 2000;15(1-2):97-108.
12. Bircher J, Kuruvilla S. Defining health by addressing individual, social, and environmental determinants: new opportunities for health care and public health. J Public Health Policy. 2014;35(3):363-86.
13. Singh A, Hussain S, Najmi AK. Role of Cochrane Reviews in informing US private payers' policies. J Evid Based Med. 2017;10(4):293-331.
14. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes. Public Health Rep. 2014;129 Suppl 2(Suppl 2):19-31.
15. King BA, Jama AO, Marynak KL, Promoff GR. Attitudes Toward Raising the Minimum Age of Sale for Tobacco Among U.S. Adults. Am J Prev Med. 2015;49(4):583-8.
16. Lee YH, Lee YM, Sung TY, Yoon JH, Song DE, Kim TY, et al. Is Male Gender a Prognostic Factor for Papillary Thyroid Microcarcinoma? Ann Surg Oncol. 2017;24(7):1958-64.
17. Udry JR. The nature of gender. Demography. 1994;31(4):561-73.
18. Vaidakis N. Conceptual controversies regarding the terms Gender and Sex. Psychiatriki. 2020;31(3):271-4.
19. Madsen TE, Simmons J, Choo EK, Portelli D, McGregor AJ, Napoli AM. The DISPARITY Study: do gender differences exist in Surviving Sepsis Campaign resuscitation bundle

completion, completion of individual bundle elements, or sepsis mortality? J Crit Care. 2014;29(3):473.e7-11.
20. Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and Gender Equity in Research: rationale for the SAGER guidelines and recommended use. Res Integr Peer Rev. 2016;1:2.
21. Heidari S, Bachelet VC. Sex and gender analysis for better science and health equity. Lancet. 2018;392(10157):1500-2.
22. Arnold AP, Chen X, Itoh Y. What a difference an X or Y makes: sex chromosomes, gene dose, and epigenetics in sexual differentiation. Handb Exp Pharmacol. 2012(214):67-88.
23. Arnold AP. A general theory of sexual differentiation. J Neurosci Res. 2017;95(1-2):291-300.
24. Wang F, Shu X, Meszoely I, Pal T, Mayer IA, Yu Z, et al. Overall Mortality After Diagnosis of Breast Cancer in Men vs Women. JAMA Oncol. 2019;5(11):1589-96.
25. Miao H, Verkooijen HM, Chia KS, Bouchardy C, Pukkala E, Larønningen S, et al. Incidence and outcome of male breast cancer: an international population-based study. J Clin Oncol. 2011;29(33):4381-6.
26. Greif JM, Pezzi CM, Klimberg VS, Bailey L, Zuraek M. Gender differences in breast cancer: analysis of 13,000 breast cancers in men from the National Cancer Data Base. Ann Surg Oncol. 2012;19(10):3199-204.
27. Li X, Yang J, Krishnamurti U, Huo L, Ward KC, O'Regan R, et al. Hormone Receptor-Positive Breast Cancer Has a Worse Prognosis in Male Than in Female Patients. Clin Breast Cancer. 2017;17(5):356-66.
28. Alswat KA. Gender Disparities in Osteoporosis. J Clin Med Res. 2017;9(5):382-7.
29. NIH consensus group. Osteoporosis prevention, diagnosis, and therapy. Jama. 2001;285(6):785-95.
30. Keller K, Rappold L, Gerhold-Ay A, Hobohm L, Hasenfuß G, Konstantinides SV, et al. Sex-specific differences in pulmonary embolism. Thromb Res. 2019;178:173-81.
31. Agarwal S, Clark D, 3rd, Sud K, Jaber WA, Cho L, Menon V. Gender Disparities in Outcomes and Resource Utilization for Acute Pulmonary Embolism Hospitalizations in the United States. Am J Cardiol. 2015;116(8):1270-6.
32. Shiraev TP, Omari A, Rushworth RL. Trends in pulmonary embolism morbidity and mortality in Australia. Thromb Res. 2013;132(1):19-25.
33. Kantarci K, Morrow MM, Miller VM. Incorporating Sex as a Biological Variable into Clinical and Translational Research Training. J Womens Health (Larchmt). 2020;29(6):865-7.
34. Ritz SA, Antle DM, Côté J, Deroy K, Fraleigh N, Messing K, et al. First steps for integrating sex and gender considerations into basic experimental biomedical research. Faseb j. 2014;28(1):4-13.
35. Wald C, Wu C. Biomedical research. Of mice and women: the bias in animal models. Science. 327. United States2010. p. 1571-2.
36. Raz L, Miller VM. Considerations of sex and gender differences in preclinical and clinical trials. Handb Exp Pharmacol. 2012(214):127-47.
37. Whitehead CL, Walker SP. Consider pregnancy in COVID-19 therapeutic drug and vaccine trials. Lancet. 2020;395(10237):e92.
38. Gopalakrishnan S, Ganeshkumar P. Systematic Reviews and Meta-analysis: Understanding the Best Evidence in Primary Healthcare. J Family Med Prim Care. 2013;2(1):9-14.

39. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gülmezoglu AM, et al. How to increase value and reduce waste when research priorities are set. Lancet. 2014;383(9912):156-65.
40. Ferreira González I, Urrútia G, Alonso-Coello P. Systematic reviews and meta-analysis: scientific rationale and interpretation. Rev Esp Cardiol. 2011;64(8):688-96.
41. Higgins JPT, Lasserson T, Chandler J, Tovey D, Churchill R. Methodological Expectations of Cochrane Intervention Reviews.
Accessed: http://methods.cochrane.org/sites/default/files/public/uploads/mecir_printed_booklet_final.pdf. London: Cochrane; 2017.
42. JPT. H, J. T, J. C, M. C, T. L, Page M, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane.; 2021. Available from: www.training.cochrane.org/handbook.
43. Chalmers I. The James Lind Initiative. J R Soc Med. 962003. p. 575-6.
44. Shah HM, Chung KC. Archie Cochrane and his vision for evidence-based medicine. Plast Reconstr Surg. 2009;124(3):982-8.
45. Hill GB. Archie Cochrane and his legacy. An internal challenge to physicians' autonomy? J Clin Epidemiol. 2000;53(12):1189-92.
46. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4(1):1.
47. Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med. 2013;158(4):280-6.
48. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med. 2019;170(1):51-8.
49. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med. 2009;151(4):W65-94.
50. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. Bmj. 2013;346:e5595.
51. Hingorani AD, Windt DA, Riley RD, Abrams K, Moons KG, Steyerberg EW, et al. Prognosis research strategy (PROGRESS) 4: stratified medicine research. Bmj. 2013;346:e5793.
52. Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. PLoS Med. 2013;10(2):e1001380.
53. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med. 2013;10(2):e1001381.
54. Riley RD, Moons KGM, Snell KIE, Ensor J, Hooft L, Altman DG, et al. A guide to systematic review and meta-analysis of prognostic factor studies. Bmj. 2019;364:k4597.
55. Khangura S, Polisena J, Clifford TJ, Farrah K, Kamel C. Rapid review: an emerging approach to evidence synthesis in health technology assessment. Int J Technol Assess Health Care. 2014;30(1):20-7.

56. Tricco AC, Tetzlaff J, Sampson M, Fergusson D, Cogo E, Horsley T, et al. Few systematic reviews exist documenting the extent of bias: a systematic review. J Clin Epidemiol. 2008;61(5):422-34.
57. Macdonald H, Loder E, Abbasi K. Living systematic reviews at The BMJ. Bmj. 2020;370:m2925.
58. Lefebvre C, Manheimer E, Glanville J. Chapter 6: Searching for studies. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from wwwcochrane-handbookorg.
59. Jenkins M. Evaluation of methodological search filters--a review. Health Info Libr J. 2004;21(3):148-63.
60. Beynon R, Leeflang MM, McDonald S, Eisinga A, Mitchell RL, Whiting P, et al. Search strategies to identify diagnostic accuracy studies in MEDLINE and EMBASE. Cochrane Database Syst Rev. 2013(9):Mr000022.
61. Glanville J, Bayliss S, Booth A, Dundar Y, Fernandes H, Fleeman ND, et al. So many filters, so little time: the development of a search filter appraisal checklist. J Med Libr Assoc. 2008;96(4):356-61.
62. Johnson JL, Greaves L, Repta R. Better science with sex and gender: Facilitating the use of a sex and gender-based analysis in health research. Int J Equity Health. 2009;8:14.
63. Johnson J, Sharman Z, Vissandjée B, Stewart DE. Does a change in health research funding policy related to the integration of sex and gender have an impact? PLoS One. 2014;9(6):e99900.
64. Petkovic J, Trawin J, Dewidar O, Yoganathan M, Tugwell P, Welch V. Sex/gender reporting and analysis in Campbell and Cochrane systematic reviews: a cross-sectional methods study. Syst Rev. 2018;7(1):113.
65. López-Alcalde J, Stallings E, Cabir Nunes S, Fernández Chávez A, Daheron M, Bonfill Cosp X, et al. Consideration of sex and gender in Cochrane reviews of interventions for preventing healthcare-associated infections: a methodology study. BMC Health Serv Res. 2019;19(1):169.
66. Heidari S, Babor TF, Castro PD, Tort S, Curno M. Sex and Gender Equity in Research: rationale for the SAGER guidelines and recommended use. Research Integrity and Peer Review. 2016;1(1):2.
67. Clayton JA, Tannenbaum C. Reporting Sex, Gender, or Both in Clinical Research? Jama. 2016;316(18):1863-4.
68. Barrios D, Morillo R, Guerassimova I, Barbero E, Escobar-Morreale H, Cohen AT, et al. Sex differences in the characteristics and short-term prognosis of patients presenting with acute symptomatic pulmonary embolism. PLoS One. 122017.
69. Feng J, Rosovsky R. SEX DIFFERENCES IN PULMONARY EMBOLISM: CLINICAL CHARACTERISTICS, IN-HOSPITAL MORTALITY, AND 30-DAY READMISSIONS. Journal of the American College of Cardiology. 2020;75(11_Supplement_1):2258-.
70. Rosovsky RP, Elgendy IY, Cannegieter SC, Huisman MV, Jimenez D, del Toro J, et al. Sex Differences in Risk Factors, Clinical Presentation, Treatment and Outcomes of Patients Presenting with Acute Pulmonary Embolism. Blood. 2019;134(Supplement_1):2429-.
71. Borrero S, Aujesky D, Stone RA, Geng M, Fine MJ, Ibrahim SA. Gender differences in 30-day mortality for patients hospitalized with acute pulmonary embolism. J Womens Health (Larchmt). 2007;16(8):1165-70.

72. Monro S. Non-binary and genderqueer: An overview of the field. Int J Transgend. 202019. p. 126-31.
73. James HA, Chang AY, Imhof RL, Sahoo A, Montenegro MM, Imhof NR, et al. A community-based study of demographics, medical and psychiatric conditions, and gender dysphoria/incongruence treatment in transgender/gender diverse individuals. Biol Sex Differ. 2020;11(1):55.
74. Shansky RM, Murphy AZ. Considering sex as a biological variable will require a global shift in science culture. Nat Neurosci. 2021;24(4):457-64.
75. Duchesne A, Tannenbaum C, Einstein G. Funding agency mechanisms to increase sex and gender analysis. Lancet. 2017;389(10070):699.
76. Currie G, Delles C. Precision Medicine and Personalized Medicine in Cardiovascular Disease. Adv Exp Med Biol. 2018;1065:589-605.
77. Moyer AM, Matey ET, Miller VM. Individualized medicine: Sex, hormones, genetics, and adverse drug reactions. Pharmacol Res Perspect. 2019;7(6):e00541.
78. McGregor AJ, Chin EL, Rojek MK, Digre KB, Lopez AM, Jenkins K, et al. Sex and Gender Health Education Summit: Advancing Curricula Through a Multidisciplinary Lens. J Womens Health (Larchmt). 282019. p. 1728-36.
79. Kling JM, Rose SH, Kransdorf LN, Viggiano TR, Miller VM. Evaluation of sex- and gender-based medicine training in post-graduate medical education: a cross-sectional survey study. Biol Sex Differ. 2016;7(Suppl 1):38.
80. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009;62(10):1006-12.
81. Vergara-Merino L, Meza N, Couve-Pérez C, Carrasco C, Ortiz-Muñoz L, Madrid E, et al. Maternal and perinatal outcomes related to COVID-19 and pregnancy: An overview of systematic reviews. Acta Obstet Gynecol Scand. 2021.
82. Kastner M, Wilczynski NL, McKibbon AK, Garg AX, Haynes RB. Diagnostic test systematic reviews: bibliographic search filters ("Clinical Queries") for diagnostic accuracy studies perform well. J Clin Epidemiol. 2009;62(9):974-81.
83. Sampson M, Zhang L, Morrison A, Barrowman NJ, Clifford TJ, Platt RW, et al. An alternative to the hand searching gold standard: validating methodological search filters using relative recall. BMC Med Res Methodol. 2006;6:33.
84. Ritchie G, Glanville J, Lefebvre C. Do published search filters to identify diagnostic test accuracy studies perform adequately? Health Info Libr J. 2007;24(3):188-92.

11. Annexes

11.1 Clinical manifestations original

FAST TRACK

Clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy: living systematic review and meta-analysis

John Allotey,[1,2] Elena Stallings,[3,4] Mercedes Bonet,[5] Magnus Yap,[6] Shaunak Chatterjee,[6] Tania Kew,[6] Luke Debenham,[6] Anna Clavé Llavall,[6] Anushka Dixit,[6] Dengyi Zhou,[6] Rishab Balaji,[6] Siang Ing Lee,[1] Xiu Qiu,[7,8,9] Mingyang Yuan,[1,7] Dyuti Coomar,[1] Madelon van Wely,[10] Elizabeth van Leeuwen,[11] Elena Kostova,[10] Heinke Kunst,[12,13] Asma Khalil,[14] Simon Tiberi,[12,13] Vanessa Brizuela,[5] Nathalie Broutet,[5] Edna Kara,[3] Caron Rahn Kim,[5] Anna Thorson,[5] Olufemi T Oladapo,[5] Lynne Mofenson,[15] Javier Zamora,[3,4,16] Shakila Thangaratinam,[2,17] for PregCOV-19 Living Systematic Review Consortium

For numbered affiliations see end of the article.

Correspondence to:
S Thangaratinam
s.thangaratinam.1@bham.ac.uk
(or @thangaratinam on Twitter:
ORCID 0000-0002-4254-460X)

Additional material is published online only. To view please visit the journal online.

cite this as: *BMJ* 2020;370:m3320

http://dx.doi.org/10.1136/bmj.m3320

Accepted: 23 August 2020

AbstrAct

Objective

To determine the clinical manifestations, risk factors, and maternal and perinatal outcomes in pregnant and recently pregnant women with suspected or confirmed coronavirus disease 2019 (covid-19).

Design

Living systematic review and meta-analysis.

Data sOurces

Medline, Embase, Cochrane database, WHO COVID-19 database, China National Knowledge Infrastructure (CNKI), and Wanfang databases from 1 December 2019 to 26 June 2020, along with preprint servers, social media, and reference lists.

stuDy selectiOn

Cohort studies reporting the rates, clinical manifestations (symptoms, laboratory and radiological findings), risk factors, and maternal and perinatal outcomes in pregnant and recently pregnant women with suspected or confirmed covid-19.

Data extractiOn

At least two researchers independently extracted the data and assessed study quality. Random effects meta-analysis was performed, with estimates pooled as odds ratios and proportions with 95% confidence intervals. All analyses will be updated regularly.

results

77 studies were included. Overall, 10% (95% confidence interval 7% to14%; 28 studies, 11 432 women) of pregnant and recently pregnant women attending or admitted to hospital for any reason were diagnosed as having suspected or confirmed covid-19. The most common clinical manifestations of covid-19 in pregnancy were fever (40%) and cough (39%). Compared with non-pregnant women of reproductive age, pregnant and recently pregnant women with covid-19 were less likely to report symptoms of fever (odds ratio 0.43, 95% confidence interval 0.22 to 0.85; I^2=74%; 5 studies; 80 521 women) and myalgia (0.48, 0.45 to 0.51; I^2=0%; 3 studies; 80 409 women) and were more likely to need admission to an intensive care unit (1.62, 1.33 to 1.96; I^2=0%) and invasive ventilation (1.88, 1.36 to 2.60; I^2=0%; 4 studies, 91 606 women). 73 pregnant women (0.1%, 26 studies, 11 580 women) with confirmed covid-19 died from any cause. Increased maternal age (1.78, 1.25 to 2.55; I^2=9%; 4 studies; 1058 women), high body mass index (2.38, 1.67 to 3.39; I^2=0%; 3 studies; 877 women), chronic hypertension (2.0, 1.14 to 3.48; I^2=0%; 2 studies; 858 women), and pre-existing diabetes (2.51, 1.31 to 4.80; I^2=12%; 2 studies; 858 women) were associated with severe covid-19 in pregnancy. Pre-existing maternal comorbidity was a risk factor for admission to an intensive care unit (4.21, 1.06 to 16.72; I^2=0%; 2 studies; 320 women) and invasive ventilation (4.48, 1.40 to 14.37; I^2=0%; 2 studies; 313 women). Spontaneous preterm birth rate was 6% (95% confidence interval 3% to 9%; I^2=55%; 10 studies; 870 women) in women with covid-19. The odds of any preterm birth (3.01, 95% confidence interval 1.16 to 7.85; I^2=1%; 2 studies; 339 women) was high in pregnant women with covid-19 compared with those without the disease. A quarter of all neonates born to mothers with covid-19 were admitted to the neonatal unit (25%) and were at increased risk of admission (odds ratio 3.13, 95% confidence interval 2.05 to 4.78, I^2=not estimable; 1 study, 1121 neonates) than those born to mothers without covid-19.

WhAt is AlreAdy knoWn on this topic

Pregnant women are considered to be a high risk group for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, and the potential adverse effects of the virus on maternal and perinatal outcomes are of concern

In non-pregnant populations admitted to hospital with coronavirus disease 2019 (covid-19) the most common symptoms are fever, cough, and dyspnoea, reported in more than two thirds of individuals

Advancing age, high body mass index, non-white ethnicity, and pre-existing comorbidities are risk factors for severe covid-19 in the general population

WhAt this study Adds

Pregnant and recently pregnant women with covid-19 diagnosed in hospital are less likely to manifest symptoms of fever and myalgia than non-pregnant women of reproductive age and might be at increased risk of admission to an intensive care unit

Risk factors for severe covid-19 in pregnancy include increasing maternal age, high body mass index, and pre-existing comorbidities

Pregnant women with covid-19 are more likely to experience preterm birth and their neonates are more likely to be admitted to a neonatal unit

cOnclusiOn

Pregnant and recently pregnant women are less likely to manifest covid-19 related symptoms of fever and myalgia than non-pregnant women of reproductive age and are potentially more likely to need intensive care treatment for covid-19. Pre-existing comorbidities, high maternal age, and high body mass index seem to be risk factors for severe covid-19. Preterm birth rates are high in pregnant women with covid-19 than in pregnant women without the disease.

systematic review registratiOn

PROSPERO CRD42020178076.

reaDers' nOte

This article is a living systematic review that will be updated to reflect emerging evidence. Updates may occur for up to two years from the date of original publication.

introduction

Since the first report (December 2019) of the novel coronavirus disease 2019 (covid-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the number of confirmed cases and associated mortality and morbidity have increased rapidly.[1 2] Pregnant women are considered a high risk group because of concerns about the effect of covid-19 on them during and after pregnancy, and on their neonates.[3] Quantification of the rates of covid-19, its risk factors, clinical manifestations, and outcomes is key to planning clinical maternal care and management in an evolving pandemic scenario.[4]

Publications on covid-19 in pregnancy have risen steeply through individual case reports, case series, observational studies, and systematic reviews. As of 26 June 2020, more than 86 reviews have been published in this area,[5-10] with a further 94 registered in PROSPERO.[8 11] The early reviews mostly included case reports and case series that were often inappropriately meta-analysed, leading to biased estimates.[12] Subsequent reviews differed little from each other, often including similar primary studies, many with duplicate data. These reviews became quickly outdated as new evidence emerged. To date, no review has comprehensively evaluated the comparative data concerning pregnant and recently pregnant women and non-pregnant women with covid-19. Moreover, the sampling frames in primary studies have varied, ranging from universal SARS-CoV-2 testing for all pregnant women admitted to hospital[13 14] to symptom based testing.[15 16] Testing strategies have also differed within and between countries, with diagnosis in many early studies based on epidemiological risk assessment and clinical features without confirmed infection, which need to be considered in the analysis.[17] Limitations in the external and internal validity of studies make it challenging for guideline developers and policy makers to make evidence based recommendations for the management of pregnant and recently pregnant women with covid-19.

We began a living systematic review to determine the clinical manifestations of covid-19 in pregnant and recently pregnant women, identify the risk factors for complications, and quantify maternal and perinatal outcomes. This systematic review will be updated on a regular basis.

Methods

Our systematic review is based on a prospectively registered protocol (PROSPERO CRD42020178076; registered 22 April 2020)[18] to evaluate a series of research questions on covid-19 during and after pregnancy. We report our findings on the rates, clinical manifestations, risk factors, and maternal and perinatal outcomes in women with covid-19 in line with the preferred reporting items for systematic reviews and meta-analyses (PRISMA) recommendations (see appendix 1). As more relevant data become available, we shall address the research questions in our published protocol. Each cycle of our living systematic review involves weekly search updates (rounds), with analysis performed every 2-4 weeks for our monthly reporting through a dedicated website, with early analysis if new definitive evidence emerges. We plan to regularly review the planned frequency of updates.

literature search

We performed a systematic search of major databases: Medline, Embase, Cochrane database, WHO (World Health Organization) COVID-19 database, China National Knowledge Infrastructure (CNKI), and Wanfang databases from 1 December 2019 to 26 June 2020 for relevant studies on covid-19 in pregnant and recently pregnant women. To identify potential studies, we coordinated our search efforts with the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre), the WHO Library, and the Cochrane Gynaecology and Fertility group. Additional searches were conducted of preprint servers, blogs, websites that serve as repositories for covid-19 studies, social media, guidelines, and reference lists of included studies and unpublished data. We also searched the Living Overview of the Evidence (LOVE) platform from 11 to 26 June 2020.[19] We contacted established groups that were coordinating or conducting surveillance and studies in pregnant women with covid-19, such as the WHO Maternal, Newborn, Child and Adolescent health (MNCAH) covid-19 research network and the International Network of Obstetric Survey Systems (INOSS) for information on published and upcoming data. No language restrictions were applied. Appendix 2 provides details of the search strategies and databases searched.

study selection

Two reviewers independently selected studies using a two stage process: they first screened the titles and abstracts of studies and then assessed the full text of the selected studies in detail for eligibility. A total of eight reviewers contributed to study selection.

Disagreements were resolved through discussion with a third reviewer (ST or JA). We excluded studies if the duplicate data for all outcomes of interest were published elsewhere, as reported by the study authors, or when the characteristics of the mother or neonate matched the setting, characteristics, and duration of another study. When we suspected an overlap of data between studies, the study that provided comparative data was included. When there was uncertainty about duplicate data, we contacted the authors of primary studies.

We defined women as having confirmed covid-19 if they had laboratory confirmation of covid-19 infection irrespective of clinical signs and symptoms.[20] Women with a diagnosis based only on clinical or radiological findings were defined as having suspected covid-19. The recently pregnant group comprised women in the postpartum and post-abortion period. We included studies that compared covid-19 rates, clinical manifestations (symptoms, laboratory and radiological results), risk factors, and associated mortality and morbidity between pregnant and recently pregnant and non-pregnant women of reproductive age, and those that compared maternal and perinatal outcomes in pregnant women with and without covid-19. Studies on non-comparative cohorts with a minimum of 10 participants were included if they reported on the rates and clinical manifestations of covid-19 and relevant outcomes in pregnant and recently pregnant women. We defined cohort studies as those that sampled participants on the basis of exposure, followed-up participants over time, and ascertained the outcomes.[21] The PROSPERO protocol provides a full list of the risk factors, clinical features, and outcomes evaluated.[18]

The sampling frames for detecting covid-19 included universal screening and testing, when all women were assessed for covid-19 using reverse transcriptase polymerase chain reaction (RT-PCR) for SARS-CoV-2 or chest computed tomography; risk based testing on the basis of epidemiological history and clinical manifestations by National Health Commission of China (NHCC) guidelines[17]; and symptom based when testing was performed on women with symptoms and those with a history of contact with affected individuals. We defined the population as being selected when only specific groups of women were included, such as those undergoing caesarean section or in the third trimester. We categorised studies as a high risk group if only women with any pre-existing medical or obstetric risk factors were included, low risk if women did not have any risk factors, and any risk if all women were included.

study quality assessment and data extraction

The quality of the comparative cohort studies was assessed for selection, comparability, and outcome ascertainment bias using the Newcastle Ottawa scale.[22] Studies achieving four stars for selection, two for comparability, and three for ascertainment of the outcome were considered to have a low risk of bias. Studies achieving two or three stars for selection, one for comparability, and two for outcome ascertainment were considered to have a medium risk of bias, and any study achieving one star for selection or outcome ascertainment, or zero for any of the three domains, was regarded as having a high risk of bias. We assessed the quality of studies reporting on the prevalence of clinical manifestations or outcomes for internal and external validity using an existing tool.[23] The following were considered as low risk of bias for external validity: representative of national population for relevant variables (population), representative of target population (sampling frame), random selection (selection bias), and more than 75% response rate in individuals with and without the outcome (non-response bias).[23] Two independent reviewers extracted data using a pre-piloted form.

statistical analysis

We pooled the comparative dichotomous data using random effects meta-analysis and summarised the findings as odds ratios with 95% confidence intervals. To combine comparative continuous data with dichotomous data we transformed standardised mean differences to logarithm odds ratios, assuming a normal underlying distribution.[24] We pooled the dichotomous non-comparative data for rates of clinical manifestations and maternal and perinatal outcomes as proportions with 95% confidence intervals using Dersimonian and Laird random effects meta-analysis after transforming data using Freeman-Tukey double arcsin transformation. Heterogeneity was reported as I^2 statistics. We undertook subgroup analysis by country status (high versus low and middle income), sampling frame (universal, risk based, and symptom based testing, including not reported), and risk status of women in the studies (high, low, any). Sensitivity analysis was performed by restricting the analysis to women with confirmed covid-19, study quality (high, low), and population (unselected, selected). All analyses were done with Stata (version 16).

Patient and public involvement

The study was supported by Katie's Team, a dedicated patients and public involvement group in Women's Health. The team was involved in the conduct, interpretation, and reporting of this living systematic review through participation in virtual meetings.

results

After removing duplicates from 49 684 citations, 20 625 unique citations were identified and 77 cohort studies (55 comparative, 22 non-comparative) were included in the systematic review (fig 1).

characteristics of included studies

Of the 77 studies, 26 (34%) were from the United States, 24 from China (31%), seven from Italy, six from Spain, three each from the United Kingdom and France, and one each from Belgium, Brazil, Denmark, Israel, Japan,

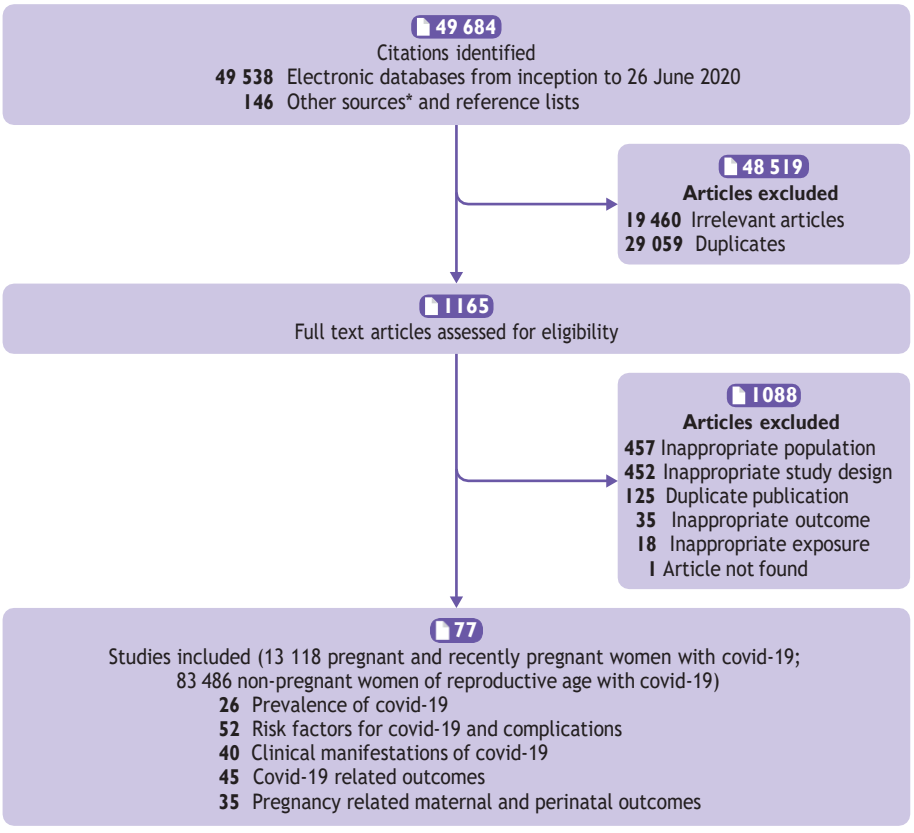


Fig 1 | study selection process. *twitter, national reports, blog by j thornton, Obg Project, cOviD-19 and Pregnancy cases, www.obgproject.com/2020/04/07/covid-19-research-watch-with-dr-jim-thornton/ (accessed 12 may 2020); ePPi-centre, cOviD-19: a living systematic map of evidence, http://eppi.ioe.ac.uk/cms/Projects/DepartmentofHealthandsocialcare/Publishedreviews/cOviD-19livingsystematicmapoftheevidence/tabid/3765/Default.aspx (accessed 12 may 2020); norwegian institute of Public Health, niPH systematic and living map on cOviD-19 evidence, www.nornesk.no/forskningskart/niPH_mainmap.html (accessed 19 may 2020); johns Hopkins university center for Humanitarian Health; cOviD-19, maternal and child Health, nutrition, http://hopkinshumanitarianhealth.org/empower/advocacy/covid-19/covid-19-children-and-nutrition/ (accessed 2 june 2020); researchgate, cOviD-19 research community, www.researchgate.net/community/cOviD-19 (accessed 2 june 2020); and living Overview of the evidence, coronavirus disease (cOviD-19), https://app.iloveevidence.com/loves/5e6fdb9669c00e4ac072701d?population=5d062d5fc80dd41e58ba8459 (accessed 16 june 2020)

Mexico, the Netherlands, and Portugal. All the studies tested respiratory samples using RT-PCR to confirm the presence of SARS-CoV-2; 23 studies additionally diagnosed covid-19 based on clinical suspicion. Eight studies (95 247 women) compared pregnant populations with non-pregnant populations,[25-32] and four studies (2230 women) compared pregnant women with covid-19 versus pregnant women without covid-19.[33-36] Forty cohort studies reported on clinical manifestations (13 018 pregnant, 85 084 non-pregnant women),[25-32 35-66] 45 studies reported on covid-19 related maternal outcomes (14 094 pregnant, 85 169 non-pregnant women),[25-32 35-51 53-59 61-74] and 35 studies reported on pregnancy related maternal (6279 women) and perinatal outcomes (2557 neonates)[13 25 27 29 30 32-41 43-47 49-50 54 55 57 59 61 62 64-67 69 70 75] (see appendix 3). The sampling frames included universal testing (29 studies), risk based NHCC guidelines (22 studies), and symptom based (19 studies) strategies. Eleven studies did not report the sampling strategy.

Quality of included studies

Overall, 67% (37/55) of the comparative cohort studies evaluated using the Newcastle Ottawa scale had an overall low risk of bias (see appendix 4a). Forty nine (89%) had a low risk of bias for study selection and six (11%) had a medium risk. The risk of bias for comparability of cohorts was low in nine of the studies (16%), medium in 45 (82%), and high in one (2%). For outcome assessment of the cohorts, 12 (22%) studies had a low risk of bias, 42 (76%) a medium risk, and one (2%) a high risk. Quality assessment of the prevalence studies for external validity showed a low risk of bias for representativeness in 13% (10/76) of the studies, sampling in 26% (20/76), selection in 74% (56/76), and non-response in 96% (73/76). For internal validity, there was low risk of bias for data collection in 95% (72/76) of the studies, case definition in 36% (27/76), measurement in 99% (75/76), differential verification in 86% (65/76), adequate follow-up in 22% (17/76), and appropriate numerator and denominator in 83% (63/76) (see appendix 4b).

rates of covid-19 in pregnant and recently pregnant women

The overall rate of covid-19 diagnosis in pregnant and recently pregnant women attending or admitted to hospital for any reason was 10% (95% confidence interval 7% to 14%; 26 studies, 11 432 women; fig 2). Rates varied by sampling strategy: of the women sampled by universal screening, 7% (4% to 10%; 18 studies, 6247 women) were diagnosed as having covid-19 compared with 18% (10% to 28%; 8 studies, 4928 women) of women sampled on the basis of symptoms. All studies with a prevalence rate for covid-19 greater than 15% were from the US, except for one study, which was from France.[76] One in 20 asymptomatic mothers (5%, 2% to 9%; 11 studies) attending or admitted to hospital had a diagnosis of covid-19 (see appendix 5a). Three quarters (74%, 51% to 93%; 11 studies) of the 162 pregnant women with covid-19 in the universal screening population were asymptomatic (see appendix 5b). Based on data from a small number of studies, a diagnosis of covid-19 in pregnancy was associated with maternal obesity (odds ratio 1.75, 95% confidence interval 1.34 to 2.30; 1 study, 1080 women), pre-existing comorbidities (1.64, 1.25 to 2.13; 1 study, 1121 women), asthma (1.71, 1.03 to 2.84; 2 studies, 1250 women), history of covid-19 in the support person (44.56, 14.90 to 133.28; 1 study, 199 women), and gestational diabetes (2.42, 1.55 to 3.79; 1 study, 1121 women) (see appendix 6a).

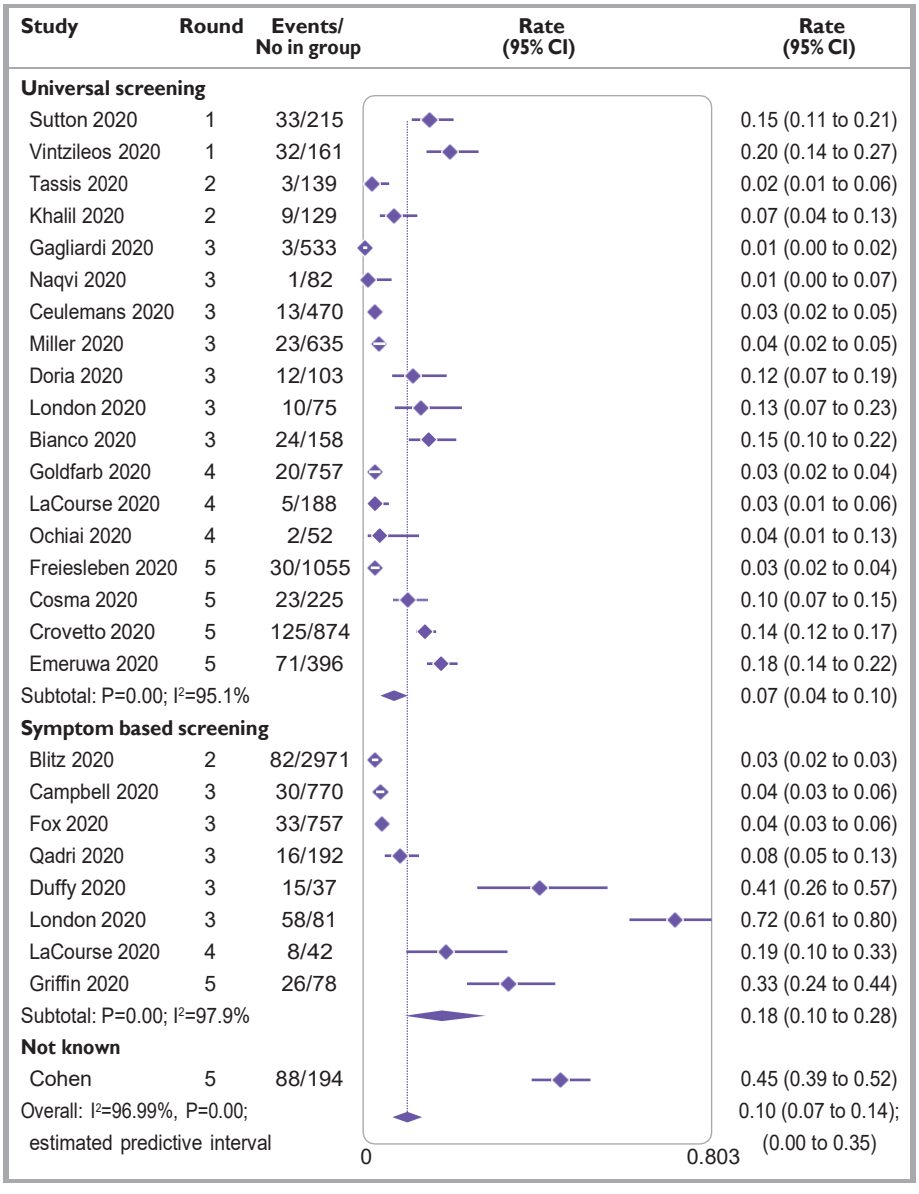


Fig 2 | Prevalence of severe acute respiratory syndrome coronavirus 2 in pregnant and recently pregnant women identified by various sampling strategies. meta-analysis includes one study (liao 2020) screened using national Health commission china criteria with no events. symptom based screening includes screening based on symptoms or history of contact with individuals with covid-19. round number represents search strategy updates in the living systematic review

clinical manifestations of covid-19 during pregnancy and after delivery

The most common symptoms reported by pregnant and recently pregnant women with suspected or confirmed covid-19 were fever (40%) and cough (39%); lymphopaenia (35%) and raised C reactive protein levels (49%) were the most common laboratory findings (fig 3). Compared with non-pregnant women of reproductive age with covid-19, pregnant and recently pregnant women with the disease were less likely to manifest symptoms of fever (0.43, 0.22 to 0.85; 5 studies, 80 521 women) and myalgia (0.48, 0.45 to 0.51; 3 studies, 80 409 women) (fig 4). A history of pre-existing diabetes was more often observed in pregnant women with covid-19 than in non-pregnant women with the disease (1.78, 1.03 to 3.05; 3 studies, 91 595 women) (see appendix 6b). Sensitivity analysis restricted to various sampling frames showed lower estimates of fever, cough, and dyspnoea in the universal screening population and higher estimates in the symptom based population (see appendix 7). The rates of clinical manifestations were similar to the overall estimates when the analysis was restricted to only women with RT-PCR confirmed covid-19, unselected populations, and women with any risk (see appendix 7).

Outcomes related to covid-19 in pregnant and recently pregnant women

Overall, 73 pregnant women (26 studies, 11 580 women) with confirmed covid-19 died from any cause (0.1%, 95% confidence interval 0.0% to 0.7%). Severe covid-19 was diagnosed in 13% (6% to 21%; 21 studies, 2271 women) of pregnant and recently pregnant women with suspected or confirmed covid-19; 4% (2% to 7%; 17 studies, 10 901 women) of the pregnant women with covid-19 were admitted to an intensive care unit, 3% (1% to 5%; 13 studies, 10 713 women) required invasive ventilation, and 0.4% (0.1% to 0.9%; 9 studies, 1935 women) required extracorporeal membrane oxygenation (fig 3). Appendix 8 provides the rates of complications by sampling strategy. Compared with non-pregnant women of reproductive age with covid-19, the odds of admission to the intensive care unit (1.62, 95% confidence interval 1.33 to 1.96) and need for invasive ventilation (1.88, 1.36 to 2.60) were higher in pregnant and recently pregnant women (four studies, 91 606 women) (table 1). Maternal risk factors associated with severe covid-19 were increasing age (1.78, 1.25 to 2.55; 4 studies, 1058 women), high body mass index (2.38, 1.67 to 3.39; 3 studies, 877 women), chronic hypertension (2.0, 1.14 to 3.48; 2 studies, 858 women), and pre-existing diabetes (2.51, 1.31 to 4.80; 2 studies, 858 women) (fig 5). Pre-existing maternal comorbidity was associated with admission to an intensive care unit (4.21, 1.06 to 16.72; 2 studies, 320 women) and the need for invasive ventilation (4.48, 1.40 to 14.37; 2 studies, 313 women) (table 2).

maternal and perinatal outcomes in pregnant and recently pregnant women with covid-19

In pregnant and recently pregnant women with covid-19 the rate of overall preterm birth was 17% (95% confidence interval 13% to 21%; 30 studies, 1872 women) and of spontaneous preterm birth was 6% (3% to 9%; 10 studies, 870 women) (fig 3). In pregnant and recently pregnant women with covid-19 compared with pregnant and recently pregnant women without the disease, the odds of any preterm birth (3.0, 95% confidence interval 1.15 to 7.85; 2 studies, 339 women) were higher, but no differences were observed in other maternal outcomes (table 1). Eighteen stillbirths (27 studies; 2837 offspring) and six neonatal deaths (26 studies; 1728 neonates) occurred among pregnant and recently pregnant women with covid-19, resulting in negligible risks (fig 3). Overall, 25% (95% confidence interval 14% to 37%; 17 studies, 1348 women) of neonates born to women with covid-19 were admitted to the neonatal unit (fig 3), with a higher risk of admission (odds ratio 3.13, 95% confidence interval 2.05 to 4.78; 1 study, 1121 neonates) than those born to mothers without the disease in one study with historical controls. No differences were observed for other perinatal outcomes. Appendix 9 provides the rates of covid-19 related and pregnancy related outcomes for the individual studies.

discussion

In this living systematic review, we found that one in 10 pregnant or recently pregnant women who are attending or admitted to hospital for any reason are diagnosed as having suspected or confirmed covid-19, although the rates vary by sampling strategy. The covid-19 related symptoms of fever and myalgia manifest less often in pregnant and recently pregnant women than in non-pregnant women of reproductive age. Whereas testing for SARS-CoV-2 in non-pregnant women is based on symptoms or contact history, testing in pregnant women is usually done when they are in hospital for reasons that might not be related to covid-19. Pregnant or recently pregnant women with covid-19 seem to be at increased risk of requiring admission to an intensive care unit or invasive ventilation. Increased maternal age, high body mass index, and pre-existing comorbidities might be associated with severe disease. Pregnant women with covid-19 are at increased risk of delivering preterm and their babies being admitted to the neonatal unit. But overall rates of spontaneous preterm births are not high. Stillbirth and neonatal death rates are low in women with suspected or confirmed covid-19. All comparative findings are based on small numbers of studies, despite the large sample sizes. Substantial heterogeneity was observed in the estimates for rates of clinical manifestations and outcomes, which varied by sampling frames, participant selection, and risk status of the participants.

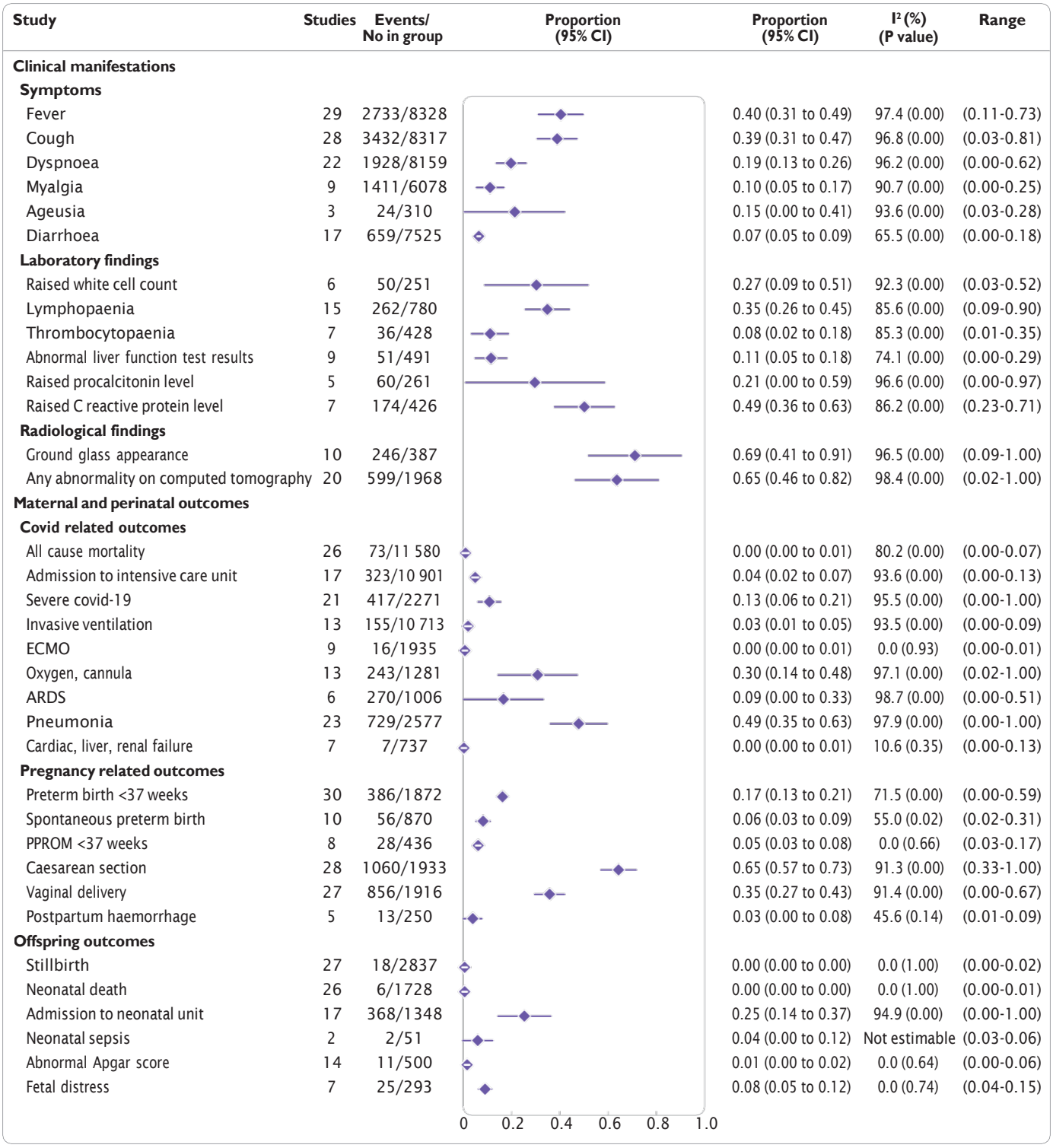
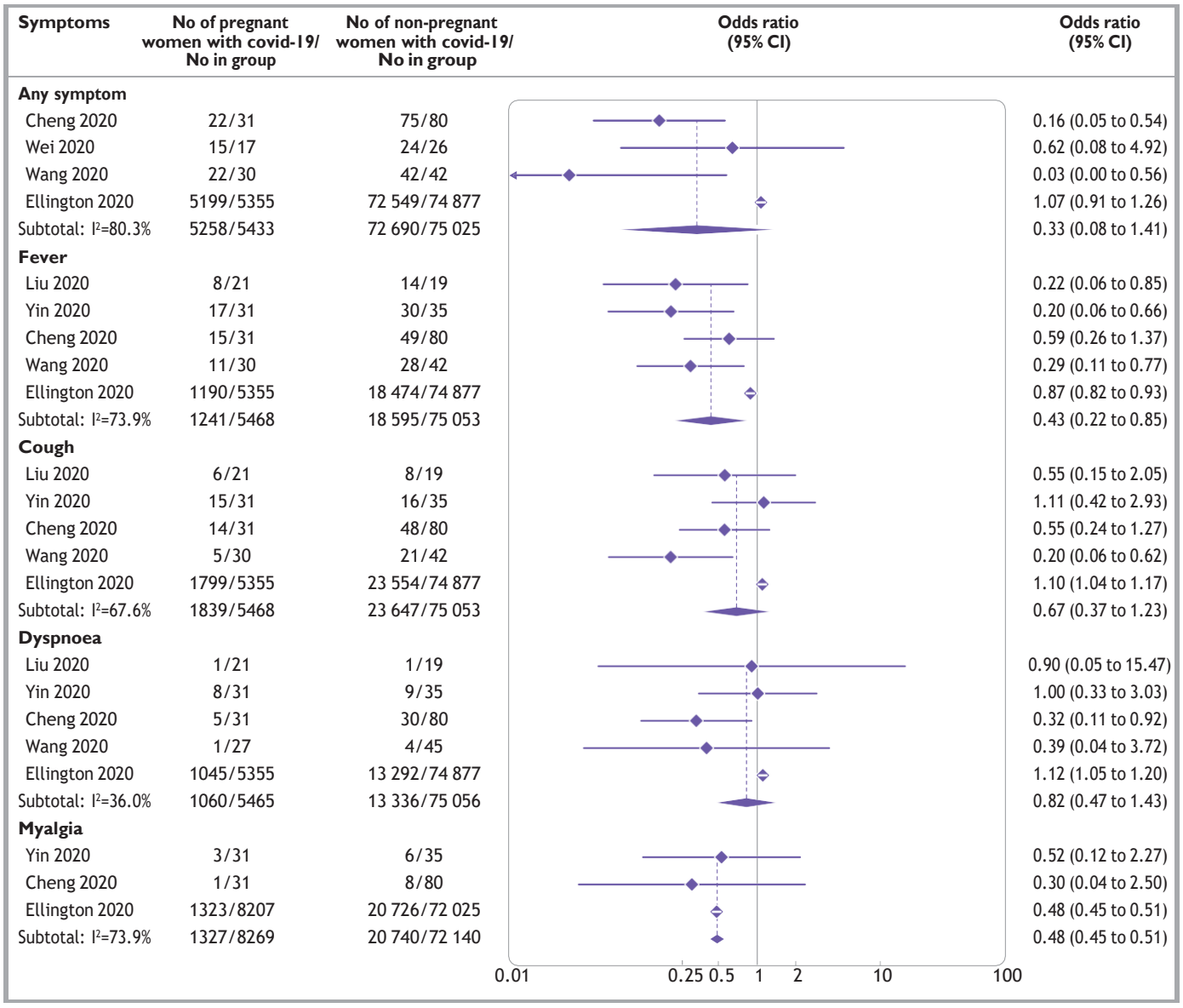


Fig 3 | rates of clinical manifestations of coronavirus disease (covid-19) in pregnant women and recently pregnant women with suspected or confirmed covid-19 and associated maternal and perinatal outcomes. ecmO=extracorporeal membrane oxygenation; arDs=acute respiratory distress syndrome; PPrOm=preterm premature rupture of membranes

strengths and limitations of this review

In this unprecedented pandemic situation, where evidence is rapidly produced and published in various formats, our living systematic review underpinned by robust methods and continually updated at regular intervals is relevant for several reasons. Firstly, it addresses important research questions relevant to clinical decision making and policies. Secondly, uncertainties remain for key outcomes that require further evidence. Thirdly, the rapid turnover of evidence in various formats requires assessments of study quality and regular updating of the findings.



Note: Weights are from random effects analysis

Fig 4 | clinical manifestations of coronavirus disease (covid-19) in pregnant and recently pregnant women compared with non-pregnant women of reproductive age with covid-19

Finally, our living systematic review will produce a strong evidence base for living guidelines on covid-19 and pregnancy.

We undertook a comprehensive search and coordinated our efforts with key organisations and research groups, such as WHO, the Cochrane Centre, and EPPI-Centre. To minimise risk of bias we restricted our meta-analysis to cohort studies, and we reported the quality of the included studies. By contacting the authors and obtaining reports not published in PubMed, we minimised the risk of missing relevant studies. Our systematic review has a large sample size and it is continuously increasing. Our living meta-analyses framework will enable us to rapidly update the findings as new data emerge. We undertook extensive work to ensure that duplicate data are not included. Our various comparative analyses allowed us to comprehensively assess the association between pregnancy and covid-19 related outcomes, covid-19 and pregnancy outcomes, risk factors for SARS-CoV-2 infection, and complications. Our review helps to understand the variations in estimates through sensitivity analyses by sampling strategies, population characteristics, and risk factors, and it provides confidence in the rates of reported outcomes.

Our systematic review also has limitations. The primary studies used varied sampling frames to identify women with covid-19, comprised women with suspected and confirmed covid-19, and primarily reported on pregnant women who required visits to hospital, including for childbirth, thereby affecting the generalisability of the estimates. Although our

table 1 | Outcomes in pregnant and recently pregnant women with coronavirus disease 2019 (covid-19)

Outcomes	no of studies	women (no with event/no in group (%)) Pregnant women with covid-19	comparison group	Odds ratio (95% ci)	i^2 (%)
comparison group: non-pregnant women of reproductive age with covid-19					
All cause mortality	4	16/8282 (0.2)	208/83 327 (0.2)	0.81 (0.49 to 1.33)	0
ICU admission	4	121/8276 (1.5)	758/83 330 (0.9)	1.62 (1.33 to 1.96)	0
Invasive ventilation	4	43/8276 (0.5)	226/83 330 (0.3)	1.88 (1.36 to 2.60)	0
ECMO	1	0/31 (0)	0/80 (0)	2.56 (0.05 to 131.60)	NE
Oxygen through nasal cannula	2	8/48 (16.7)	49/106 (46.2)	0.21 (0.04 to 1.13)	65.7
ARDS	1	0/17 (0)	0/26 (0)	1.51 (0.03 to 79.93)	NE
Major organ failure	1	0/17 (0)	0/26 (0)	1.51 (0.03 to 79.93)	NE
comparison group: pregnant women without covid-19					
Maternal outcomes:					
All cause mortality	1*	5/427 (1.2)	0/694 (0)	18.08 (1.00 to 327.83)	NE
ICU admission	1*	40/427 (9.4)	1/694 (0.1)	71.63 (9.81 to 523.06)	NE
Preterm birth <37 weeks	2	7/44 (15.9)	18/295 (6.1)	3.01 (1.16 to 7.85)	0.9
Caesarean section	3*	184/491 (37.5)	577/1676 (34.4)	2.02 (0.67 to 6.10)	87.5
Perinatal outcomes:					
Stillbirth	1*	3/427 (0.7)	2/694 (0.3)	2.45 (0.41 to 14.71)	NE
Neonatal death	1*	2/427 (0.5)	1/694 (0.1)	3.26 (0.30 to 36.07)	NE
Admission to neonatal unit	1*	64/427 (15.0)	37/694 (5.3)	3.13 (2.05 to 4.79)	NE
Abnormal Apgar score at 5 minutes	1	0/30 (0)	12/740 (1.6)	0.96 (0.06 to 16.51)	NE
Fetal distress	1	3/34 (8.8)	12/242 (5.0)	1.86 (0.50 to 6.94)	NE

ICU=intensive care unit; ECMO=extracorporeal membrane oxygenation; ARDS=acute respiratory distress syndrome; NE=not estimable.
The denominator is number of pregnancies for all outcomes.
*Historical comparative cohort in UK Obstetric Surveillance System study.

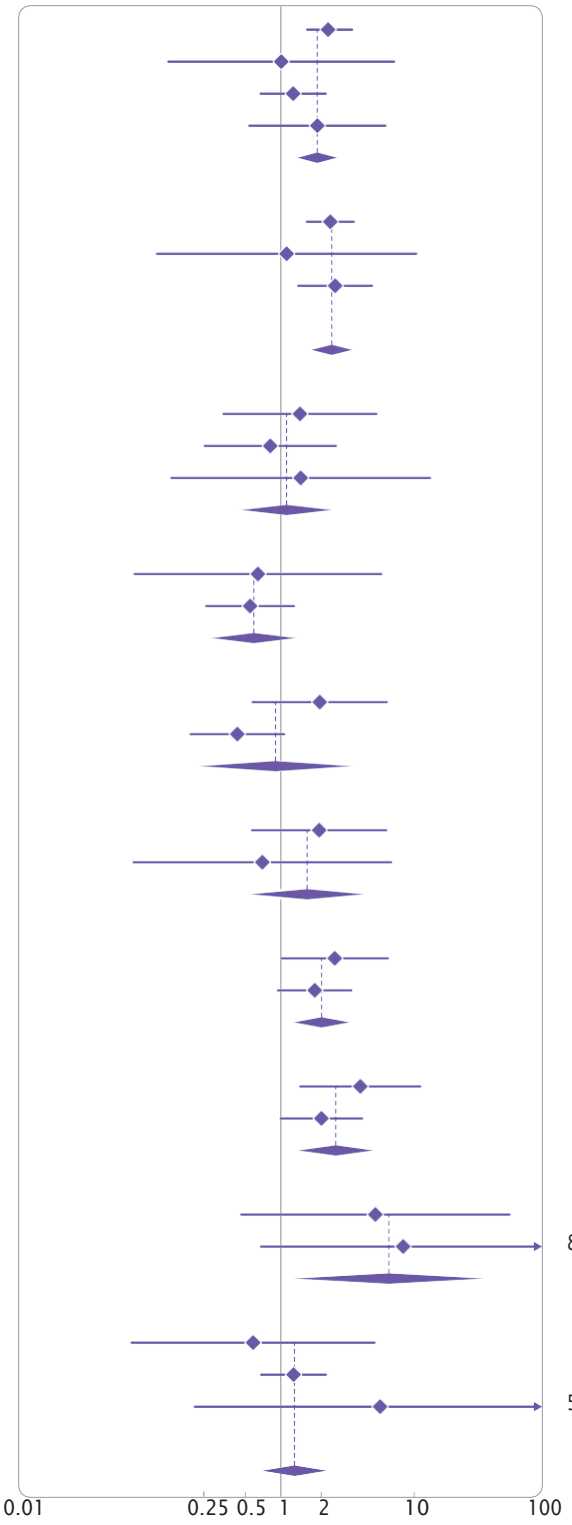
sensitivity analyses aimed to tackle some of these problems, the numbers and sample sizes of the individual studies were too small to identify differences between the subgroups. The timing of assessment of the clinical manifestations of disease was generally not available. The definitions of symptoms, tests, and outcomes were heterogeneous. Furthermore, poor reporting of the criteria for caesarean section, admissions to the neonatal unit, and the causes of preterm birth, made it difficult to disentangle iatrogenic effect from the true impact of the disease. There is a paucity of comparative data to assess the risk of severe disease in pregnant women compared with non-pregnant women in similarly aged groups, and to compare pregnancy outcomes in women with and without covid-19. Not many studies reported outcomes by trimester for symptom onset, making it difficult to assess the rates of miscarriage and postpartum complications. For some outcomes, the findings were influenced by a single large study.[26] Many studies had to be excluded as we could not rule out potential overlap in the study populations.

comparison with existing evidence

Alongside the spread of the pandemic, a shift has occurred in the types of studies published, with initial studies involving pregnant women from epidemic regions in China, followed by reports of large regional and national datasets from the US, UK, Netherlands, Spain, and, more recently, Latin American countries. The study design has also changed from initial small case series and case reports to large observational data, with recent studies also providing comparative data. The prevalence of covid-19 varied widely between studies, particularly when sampling was done based on symptoms or history of contact, highlighting the variations in criteria for testing. Moreover, the findings only relate to those women attending hospital for any reason. The true prevalence of covid-19 in pregnancy is likely to be lower when all pregnant women are included.

In the recent cohort study of all individuals admitted with covid-19 in the UK, the cluster of respiratory symptoms of cough, fever, and breathlessness were observed in more than two thirds of individuals,[77] similar to reported rates in the US and China.[78-80] But in our review, fewer pregnant and recently pregnant women with covid-19 manifested these symptoms than the non-pregnant population, indicating possible high rates of asymptomatic presentation in this population. This is likely because of the strategy of universal screening for covid-19 in pregnancy and the low thresholds for testing than in non-pregnancy. Despite the possibility of the above strategies detecting pregnant women with mild disease, we observed an increase in admissions to the intensive care unit and need for invasive ventilation compared with non-pregnant women of reproductive age with covid-19. The findings were mainly influenced by the recent large Centers for Disease Control and Prevention report from the US.[26] Pregnancy status was not ascertained in a large proportion of women of reproductive age in the CDC report that could affect the estimates. Furthermore, the outcomes for which the data were missing were considered to be absent in the report, thereby incurring bias. The pooled estimates for severe covid-19 and admission to an intensive care unit were, however, still relatively high in the non-comparative data, indicative of a potential high risk in pregnancy. This is supported by the recent analysis in a Swedish study suggesting a high risk of admission to an intensive care unit and invasive ventilation in pregnant women than non-pregnant women.[81]

Risk factors	No of pregnant women with risk factor and severe covid-19/No in group	No of pregnant women with risk factor without severe covid-19/No in group	Odds ratio (95% CI)	Odds ratio (95% CI)
Age*				
Kayem 2020	59/128	135/489		2.24 (1.50 to 3.35)
Martinez-Perez 2020	2/4	39/78		1.00 (0.13 to 7.46)
Khoury 2020	22/75	43/166		1.19 (0.65 to 2.18)
Chen 2020 (continuous age)	n/9	n/109		1.87 (0.55 to 6.42)
Subtotal: I^2=9%	83/216	219/842		1.78 (1.25 to 2.55)
Body mass index				
Kayem 2020	46/128	93/489		2.39 (1.56 to 3.66)
Martinez-Perez 2020	1/4	18/78		1.11 (0.11 to 11.35)
Khoury 2020	43/62	55/116		2.51 (1.31 to 4.81)
Wu 2020	0/0	0/13		Excluded
Subtotal: I^2=0%	90/194	166/696		2.38 (1.67 to 3.39)
Multiparity				
Chen 2020	5/9	46/97		1.39 (0.35 to 5.47)
Savasi 2020	8/14	39/63		0.82 (0.25 to 2.66)
Martinez-Perez 2020	3/4	53/78		1.42 (0.14 to 14.29)
Subtotal: I^2=0%	16/27	138/238		1.07 (0.46 to 2.46)
Third trimester				
Yan 2020	7/8	99/108		0.64 (0.07 to 5.76)
Andrikopoulou 2020	22/34	94/124		0.59 (0.26 to 1.32)
Subtotal: I^2=0%	29/42	193/232		0.59 (0.28 to 1.27)
Non-white				
Savasi 2020	6/14	18/63		1.88 (0.57 to 6.17)
Khoury 2020	54/65	143/156		0.45 (0.19 to 1.06)
Subtotal: I^2=73%	60/79	161/219		0.86 (0.21 to 3.50)
Any comorbidity				
Savasi 2020	6/14	18/63		1.88 (0.57 to 6.17)
Martinez-Perez 2020	1/4	25/78		0.71 (0.07 to 7.14)
Subtotal: I^2=0%	7/18	43/141		1.53 (0.53 to 4.41)
Chronic hypertension				
Kayem 2020	7/128	11/489		2.51 (0.95 to 6.62)
Khoury 2020	18/75	25/166		1.78 (0.90 to 3.51)
Subtotal: I^2=0%	25/203	36/655		2.00 (1.14 to 3.48)
Pre-existing diabetes				
Kayem 2020	7/128	7/489		3.98 (1.37 to 11.57)
Khoury 2020	16/75	20/166		1.98 (0.96 to 4.08)
Subtotal: I^2=12%	23/203	27/655		2.51 (1.31 to 4.80)
Pre-eclampsia				
Yan 2020	1/8	3/108		5.00 (0.46 to 54.51)
Martinez-Perez 2020	1/4	3/78		8.33 (0.66 to 105.71)
Subtotal: I^2=0%	2/12	6/186		6.35 (1.11 to 36.22)
Gestational diabetes				
Andrikopoulou 2020	1/34	6/124		0.60 (0.07 to 5.13)
Kayem 2020	17/128	54/489		1.23 (0.69 to 2.21)
Martinez-Perez 2020	0/4	1/78		5.74 (0.20 to 161.79)
Yan 2020	0/0	9/116		Excluded
Subtotal: I^2=0%	18/166	70/807		1.23 (0.70 to 2.14)



Note: Weights are from random effects analysis

Fig 5 | risk factors associated with severe coronavirus disease 2019 (covid-19) in pregnant and recently pregnant women. symptom based screening: savasi v, Kayem g; nHcc (national Health commission china). criteria based screening: chen, wu, yan. all other studies used universal screening. cut-off for age is 35 years or more, and for body mass index is 30 or more. *includes one study with continuous measurement of risk factor

table 2 | maternal characteristics associated with severe coronavirus disease 2019 (covid-19) and all cause death in pregnant and recently pregnant women with a diagnosis of covid-19

maternal risk factors and outcomes	no of studies	total no of women	Pregnant women (no with risk factor/no in group (%))		Odds ratio (95% ci)	i^2 (%)
			with outcome	without outcome		
Age ≥35 years:						
Severe disease	4	1058	216*	842*	1.78 (1.25 to 2.55)	9
ICU admission	2	260	8/87 (9.2)	8/173 (4.6)	2.44 (0.43 to 14.01)	63
Invasive ventilation	1	178	3/65 (4.6)	2/113 (1.8)	2.69 (0.44 to 16.51)	NE
Maternal death	1	288	20/154 (13.0)	16/134 (11.9)	1.10 (0.55 to 2.22)	NE
Multiparity:						
Severe disease	3	265	16/154 (10.4)	11/111 (9.9)	1.07 (0.46 to 2.46)	0
ICU admission	1	42	4/22 (18.2)	4/20 (20.0)	0.89 (0.19 to 4.15)	NE
Body mass index ≥30:						
Severe disease	3	877	90/256 (35.2)	104/621 (16.7)	2.38 (1.67 to 3.39)	0
ICU admission	1	142	3/22 (13.6)	4/120 (3.3)	4.58 (0.95 to 22.09)	NE
Invasive ventilation	1	135	5/21 (23.8)	6/114 (5.3)	5.63 (1.54 to 20.59)	NE
Maternal death	2	596	6/62 (9.7)	37/534 (6.9)	2.57 (0.97 to 6.82)	0
Non-white ethnicity:						
Severe disease	2	298	60/221 (27.1)	19/77 (24.7)	0.86 (0.21 to 3.50)	73
ICU admission	1	42	5/20 (25.0)	3/22 (13.6)	2.11 (0.43 to 10.28)	NE
Maternal death	2	596	31/220 (14.1)	12/376 (3.2)	2.40 (0.94 to 6.11)	0
Any comorbidity:						
Severe disease	2	159	7/50 (14.0)	11/109 (10.1)	1.53 (0.53 to 4.41)	0
ICU admission	2	320	4/37 (10.8)	11/283 (3.9)	4.21 (1.06 to 16.72)	0
Invasive ventilation	2	313	6/36 (16.7)	10/277 (3.6)	4.48 (1.40 to 14.37)	0
Chronic hypertension:						
Severe disease	2	858	25/61 (41.0)	178/797 (22.3)	2.0 (1.14 to 3.48)	0
ICU admission	1	141	2/5 (40.0)	5/136 (3.7)	17.47 (2.37 to 129.02)	NE
Invasive ventilation	1	134	4/5 (80.0)	7/129 (5.4)	69.71 (6.85 to 709.34)	NE
Maternal death	2	596	5/29 (17.2)	38/567 (6.7)	3.38 (1.17 to 9.75)	0
Pre-existing diabetes:						
Severe disease	2	858	23/50 (46.0)	180/808 (22.3)	2.51 (1.31 to 4.80)	12
ICU admission	2	181	1/7 (14.3)	14/174 (8.0)	2.88 (0.44 to 18.96)	0
Invasive ventilation	1	132	1/6 (16.7)	9/126 (7.1)	2.60 (0.27 to 24.71)	NE
Maternal death	2	596	10/52 (19.2)	33/544 (6.1)	6.63 (0.27 to 161.45)	91
Asthma:						
Severe disease	3	857	17/61 (27.9)	149/796 (18.7)	1.86 (0.88 to 3.93)	22
Maternal death	2	596	3/22 (13.6)	40/574 (7.0)	2.04 (0.61 to 6.85)	0
Smoking:						
Severe disease	3	776	5/23 (21.7)	141/753 (18.7)	1.67 (0.64 to 4.40)	0
ICU admission	1	42	1/2 (50.0)	7/40 (17.5)	4.71 (0.26 to 84.77)	NE
Maternal death	1	308	0/10 (0)	7/298 (2.3)	1.85 (0.10 to 34.60)	NE
Gestation ≥28 weeks:						
Severe disease	2	274	29/222 (13.1)	13/52 (25.0)	0.59 (0.28 to 1.27)	0
Maternal death	1	273	22/190 (11.6)	12/83 (14.5)	0.78 (0.36 to 1.65)	NE
Gestational diabetes:						
Severe disease	4	973	18/88 (20.5)	148/885 (16.7)	1.23 (0.70 to 2.14)	0
Pre-eclampsia:						
Severe disease	2	198	2/8 (25.0)	10/190 (5.3)	6.36 (1.12 to 36.22)	0
ICU admission	1	42	6/6 (100.0)	2/36 (5.6)	179.40 (7.69 to 4186.05)	NE

ICU=intensive care unit; NE=not estimable.
*Includes one or more studies with continuous measurement of risk factor.

Similar to the general population, high body mass index and pre-existing comorbidity seemed to be risk factors for severity of covid-19 in pregnancy, including admission to an intensive care unit and invasive ventilation.[77] Complications related to covid-19 did not seem to be increased in women presenting in the third trimester or in multiparous women—but existing sample sizes are not large. Both chronic hypertension and pre-existing diabetes were associated with maternal death in pregnant women with covid-19, which are known risk factors in the general population. But it is not known if covid-19 was the direct cause of death for these women, and the numbers of studies are small. We observed an increase in rates of preterm birth in pregnant women with covid-19 compared with those without the disease. These preterm births could be medically indicated, as the overall rates of spontaneous preterm births in pregnant women with covid-19 was broadly similar to those observed in the pre-pandemic period. Although more than 60% of pregnant women underwent caesarean section in the non-comparative studies, we did not find a statistically significant difference in comparative studies of pregnant women with and without covid-19. The precision of the estimates is expected to improve with the publication of more data in the future. The overall

rates of stillbirths and neonatal deaths do not seem to be higher than the background rates. The indications for admissions to the neonatal unit, observed in about a quarter of neonates delivered to mothers with covid-19, have not been reported. Local policies on observation and quarantine of infants with exposure to SARS-CoV-2 might have influenced these rates.

relevance for clinical practice and research

Based on existing data, healthcare professionals should be aware that pregnant and recently pregnant women with covid-19 might manifest fewer symptoms than the general population, with the overall pattern similar to that of the general population. Emerging comparative data indicate the potential for an increase in the rates of admission to intensive care units and invasive ventilation in pregnant women compared with non-pregnant women. Mothers with pre-existing comorbidities will need to be considered as a high risk group for covid-19, along with those who are obese and of greater maternal age. Clinicians will need to balance the need for regular multidisciplinary antenatal care to manage women with pre-existing comorbidities against unnecessary exposure to the virus, through virtual clinic appointments when possible. Pregnant women with covid-19 before term gestation might need to be managed in a unit with facilities to care for preterm neonates.

Further data are needed to assess robustly if pregnancy related maternal and neonatal complications are increased in women with covid-19 than those without the disease. Similarly, the association between other risk factors such as ethnicity and pregnancy specific risk factors such as pre-eclampsia and gestational diabetes on both covid-19 related and pregnancy related outcomes needs evaluation. Pre-eclampsia was reported to be associated with severe covid-19 in small studies, but it requires further assessment as the clinical presentation of severe pre-eclampsia could mimic worsening covid-19.[82] Robust collection of maternal data by trimester of exposure, including the periconception period, is required to determine the effects of covid-19 on early pregnancy outcomes, fetal growth, and risk of stillbirth.

Systematic reviews are considered to be the highest quality evidence informing guidelines, and poor quality reviews will have a direct impact on clinical care. Despite the urgent need for evidence on the impact of covid-19 in pregnant women, systematic reviews and meta-analyses still need to adhere to the reporting guidelines on search criteria, quality assessment, and analysis. This is particularly important as large numbers of non-peer reviewed scientific papers and reports are currently available in the public domain in multiple versions. Primary studies need to explicitly state if duplicate data have been included to avoid double counting of participants in evidence synthesis. Individual participant data meta-analysis of the emerging cohorts is critical to assess both differential presentation and outcomes by underlying risk factors, and to determine the differential effects of interventions to reduce the rates of complications. With the establishment of several national and global prospective cohorts, we expect the sample size of our meta-analysis to increase further in the coming months. Our living systematic review and meta-analysis with its regular search and analyses updates is ideally placed to assess the impact of new findings on the rapidly growing evidence base.

autHOr aFFiliatiOns

[1]Institute of Applied Health Research, University of Birmingham, Birmingham, UK

[2]WHO Collaborating Centre for Global Women's Health, Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK

[3]Clinical Biostatistics Unit, Hospital Universitario Ramón y Cajal (IRYCIS), Madrid, Spain

[4]CIBER Epidemiology and Public Health (CIBERESP), Madrid, Spain

[5]UNDP/UNFPA/UNICEF/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP), Department of Sexual and Reproductive Health and Research, World Health Organization, Geneva, Switzerland

[6]Birmingham Medical School, University of Birmingham, Birmingham, UK

[7]Division of Birth Cohort Study, Guangzhou Women and Children's Medical Centre, Guangzhou Medical University, Guangzhou, China

[8]Department of Woman and Child Health Care, Guangzhou Women and Children's Medical Centre, Guangzhou Medical University, Guangzhou, China

[9]Department of Obstetrics and Gynaecology, Guangzhou Women and Children's Medical Centre, Guangzhou Medical University, Guangzhou, China

[10]Netherlands Satellite of the Cochrane Gynaecology and Fertility Group, Amsterdam University Medical Centre, Amsterdam, Netherlands

[11]Department of Obstetrics and Gynaecology, Amsterdam University Medical Centre, Amsterdam, Netherlands

[12]Blizard Institute, Queen Mary University of London, London, UK

[13]Barts Health NHS Trust, London, UK

[14]St George's, University of London, London, UK

[15]Elizabeth Glaser Paediatric AIDS Foundation, Washington DC, USA

[16]Women's Health Research Unit, Queen Mary University of London, London, UK

[17]Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK

Other members of the PregCOV-19 Living Systematic Review Consortium are Uma Ram, Ajith S Nair, Pura Rayco-Solon, and Hector Pardo-Hernandez. The Cochrane Gynaecology and Fertility Group thank Marijke Strikwerda and Bethany Clark for help with searches and data extraction. The PregCOV-19 Living Systematic Review Group would also like to thank Katie's Team for its contribution towards the development and reporting of this work, James Thomas from the EPPI-Centre for helping with search updates, and the PregCOV-19 Living Systematic Review steering committee members, Pisake Lumbiganon, Carolina Carvalho Ribeiro do Valle, Clare Whitehead, David Lissauer, Joao Paulo Souza, and Marian Knight, who provided guidance throughout.

Contributors: JA and ES are joint first authors. ST, MB, and JA conceptualised the study. MY, SC, LD, TK, ACL, AD, DZ, RB, SL, XQ, and MYuan selected the studies. JA, ES, MY, LD, DZ, XQ, and MYuan extracted the data. JZ conducted the analyses. All coauthors contributed to the writing of the manuscript and approved the final version. ST, JA, ES, and JZ are the guarantors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted

Funding: The project was partially funded by the World Health Organization.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare:

partial funding by the World Health Organization; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required.

Data sharing: No additional data available.

The corresponding author (ST) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been disclosed.

Dissemination to participants and related patient and public communities: The PregCov-19 LSR Group will disseminate the findings through a dedicated website (www.birmingham.ac.uk/research/who-collaborating-centre/pregcov/index.aspx) and social media.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;395:497-506. doi:10.1016/S0140-6736(20)30183-5
2 World Health Organization (WHO). Coronavirus disease. (COVID-19) Pandemic, https://www.who.int/emergencies/diseases/novel-coronavirus-2019 (accessed 7 May 2020)
3 Cabinet Office. Guidance. Staying alert and safe (social distancing). Coronavirus (COVID-19) Guidance and support. Updated 22 May 2020 https://www.gov.uk/government/publications/staying-alert-and-safe-social-distancing/staying-alert-and-safe-social-distancing (accessed 24 May 2020).
4 RCOG. (COVID-19) Infection in Pregnancy, https://www.rcog.org.uk/en/guidelines-research-services/guidelines/coronavirus-pregnancy/
5 Zaigham M, Andersson O. Maternal and perinatal outcomes with COVID-19: A systematic review of 108 pregnancies. *Acta Obstet Gynecol Scand* 2020;99:823-9. doi:10.1111/aogs.13867
6 Parazzini F, Bortolus R, Mauri PA, Favilli A, Gerli S, Ferrazzi E. Delivery in pregnant women infected with SARS-CoV-2: A fast review. *Int J Gynaecol Obstet* 2020;150:41-6. doi:10.1002/ijgo.13166
7 Elshafeey F, Magdi R, Hindi N, et al. A systematic scoping review of COVID-19 during pregnancy and childbirth. *Int J Gynaecol Obstet* 2020;150:47-52. doi:10.1002/ijgo.13182
8 Di Mascio D, Khalil A, Saccone G, et al. Outcome of coronavirus spectrum infections (SARS, MERS, COVID-19) during pregnancy: a systematic review and meta-analysis. *Am J Obstet Gynecol MFM* 2020;2:100107.
9 Della Gatta AN, Rizzo R, Pilu G, Simonazzi G. Coronavirus disease 2019 during pregnancy: a systematic review of reported cases. *Am J Obstet Gynecol* 2020;223:36-41. doi:10.1016/j.ajog.2020.04.013
10 Cheruiyot I, Henry BM, Lippi G. Is there evidence of intra-uterine vertical transmission potential of COVID-19 infection in samples tested by quantitative RT-PCR?*Eur J Obstet Gynecol Reprod Biol* 2020;249:100-1. doi:10.1016/j.ejogrb.2020.04.034
11 Gajbhiye R, Modi D, Mahale S. Pregnancy outcomes, Newborn complications and Maternal-Fetal Transmission of SARS-CoV-2 in women with COVID-19: A systematic review of 441 cases. medRxiv [Preprint]. 2020. doi.org/10.1101/2020.04.11.20062356
12 Murad MH, Sultan S, Haffar S, Bazerbachi F. Methodological quality and synthesis of case series and case reports. *BMJ Evid Based Med* 2018;23:60-3. doi:10.1136/bmjebm-2017-110853
13 Breslin N, Baptiste C, Gyamfi-Bannerman C, et al. Coronavirus disease 2019 infection among asymptomatic and symptomatic pregnant women: Two weeks of confirmed presentations to an affiliated pair of New York City hospitals. *Am J Obstet Gynecol MFM* 2020;2:100118.
14 Vintzileos WS, Muscat J, Hoffmann E, et al. Screening all pregnant women admitted to labor and delivery for the virus responsible for coronavirus disease 2019. *Am J Obstet Gynecol* 2020;223:284-6.
15 Xu L, Yang Q, Shi H, et al. Clinical presentations and outcomes of SARS-CoV-2 infected pneumonia in pregnant women and health status of their neonates. *Sci Bull (Beijing)* 2020;65:1537-42.
16 Blitz MJ, Grunebaum A, Tekbali A, et al. Intensive care unit admissions for pregnant and nonpregnant women with coronavirus disease 2019. *Am J Obstet Gynecol* 2020;223:290-1.
17 Released by National Health Commission & National Administration of Traditional Chinese Medicine on 3 March 2020. Trial Version 7. *Chin Med J (Engl)* 2020;133:1087-95.
18 Allotey J, Bonet M, Zamora J, et al. COVID-19 in pregnant women: a Living Systematic Review on prevalence, presentation, prognosis and treatment. PROSPERO 2020 CRD42020178076. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020178076.
19 Living Overview of the Evidence. (LOVE) Platform. app.iloveevidence.com. (accessed 8 July 2020)
20 World Health Organization. Global surveillance for COVID-19 caused by human infection with COVID-19 virus: interim guidance, 20 March 2020. World Health Organization. https://apps.who.int/iris/handle/10665/331506. 2020.
21 Dekkers OM, Egger M, Altman DG, Vandenbroucke JP. Distinguishing case series from cohort studies. *Ann Intern Med* 2012;156:37-40. doi:10.7326/0003-4819-156-1-201201030-00006
22 Wells G. Proceedings or the Third Symposium on Systematic Reviews beyond the Basics. SBOD. Improving Quality and Impact; The Newcastle–Ottawa Scale (NOS) for Assessing the Quality of non-randomised Studies in Meta-analysis. July 3-5 2000, Oxford.
23 Hoy D, Brooks P, Woolf A, et al. Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement. *J Clin Epidemiol* 2012;65:934-9. doi:10.1016/j.jclinepi.2011.11.014
24 Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Stat Med* 2000;19:3127-31. doi:10.1002/1097-0258(20001130)19:22<3127::AID-SIM784>3.0.CO;2-M
25 Cheng B, Jiang T, Zhang L, et al. Clinical Characteristics of Pregnant Women with Coronavirus Disease 2019 in Wuhan, China. *SSRN* 2020; https://ssrn.com/abstract=3555240
26 Ellington S, Strid P, Tong VT, et al. Characteristics of Women of Reproductive Age with Laboratory-Confirmed SARS-CoV-2 Infection by Pregnancy Status - United States, January 22-June 7, 2020. *MMWR Morb Mortal Wkly Rep* 2020;69:769-75. doi:10.15585/mmwr.mm6925a1
27 Liu F, Liu H, Li J, Hou L, Lan W, Wang D. Clinico-Radiological Features and Outcomes in Pregnant Women with COVID-19: Compared with Age-Matched Non-Pregnant Women. *SSRN* 2020. https://ssrn.com/abstract=3556647
28 Mohr-Sasson A, Chayo J, Bart Y, et al. Laboratory characteristics of pregnant compared to non-pregnant women infected with SARS-CoV-2. *Arch Gynecol Obstet* 2020; published online 22 June. doi:10.1007/s00404-020-05655-7
29 Qiancheng X, Jian S, Lingling P, et al, sixth batch of Anhui medical team aiding Wuhan for COVID-19. Coronavirus disease 2019 in pregnancy. *Int J Infect Dis* 2020;95:376-83. doi:10.1016/j.ijid.2020.04.065
30 Wang Z, Wang Z, Xiong G. Clinical characteristics and laboratory results of pregnant women with COVID-19 in Wuhan, China. *Int J Gynaecol Obstet* 2020. doi:10.1002/ijgo.13265
31 Wei L, Gao X, Chen S, et al. Clinical Characteristics and Outcomes between Pregnant and Non-Pregnant Women with Coronavirus Disease 2019: A Retrospective Cohort Study. SSRN2020. https://ssrn.com/abstract=3569858
32 Yin M, Zhang L, Deng G, et al. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infection During Pregnancy In China: A Retrospective Cohort Study. medRxiv [Preprint]. 2020. doi.org/10.1101/2020.04.07.20053744.
33 Campbell KH, Tornatore JM, Lawrence KE, et al. Prevalence of SARS-CoV-2 Among Patients Admitted for Childbirth in Southern Connecticut. *JAMA* 2020. doi:10.1001/jama.2020.8904
34 Knight M, Bunch K, Vousden N, et al. Characteristics and outcomes of pregnant women hospitalised with confirmed SARS-CoV-2 infection in the UK: a national cohort study using the UK Obstetric Surveillance System (UKOSS). 2020.doi.org/10.1101/2020.05.08.20089268.
35 Li N, Han L, Peng M, et al. Maternal and neonatal outcomes of pregnant women with COVID-19 pneumonia: a case-control study. *Clin Infect Dis* 2020;ciaa352. doi:10.1093/cid/ciaa352
36 Liao J, He X, Gong Q, Yang L, Zhou C, Li J. Analysis of vaginal delivery outcomes among pregnant women in Wuhan, China during the COVID-19 pandemic. *Int J Gynaecol Obstet* 2020;150:53-7. doi:10.1002/ijgo.13188
37 Khalil A, Hill R, Ladhani S, Pattisson K, O'Brien P. Severe acute respiratory syndrome coronavirus 2 in pregnancy: symptomatic pregnant women are only the tip of the iceberg. *Am J Obstet Gynecol* 2020;223:296-7. doi:10.1016/j.ajog.2020.05.005
38 Khan S, Jun L, Nawsherwan, et al. Association of COVID-19 with pregnancy outcomes in health-care workers and general women. *Clin Microbiol Infect* 2020;26:788-90. doi:10.1016/j.cmi.2020.03.034
39 Nie R, Wang S-s, Yang Q, et al. Clinical features and the maternal and neonatal outcomes of pregnant women with coronavirus disease 2019. medRxiv [Preprint]. 2020. doi.org/10.1101/2020.03.22.20041061.
40 Yan J, Guo J, Fan C, et al. Coronavirus disease 2019 in pregnant women: a report based on 116 cases. *Am J Obstet Gynecol* 2020;223:P111.E1-111.E14. doi:10.1016/j.ajog.2020.04.014

41 Zeng L, Xia S, Yuan W, et al. Neonatal Early-Onset Infection With SARS-CoV-2 in 33 Neonates Born to Mothers With COVID-19 in Wuhan, China. *JAMA Pediatr* 2020. doi:10.1001/jamapediatrics.2020.0878
42 Liu F, Lan W-s, Gan Q. Clinical and CT features of coronavirus disease in pregnant women. *Radiol Practice* 2020;35:417-20.
43 Lokken EM, Walker CL, Delaney S, et al. Clinical Characteristics of 46 Pregnant Women with a SARS-CoV-2 Infection in Washington State. *Am J Obstet Gynecol* 2020. doi:10.1016/j.ajog.2020.05.031
44 Dória M, Peixinho C, Laranjo M, Mesquita Varejão A, Silva PT. Covid-19 during pregnancy: A case series from an universally tested population from the north of Portugal. *Eur J Obstet Gynecol Reprod Biol* 2020;250:261-2. doi:10.1016/j.ejogrb.2020.05.029
45 Yang H, Hu B, Zhan S, Yang LY, Xiong G. Effects of SARS-CoV-2 infection on pregnant women and their infants: A retrospective study in Wuhan, China. *Arch Pathol Lab Med* 2020; published online 18 May.
46 Pereira A, Cruz-Melguizo S, Adrien M, Fuentes L, Marin E, Perez-Medina T. Clinical course of coronavirus disease-2019 in pregnancy. *Acta Obstet Gynecol Scand* 2020;99:839-47. doi:10.1111/aogs.13921
47 Fox NS, Melka S. COVID-19 in Pregnant Women: Case Series from One Large New York City Obstetrical Practice. *Am J Perinatol* 2020;37:1002-4. doi:10.1055/s-0040-1712529
48 Savasi VM, Parisi F, Patanè L, et al. Clinical Findings and Disease Severity in Hospitalized Pregnant Women With Coronavirus Disease 2019 (COVID-19). *Obstet Gynecol* 2020;136:252-8. doi:10.1097/AOG.0000000000003979
49 Zeng Y, Lin L, Yan Q, et al. Update on clinical outcomes of women with COVID-19 during pregnancy. *Int J Gynaecol Obstet* 2020;150:264-6. doi:10.1002/ijgo.13236
50 Liu P, Zheng J, Yang P, et al. The immunologic status of newborns born to SARS-CoV-2-infected mothers in Wuhan, China. *J Allergy Clin Immunol* 2020;146:101-9.
51 Ceulemans D, Thijs I, Schreurs A, et al. Screening for COVID-19 at childbirth: is it effective?*Ultrasound Obstetr Gynecol* 2020;56:113-4.
52 Duffy CR, Hart JM, Modest AM, et al. Lymphopenia and Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infection Among Hospitalized Obstetric Patients. *Obstet Gynecol* 2020;136:229-31. doi:10.1097/AOG.0000000000003984
53 Andrikopoulou M, Madden N, Wen T, et al. Symptoms and Critical Illness Among Obstetric Patients With Coronavirus Disease 2019 (COVID-19) Infection. *Obstet Gynecol* 2020;136:291-9. doi:10.1097/AOG.0000000000003996
54 Servei Català de la Salut. Àrea de Sistemes d'Informació. Informació visualitzada a les 06: 11h del divendres 29 de maig de 2020.
55 Kayem G, Lecarpentier E, Deruelle P, et al. A snapshot of the Covid-19 pandemic among pregnant women in France. *J Gynecol Obstet Hum Reprod* 2020;49:101826. doi:10.1016/j.jogoh.2020.101826
56 Knight M, Bunch K, Vousden N, et al, UK Obstetric Surveillance System SARS-CoV-2 Infection in Pregnancy Collaborative Group. Characteristics and outcomes of pregnant women admitted to hospital with confirmed SARS-CoV-2 infection in UK: national population based cohort study. *BMJ* 2020;369:m2107. doi:10.1136/bmj.m2107
57 Martínez-Perez O, Vouga M, Cruz Melguizo S, et al. Association Between Mode of Delivery Among Pregnant Women With COVID-19 and Maternal and Neonatal Outcomes in Spain. *JAMA* 2020. doi:10.1001/jama.2020.10125
58 Profile of pregnant and children and adolescents with COVID-19. Brazilian Ministry of Health Covid-19 Bulletin. No 17. Epidemic week 21. 17-23 May. https://www.saude.gov.br/images/pdf/2020/May/29/2020-05-25---BEE17---Boletim-do-COE.pdf
59 Yuan L, Zhang J, Dong X. Clinical characteristics and prognosis analysis of COVID-19 infection in 28 pregnancy women. *Xinjiang Yike Daxue Xuebao* 2020;43.
60 Griffin I, Benarba F, Peters C, et al. The Impact of COVID-19 Infection on Labor and Delivery, Newborn Nursery, and Neonatal Intensive Care Unit: Prospective Observational Data from a Single Hospital System. *Am J Perinatol* 2020;37:1022-30. doi:10.1055/s-0040-1713416
61 Sentilhes L, De Marcillac F, Jouffrieau C, et al. COVID-19 in pregnancy was associated with maternal morbidity and preterm birth. *Am J Obstet Gynecol* 2020.
62 Cosma S, Carosso A, Cusato J, et al. COVID-19 and first trimester spontaneous abortion: a case-control study of 225 pregnant patients. medRxiv [Preprint]. 2020. doi.org/10.1101/2020.06.19.20135749
63 Crovetto F, Crispi F, Llurba E, Figueras F, Gomez-Roig MD, Gratacos E. Seroprevalence and clinical spectrum of SARS-Cov-2 infection in the first versus third trimester of pregnancy. medRxiv [Preprint]. 2020 doi.org/10.1101/2020.06.17.20134098
64 Khoury R, Bernstein PS, Debolt C, et al. Characteristics and Outcomes of 241 Births to Women With Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infection at Five New York City Medical Centers. *Obstet Gynecol* 2020;136:273-82. doi:10.1097/AOG.0000000000004025
65 Maraschini A, Corsi E, Salvatore MA, Donati S. Coronavirus and birth in Italy: results of a national population-based cohort study. medRxiv [Preprint]. 2020. doi.org/10.1101/2020.06.11.20128652
66 NVOG: Nederlandse Vereniging voor Obstetrie en Gynaecologie. Update registratie COVID-19 positieve zwangeren in NethOSS Geplaatst op 19 June 2020.
67 Pierce-Williams RAM, Burd J, Felder L, et al. Clinical course of severe and critical COVID-19 in hospitalized pregnancies: a US cohort study. *Am J Obstet Gynecol MFM* 2020:100134.
68 Dong Y, Mo X, Hu Y, Tong S. Epidemiological and Transmission Patterns of Pregnant Women with 2019 Coronavirus Disease in China. SSRN2020. doi:10.2139/ssrn.3551330
69 London V, McLaren RJr, Atallah F, et al. The Relationship between Status at Presentation and Outcomes among Pregnant Women with COVID-19. *Am J Perinatol* 2020;37:991-4. doi:10.1055/s-0040-1712164
70 Qadri F, Mariona F.Pregnancy affected by SARS-CoV-2 infection: a flash report from Michigan. *J Matern Fetal Neonatal Med* 2020;20:1-3.
71 Lumbreras-Marquez MI, Campos-Zamora M, Lizaola-Diaz de Leon H, Farber MK. Maternal mortality from COVID-19 in Mexico. *Int J Gynaecol Obstet* 2020;150:266-7. doi:10.1002/ijgo.13250
72 Romagano MP, Guerrero K, Spillane N, et al. Perinatal outcomes in critically ill pregnant women with coronavirus disease 2019. *Am J Obstetr Gynecol MFM* 2020.
73 Giannini A, Mantovani A, Vezzoli C, Franchini D, Finazzi P. Lung ultrasound for pregnant women admitted to ICU for Covid-19 pneumonia. *Minerva Anestesiol* 2020.
74 Blitz MJ, Rochelson B, Minkoff H, et al. Maternal mortality among women with coronavirus disease 2019 admitted to the intensive care unit. *Am J Obstet Gynecol* 2020;S0002-9378(20)30636-0.
75 Perlman J, Oxford C, Chang C, Salvatore C, Di Pace J. Delivery Room Preparedness and Early Neonatal Outcomes During COVID-19 Pandemic in New York City. *Pediatrics* 2020;146:e20201567. doi:10.1542/peds.2020-1567
76 Cohen J, Vignaux O, Jacquemard F. Covid-19 in pregnant women: General data from a French National Survey. *Eur J Obstet Gynecol Reprod Biol* 2020;251:267-8. doi:10.1016/j.ejogrb.2020.06.002
77 Docherty AB, Harrison EM, Green CA, et al, ISARIC4C investigators. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ* 2020;369:m1985. doi:10.1136/bmj.m1985
78 Guan WJ, Ni ZY, Hu Y, et al, China Medical Treatment Expert Group for Covid-19. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med* 2020;382:1708-20. doi:10.1056/NEJMoa2002032
79 Richardson S, Hirsch JS, Narasimhan M, et al, and the Northwell COVID-19 Research Consortium. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. *JAMA* 2020. doi:10.1001/jama.2020.6775
80 Garg S, Kim L, Whitaker M, et al. Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019 - COVID-NET, 14 States, March 1-30, 2020. *MMWR Morb Mortal Wkly Rep* 2020;69:458-64. doi:10.15585/mmwr.mm6915e3
81 Collin J, Byström E, Carnahan A, Ahrne M. Public Health Agency of Sweden's Brief Report: Pregnant and postpartum women with severe acute respiratory syndrome coronavirus 2 infection in intensive care in Sweden. *Acta Obstet Gynecol Scand* 2020;99:819-22. doi:10.1111/aogs.13901
82 Mendoza M, Garcia-Ruiz I, Maiz N, et al. Pre-eclampsia-like syndrome induced by severe COVID-19: a prospective observational study. *BJOG* 2020; published online 1 June.

Supplementary information: appendices 1-9

11.2 BMJ maternal risk factors appendices

Appendix 6a: Maternal factors associated with COVID-19 in pregnant (and recently pregnant women)

Risk factor	No. of studies	Total No. of women	COVID-19 present n/N	COVID-19 absent n/N	OR(95% CI)	12
Age 35yrs	25	14865	2522*	12343*	0.94 (0.737; 1.196)	0%
Parity 1	14	12021	1305/6693	790/5328	1.00 (0.822; 1.207)	51%
BMI 30	12	11735	1977*	9758*	1.29 (0.974; 1.694)	93%
Non-Caucasian vs Caucasian	11	8691	540/4527	577/4164	1.66 (1.012; 2.718)	87%
Multiple pregnancy	8	5231	28/112	1485/5119	0.93 (0.563; 1.525)	12%
Gestation 28 w	2	1487	267/743	372/744	0.86 (0.647; 1.148)	0%
Smoking	14	11687	81/678	1756/11009	0.67 (0.384; 1.151)	62%
Any co-morbidity	10	8811	342/2302	668/6509	1.21 (0.981; 1.502)	23%
Chronic hypertension	13	7744	79/365	1932/7379	1.40 (0.850; 2.315)	48%
Pre-existing diabetes	12	8000	49/257	1972/7743	0.98 (0.534; 1.784)	52%
Asthma	9	5620	113/276	1624/5344	0.83 (0.407; 1.690)	73%
Support person positive	3	1250	418/533	232/717	4.492 (0.629; 32.085)	93%
Gestational diabetes	10	3298	113/492	1059/5271	1.01 (0.640; 1.591)	62%

Appendix 6b. Comparison of characteristics of pregnant (and recently pregnant) women with COVID-19 vs non-pregnant reproductive aged women with COVID-19

Risk factor	No. of studies	Total No. of women	Pregnant (and recently pregnant) women with COVID-19 n/N	Non-pregnant reproductive aged women with COVID-19 n/N	OR(95% CI)	12
Age 35yrs	7	462366	30546*	431619*	0.41 (0.245; 0.683)	72%
BMI 30	2	461897	30446*	431451*	1.98 (1.737; 2.255)	0%
Non-Caucasian vs Caucasian		461825	23609/30415	294584/431410	1.61 (1.567; 1.656)	**NE**
Any co-morbidity	2	461,936	6431/30446	104706/431490	0.94 (0.563; 1.577)	34%
Chronic hypertension	3	462087	229/30532	5523/431555	0.82 (0.316; 2.117)	57%
Pre-existing diabetes	5	462262	616/30594	6436/431668	1.35 (1.238; 1.464)	0%
Asthma	2	262	7/108	16/154	0.62 (0.182; 2.094)	14%
Support person positive	1	43	4/17	19/26	0.11 (0.027; 0.467)	**NE**

**Includes one or more studies with continuous measurement of risk factor*

Appendix 7. Clinical manifestations of coronavirus disease (COVID-19) in pregnant and recently pregnant women with suspected or confirmed disease

Clinical manifestations	Subgroup	Studies	Events/N(*)	Proportion (95% CI)	I-squared	Range
Symptom						
Fever	All studies	53	8033/39429	0.396 (0.306; 0.489)	99.2%	(0.050-0.778)
	Risk based NHCC	16	218/529	0.398 (0.306; 0.489)	76.6%	(0.147-0.688)
	Universal	10	199/550	0.251 (0.136; 0.385)	86.9%	(0.111-0.531)
	Symptom based	5	885/1683	0.544 (0.446; 0.640)	93.3%	(0.441-0.667)
	Not known	22	6731/36667	0.426 (0.279; 0.578)	99.6%	(0.050-0.778)
	Confirmed Covid-19	46	4829/33558	0.413 (0.319; 0.511)	98.1%	(0.050-0.778)
	Admitted	27	890/1832	0.399 (0.328; 0.472)	86.9%	(0.111-0.688)
	All	11	6786/36755	0.452 (0.250; 0.662)	99.8%	(0.109-0.778)
	Selected	14	348/798	0.380 (0.289-0.475)	83.4%	(0.050-0.711)
	Any risk	50	7972/39249	0.403 (0.310; 0.500)	99.3%	(0.109-0.778)
	High risk	2	52/160	0.314 (0.243; 0.390)	NE	(0.050-0.711)
	HIC	22	4616/33091	0.395 (0.263;0.535)	99.0%	(0.109-0.778)
	LMIC	30	3248/5950	0.400 (0.348; 0.454)	87.7%	(0.050-0.688)
Cough	All studies	53	10379/39641	0.415 (0.330; 0.504)	99.1%	(0.029-0.829)
	Risk based NHCC	16	170/529	0.325 (0.239; 0.416)	76.7%	(0.029-0.647)
	Universal	10	208/767	0.223 (0.075; 0.415)	95.9%	(0.037-0.558)
	Symptom based	5	949/1683	0.544 (0.488; 0.599)	78.3%	(0.470-0.622)
	Not known	22	9052/36662	0.541 (0.392; 0.686)	99.6%	(0.059-0.829)
	Confirmed Covid-19	45	6770/33764	0.411 (0.330; 0.495)	97.5%	(0.000-0.818)
	Admitted	26	876/2038	0.408 (0.312; 0.508)	94.1%	(0.037-0.800)
	All	12	9091/36761	0.516 (0.321; 0.708)	99.8%	(0.172-0.783)
	Selected	14	403/798	0.364 (0.218; 0.524)	94.4%	(0.029-0.829)
	Any risk	50	10256/39461	0.416 (0.329; 0.506)	99.1%	(0.029-0.800)
	High risk	2	118/160	0.751 (0.679; 0.817)	NE	(0.100-0.829)
	HIC	21	6682 /33297	0.434 (0.303; 0.569)	99.0%	(0.037-0.829)
	LMIC	31	3495/5956	0.402 (0.337; 0.468)	92.7%	(0.029-0.800)
Dyspnoea	All studies	42	5408/39014	0.212 (0.153; 0.277)	98.7%	(0.000-0.621)
	Risk based NHCC	11	42/392	0.095 (0.052; 0.148)	51.3%	(0.000-0.258)
	Universal	7	78/489	0.099 (0.033; 0.189)	78.5%	(0.048-0.238)
	Symptom based	4	410/1498	0.289 (0.189; 0.401)	94.7%	(0.155-0.409)
	Not known	20	4878/36635	0.316 (0.211; 0.430)	99.3%	(0.080-0.621)
	Confirmed Covid-19	36	3345/33192	0.200 (0.150; 0.255)	94.6	(0.000-0.659)
	Admitted	20	392/1537	0.206 (0.148; 0.270)	84.3%	(0.033-0.545)
	All	11	4780/36755	0.214 (0.104; 0.348)	99.6%	(0.064-0.556)

	Selected	11	236/722	0.226 (0.102; 0.378)	94.0%	(0.000-0.621)
	Any risk	40	5309/38854	0.194 (0.137; 0.258)	98.6%	(0.000-0.556)
	High risk	1	87/140	0.621 (0.526; 0.702)	NE	(0.621-0.621)
	HIC	18	3390/32845	0.258 (0.171; 0.354)	97.8%	(0.053-0.621)
	LMIC	23	1958/5781	0.181 (0.114; 0.259)	96.4%	(0.000-0.600)
Myalgia	All studies	22	5196/34663	0.189 (0.118; 0.271)	98.4%	(0.000-0.667)
	Risk based NHCC	7	13/184	0.058 (0.023; 0.103)	5.8%	(0.000-0.152)
	Universal	2	10/60	0.146 (0.063; 0.251)	NE	(0.073-0.368)
	Symptom based	3	122/881	0.149 (0.103; 0.202)	70.7%	(0.117-0.242)
	Not known	10	5051/33538	0.328 (0.185; 0.488)	99.3%	(0.100-0.667)
	Confirmed Covid-19	18	4096/32026	0.161 (0.122; 0.205)	86.1%	(0.000-0.667)
	Admitted	12	157/1121	0.138 (0.090; 0.194)	77.3%	(0.000-0.459)
	All	5	4937/33311	0.327 (0.134; 0.553)	99.6%	(0.126-0.667)
	Selected	5	102/231	0.202 (0.006; 0.528)	94.8%	(0.000-0.636)
	Any risk	20	5100/34503	0.161 (0.094; 0.240)	98.4%	(0.000-0.667)
	High risk	1	89/140	0.636 (0.550; 0.715)	NE	(0.636-0.636)
	HIC	7	4024/31398	0.257 (0.151; 0.379)	97.0%	(0.107-0.667)
	LMIC	14	1116/2877	0.154 (0.062; 0.271)	92.8%	(0.000-0.667)
Ageusia	All studies	10	83/776	0.143 (0.065; 0.243)	89.6%	(0.025-0.553)
	Risk based NHCC	NA	NA	NA	NA	NA
	Universal	3	14/274	0.133 (0.000; 0.403)	87.6%	(0.025-0.300)
	Symptom based	2	16/112	0.123 (0.067; 0.192)	NE	(0.045-0.283)
	Not known	5	53/390	0.156 (0.041; 0.321)	90.3%	(0.048-0.553)
	Confirmed Covid-19	9	76/707	0.161 (0.071; 0.276)	90.2%	(0.025-0.586)
	Admitted	5	40/638	0.063 (0.027; 0.113)	73.7%	(0.025-0.162)
	All	2	16/56	0.280 (0.164; 0.411)	NE	(0.283-0.300)
	Selected	3	27/82	0.253 (0.022; 0.597)	89.7%	(0.048-0.553)
	Any risk	10	83/776	0.143 (0.065; 0.243)	89.6%	(0.025-0.553)
	High risk	NA	NA	NA	NA	NA
	HIC	7	71/621	0.180 (0.064; 0.333)	92.7%	(0.025-0.553)
	LMIC	3	12/155	0.076 (0.025; 0.149)	49.1%	(0.045-0.162)
Diarrhoea	All studies	29	2236/38206	0.079 (0.058; 0.104)	93.4%	(0.000-0.529)
	Risk based NHCC	9	20/317	0.064 (0.021; 0.122)	58.3%	(0.000-0.182)
	Universal	4	12/208	0.051 (0.022; 0.089)	0.0%	(0.024-0.072)
	Symptom based	4	92/1498	0.057 (0.032; 0.088)	77.7%	(0.040-0.088)
	Not known	12	2112/36183	0.113 (0.074; 0.159)	97.2%	(0.027-0.529)
	Confirmed Covid-19	24	1620/32412	0.071 (0.050; 0.095)	78.4%	(0.015-0.545)
	Admitted	14	75/1171	0.058 (0.040; 0.078)	17.9%	(0.024-0.182)
	All	8	2069/36647	0.067 (0.042; 0.097)	96.5%	(0.032-0.333)

	Selected	7	92/388	0.119 (0.002; 0.340)	95.7%	(0.000-0.529)
	Any risk	28	2162/38066	0.063 (0.047; 0.082)	88.8%	(0.000-0.333)
	High risk	1	74/140	0.529 (0.442; 0.613)	NE	(0.529-0.529)
	HIC	12	1681/32212	0.114 (0.064; 0.175)	95.4%	(0.040-0.529)
	LMIC	16	539/5606	0.063 (0.044; 0.084)	64.0%	(0.000-0.182)
Laboratory findings						
Raised white cell count	All studies	13	159/580	0.263 (0.145; 0.401)	90.9%	(0.000-0.652)
	Risk based NHCC	6	44/264	0.215 (0.069; 0.407)	90.1%	(0.026-0.519)
	Universal	2	22/62	0.350 (0.234; 0.476)	NE	(0.268-0.524)
	Symptom based		43/66	0.652 (0.524; 0.765)	NE	(0.652-0.652)
	Not known	4	50//188	0.205 (0.089; 0.346)	63.3%	(0.000-0.312)
	Confirmed Covid-19	12	155/507	0.293 (0.183; 0.415)	87.0%	(0.046-0.652)
	Admitted	7	94/257	0.330 (0.168; 0.515)	88.6%	(0.135-0.652)
	All	2	39/131	0.278 (0.198; 0.365)	NE	(0.000-0.312)
	Selected	4	26/192	0.201 (0.030; 0.456)	90.9%	(0.026-0.357)
	Any risk	12	153/560	0.260 (0.136; 0.406)	91.6%	(0.000-0.652)
	High risk		6/20	0.300 (0.119; 0.543)	NE	(0.300-0.300)
	HIC	NA	NA	NA	NA	NA
	LMIC	13	159 /580	0.263 (0.145; 0.401)	90.9%	(0.000-0.652)
Lymphopaenia	All studies	27	659/1833	0.330 (0.255; 0.410)	90.4%	(0.000-0.900)
	Risk based NHCC	9	111/311	0.347 (0.242; 0.460)	70.9%	(0.176-0.900)
	Universal	4	48/281	0.179 (0.068; 0.324)	83.8%	(0.049-0.476)
	Symptom based	5	264/724	0.390 (0.221; 0.573)	94.8%	(0.157-0.697)
	Not known	9	236/517	0.349 (0.215; 0.495)	88.9%	(0.000-0.712)
	Confirmed Covid-19	23	599/1726	0.312 (0.228; 0.403)	92.7%	(0.049-0.712)
	Admitted	14	263/875	0.309 (0.202; 0.426)	91.2%	(0.049-0.697)
	All	4	176/602	0.408 (0.211; 0.620)	93.9%	(0.000-0.712)
	Selected	9	120/356	0.327 (0.228; 0.434)	72.6%	(0.150-0.900)
	Any risk	24	642/1762	0.343 (0.262; 0.428)	91.3%	(0.000-0.900)
	High risk	2	12/51	0.232 (0.122; 0.361)	NA	(0.150-0.290)
	HIC	9	220/798	0.278 (0.173; 0.396)	91.2%	(0.049-0.500)
	LMIC	17	283/647	0.356 (0.246; 0.474)	87.6%	(0.000-0.900)
Thrombocytopaenia	All studies	13	91/1383	0.058 (0.024; 0.103)	86.3%	(0.000-0.353)
	Risk based NHCC	6	33/250	0.097 (0.024; 0.103)	77.1%	(0.000-0.353)
	Universal	4	15/506	0.038 (0.006; 0.093)	79.6%	(0.012-0.122)
	Symptom based	2	43/627	0.038 (0.000; 0.123)	91.9%	(0.011-0.103)
	Not known	NA	NA	NA	NA	NA
	Confirmed Covid-19	12	67/1243	0.042 (0.013; 0.083)	83.5%	(0.000-0.353)
	Admitted	9	27/823	0.040 (0.011; 0.082)	78.8%	(0.000-0.353)

	All	1	40/388	0.103 (0.075; 0.138)	NE	(0.103-0.103)
	Selected	3	24/172	0.102 (0.028; 0.208)	62.5%	(0.036-0.181)
	Any risk	13	91/1383	0.058 (0.024; 0.103)	86.3%	(0.000-0.353)
	High risk	NA	NA	NA	NA	NA
	HIC	5	13/704	0.017 (0.005; 0.036)	47.4%	(0.011-0.082)
	LMIC	7	38/291	0.100 (0.036; 0.186)	72.5%	(0.000-0.353)
Abnormal liver function test	All studies	12	99/641	0.126 (0.060; 0.208)	84.8%	(0.000-0.360)
	Risk based NHCC	6	19/154	0.108 (0.031; 0.214)	66.3%	(0.000-0.294)
	Universal	2	11/219	0.049 (0.023; 0.083)	NE	(0.044-0.066)
	Symptom based	2	23/131	0.175 (0.113; 0.246)	NE	(0.156-0.204)
	Not known	2	46/137	0.328 (0.248; 0.412)	NE	(0.083-0.360)
	Confirmed Covid-19	11	90/663	0.106 (0.043; 0.189)	87.2%	(0.000-0.360)
	Admitted	10	54/488	0.117 (0.068; 0.177)	62.4%	(0.044-0.294)
	All		45/125	0.360 (0.276; 0.451)	NE	(0.360-0.360)
	Selected		0/28	0.000 (0.000; 0.123)	NE	(0.000-0.000)
	Any risk	12	99/641	0.126 (0.060; 0.208)	84.8%	(0.000-0.360)
	High risk	NA	NA	NA	NA	NA
	HIC	4	34/350	0.105 (0.041; 0.193)	79.7%	(0.044-0.204)
	LMIC	8	65/291	0.138 (0.045; 0.264)	82.5%	(0.000-0.360)
Raised procalcitonin	All studies	5	60/261	0.211 (0.001; 0.588)	96.6%	(0.000-0.968)
	Risk based NHCC	4	33/103	0.221 (0.000; 0.802)	97.4%	(0.000-0.968)
	Universal	1	27/158	0.171 (0.116; 0.239)	NE	(0.171-0.171)
	Symptom based	NA	NA	NA	NA	NA
	Not known	NA	NA	NA	NA	NA
	Confirmed Covid-19	5	59/253	0.206 (0.000; 0.595)	96.6%	(0.000-0.968)
	Admitted	4	59/233	0.269 (0.000; 0.738)	97.3%	(0.000-0.968)
	All	NA	NA	NA	NA	NA
	Selected		1/28	0.036 (0.001; 0.183)	NE	(0.036-0.036)
	Any risk	5	60/261	0.211 (0.001; 0.588)	96.6%	(0.000-0.968)
	High risk	NA	NA	NA	NA	NA
	HIC		27/158	0.171 (0.116; 0.239)	NE	(0.171-0.171)
	LMIC	4	33/103	0.221 (0.000; 0.802)	97.4%	(0.000-0.968)
Raised C-reactive protein	All studies	10	298/637	0.489 (0.362; 0.617)	89.5%	(0.100-0.714)
	Risk based NHCC	4	103/205	0.538 (0.415; 0.659)	60.1%	(0.440-0.704)
	Universal	1	15/21	0.714 80.478; 0.887)	NE	(0.714-0.714)
	Symptom based	2	64/120	0.534 (0.444; 0.623)	NE	(0.407-0.636)
	Not known	3	116/291	0.318 (0.060; 0.656)	96.5%	(0.100-0.640)
	Confirmed Covid-19	10	247/540	0.456 (0.308; 0.608)	90.9%	(0.100-0.737)
	Admitted	5	132/314	0.527 (0.312; 0.737)	92.3%	(0.233-0.714)

	All		80/125	0.640 (0.549; 0.724)	NE	(0.640-0.640)
	Selected	4	86/198	0.404 (0.237; 0.582)	79.5%	(0.100-0.607)
	Any risk	9	296/617	0.529 (0.403; 0.654)	89.0%	(0.233-0.714)
	High risk	1	2/20	0.100 (0.012; 0.317)	NE	(0.100-0.100)
	HIC	2	56/200	0.276 (0.216; 0.341)	NE	(0.233-0.407)
	LMIC	8	242/437	0.541 (0.423; 0.657)	81.1%	(0.100-0.714)
Radiological findings						
Ground glass appearance	All studies	14	338/569	0.685 (0.463; 0.872)	96.3%	(0.093-1.000)
	Risk based NHCC	10	260/367	0.736 (0.514; 0.912)	94.4%	(0.152-1.000)
	Universal	NA	NA	NA	NA	NA
	Symptom based	2	28/77	0.379 (0.272; 0.492)	NE	(0.093-1.000)
	Not known	2	50/125	0.394 (0.309; 0.482)	NE	(0.212-0.800)
	Confirmed Covid-19	11	215/397	0.621 (0.366; 0.846)	96%	(0.026-1.000)
	Admitted	7	144/225	0.749 (0.416; 0.975)	96.1%	(0.093-1.000)
	All		18/85	0.212 (0.131; 0.314)		(0.212-0.212)
	Selected	5	165/215	0.774 (0.441; 0.986)	95.1%	(0.152-1.000)
	Any risk	14	338/569	0.685 (0.463; 0.872)	96.3%	(0.093-1.000)
	High risk	NA	NA	NA	NA	NA
	HIC	2	28/77	0.379 (0.272; 0.492)	NA	(0.093-1.000)
	LMIC	12	310/492	0.698 (0.481; 0.879)	95.6%	(0.152-1.000)
CT-chest abnormality	All studies	24	694/2120	0.644 (0.471; 0.800)	98.2%	(0.024-1.000)
	Risk based NHCC	14	356/466	0.790 (0.646; 0.906)	90.8%	(0.250-1.000)
	Universal	2	22/146	0.090 (0.046; 0.145)	NA	(0.024-0.905)
	Symptom based	4	213/1149	0.397 (0.116; 0.722)	99.0%	(0.083-1.000)
	Not known	4	103/359	0.479 (0.119; 0.850)	97.6%	(0.124-0.875)
	Confirmed Covid-19	20	215/397	0.620 (0.438; 0.787)	97.1%	(0.024-1.000)
	Admitted	13	378/810	0.736 (0.501; 0.919)	97.2%	(0.185-1.000)
	All	2	54/742	0.070 (0.052; 0.090)	NA	(0.024-0.083)
	Selected	8	251/524	0.719 (0.375; 0.964)	98.1%	(0.124-1.000)
	Any risk	23	667.2076	0.645 (0.467; 0.806)	98.3%	(0.024-1.000)
	High risk	NA	NA	NA	NA	NA
	HIC	6	206/1498	0.141 (0.071; 0.228)	93.6%	(0.024-0.324)
	LMIC	18	488/622	0.813 (0.696; 0.909)	90.7%	(0.250-1.000)

N - Number of pregnant or recently pregnant women for whom manifestations were reported; CI - Confidence Interval ; CT - Computerised tomography; NHCC National Health Commission China; NA- Not available; HiC - High income Countries; LMIC - Low and Middle income Countries*

Risk based NHCC . Universal and Symptom based, Not Known=Sampling frames for detecting COVID -19; Confirmed COVID -19=Analysis restricted to women with laborato1y corifirmation of COVJD-i 9 only; Admi tted, All, Selected= Population types of women in studies; Any risk, High risk = Pregnancy risk status

Appendix 8. Prevalence of coronavirus disease (COVID-19) and pregnancy related outcomes in pregnant or recently pregnant women with suspected or confirmed disease

Outcomes	Subgroup	Studies	Events/N(*)	Proportion (95% CI)	I-squared	Range
Covid-19 related outcomes						
All-cause mortality	**All** studies	60	403/44157	0.000 (0.000; 0.005)	92.5%	(0.000-0.082)
	Risk based NHCC	13	2/526	0.000 (0.000; 0.005)	0.0%	(0.000-0.019)
	Universal	17	3/999	0.000 (0.000; 0.000)	0.0%	(0.000-0.021)
	Symptom based	12	80/5908	0.006 (0.002; 0.013)	63.8%	(0.000-0.048)
	Not known	18	318/36724	0.002 (0.000; 0.018)	97.6%	(0.000-0.082)
	Confirmed Covid-19	53	265/40052	0.0006 (0.0000; 0.0051)	90.6%	(0.000-0.127)
	Admitted	31	22/2712	0.000 (0.000; 0.002)	3.3%	(0.000-0.067)
	All	133	377/40356	0.004 (0.000; 0.020)	98.4%	(0.000-0.082)
	Selected	15	4/1045	0.000 (0.000; 0.001)	0.0%	(0.000-0.021)
	High risk	NA	NA	NA	NA	**NA**
	Any risk	59	403/44119	0.000 (0.000; 0.005)	92.6%	(0.000-0.082)
	HIC	30	58/34050	0.000 (0.000; 0.000)	0.0%	(0.000-0.012)
	LMIC	29	342/9719	0.005 (0.000; 0.015)	86.3%	(0.000-0.082)
Admission to intensive care unit	**All** studies	50	1373/41288	0.044 (0.024; 0.069)	97.5%	(0.000-0.294)
	Risk based NHCC	7	9/294	0.012 (0.000; 0.038)	31.3%	(0.000-0.069)
	Universal	15	43/1051	0.032 (0.018; 0.050)	21.4%	(0.000-0.110)
	Symptom based	11	497/5869	0.064 (0.040; 0.092)	90.0%	(0.011-0.113)
	Not known	17	824/34074	0.052 (0.01O; 0.117)	98.8%	(0.000-0.294)
	Confirmed Covid-19	47	829/39869	0.040 (0.024; 0.059)	95.2%	(0.000-0.212)
	Admitted	30	168/2999	0.041 (0.027; 0.056)	62.2%	(0.000-0.294)
	All	12	1180/37694	0.062 (0.015; 0.133)	99.4%	(0.000-0.200)
	Selected	8	25/595	0.032 (0.014; 0.054)	21.0%	(0.000-0.070)
	High risk	NA	NA	NA	NA	NA
	Any risk	49	1372/41250	0.044 (0.024; 0.069)	97.6%	(0.000-0.294)
	HIC	27	459/33978	0.043 (0.025; 0.066)	91.4%	(0.000-0.133)
	LMIC	22	870/6922	0.040 (0.018; 0.069)	91.6%	(0.000-0.294)
Severe COVID-19	**All** studies	39	633/5621	0.104 (0.065; 0.151)	94.4%	(0.000-1.000)
	Risk based NHCC	11	39/493	0.053 (0.016; 0.105)	73.0%	(0.000-0.323)
	Universal	12	114/971	0.078 (0.021; 0.159)	92.0%	(0.000-0.311)
	Symptom based	6	232/1294	0.167 (0.124; 0.215)	65.8%	(0.087-0.348)
	Not known	10	248/2863	0.168 (0.021; 0.393)	97.8%	(0.000-1.000)
	Confirmed Covid-19	36	466/2932	0.114 (0.066; 0.172)	93.8%	(0.000-1.000)
	Admitted	22	164/1303	0.090 (0.045; 0.145)	87.6%	(0.000-0.348)

	All	8	317/3431	0.107 (0.046; 0.186)	93.7%	(0.024-0.207)
	Selected	9	152/887	0.130 (0.007; 0.349)	97.8%	(0.000-1.000)
	High risk	.','	142 /599	0.465 (0.009; 0.971)	99.3%	(0.110-1.000)
	Any risk	36	491/5022	0.081 (0.053; 0.113)	88.2%	(0.000-0.348)
	HIC	20	419/2363	0.161 (0.085; 0.254)	96.2%	(0.000-1.000)
	LMIC	19	214/3258	0.051 (0.030; 0.077)	56.7%	(0.000-0.323)
Invasive ventilation	**All** studies	31	668/42026	0.027 (0.011; 0.047)	97.5%	(0.000-0.133)
	Risk based NHCC	4	2/193	0.004 (0.000; 0.024)	0.0%	(0.000-0.017)
	Universal	9	17/519	0.024 (0.008; 0.047)	32.8%	(0.000-0.077)
	Symptom based	7	276/5242	0.046 (0.030; 0.065)	80.1%	(0.000-0.093)
	Not known	11	373/36072	0.027 (0.004; 0.065)	98.6%	(0.000-0.133)
	Confirmed Covid-19	30	388/38240	0.029 (0.014; 0.047)	95.6%	(0.000-0.133)
	Admitted	17	64/1492	0.022 (0.008; 0.040)	63.0%	(0.000-0.089)
	All	10	596/40040	0.041 (0.012; 0.084)	99.2%	(0.000-0.133)
	Selected	4	8/494	0.015 (0.005; 0.029)	0.0%	(0.012-0.023)
	High risk	NA	NA	NA	NA	NA
	Any risk	31	668/42026	0.027 (0.011; 0.047)	97.5%	(0.000-0.133)
	HIC	17	190/32678	0.031 (0.011; 0.059)	93.6%	(0.000-0.133)
	LMIC	13	442/8960	0.016 (0.001; 0.043)	96.2%	(0.000-0.098)
Need for ECMO	**All** studies	13	37/33521	0.002 (0.000; 0.007)	76.0%	(0.000-0.014)
	Risk based NHCC	3	1/181	0.002 (0.000; 0.021)	0.0%	(0.000-0.009)
	Universal	NA	NA	NA	NA	NA
	Symptom based	5	16/2044	0.007 (0.004; 0.012)	0.0%	(0.000-0.010)
	Not known	5	20/31296	0.000 (0.000; 0.003)	51.7%	(0.000-0.014)
	Confirmed Covid-19	13	37/33427	0.0014 (0.0000; 0.0064)	76.5%	(0.000-0.015)
	Admitted	7	11/1887	0.003 (0.000; 0.007)	0.0%	(0.000-0.014)
	All	3	25/31420	0.004 (0.000; 0.014)	89.7%	(0.001-0.010)
	Selected	3	1/214	0.002 (0.000; 0.018)	0.0%	(0.000-0.009)
	High risk	3	3/735	0.001 (0.000; 0.007)	0.0%	(0.000-0.014)
	Any risk	10	34/32786	0.002 (0.000; 0.008)	79.2%	(0.000-0.010)
	HIC	9	34/32952	0.003 (0.000; 0.009)	81.0%	(0.000-0.014)
	LMIC	3	1/181	0.002 (0.000; 0.021)	0.0%	(0.000-0.009)
Oxygen through cannula only	**All** studies	17	261/1522	0.225 (0.115; 0.356)	96.2%	(0.022-1.000)
	Risk based NHCC	4	64/108	0.619 (0.106; 0.998)	97.0%	(0.065-1.000)
	Universal	4	29/371	0.080 (0.021; 0.168)	84.4%	(0.025-0.183)
	Symptom based	5	102/853	0.093 (0.044; 0.155)	73.3%	(0.022-0.135)
	Not known	4	66/190	0.261 (0.004; 0.684)	97.2%	(0.053-0.812)
	Confirmed Covid-19	17	248/1417	0.223 (0.116; 0.351)	95.7%	(0.025-1.000)

	Admitted	11	89/703	0.177 (0.066; 0.322)	94.7%	(0.022-1.000)
	All	1	83/617	0.135 (0.109; 0.164)	NE	(0.135-0.135)
	Selected	5	89/202	0.366 (0.038; 0.787)	97.3%	(0.053-0.879)
	High risk	1	52/64	0.812 (0.695; 0.899)	NE	(0.812-0.812)
	Any risk	16	209/1458	0.190 (0.099; 0.299)	94.8%	(0.022-1.000)
	HIC	12	189/1377	0.128 (0.054; 0.224)	94.4%	(0.022-0.812)
	LMIC	5	72/145	0.535 (0.118; 0.925)	96.7%	(0.065-1.000)
Acute respiratory distress syndrome	**All** studies	15	315/2348	0.069 (0.009; 0.165)	97.8%	(0.000-0.508)
	Risk based NHCC	1	0/17	0.000 (0.000; 0.195)	NE	(0.000-0.000)
	Universal	3	10/269	0.047 (0.000; 0.144)	50.1%	(0.029-0.133)
	Symptom based	5	17/631	0.025 (0.003; 0.061)	66.7%	(0.000-0.130)
	Not known	6	288/1431	0.128 (0.004; 0.370)	99.0%	(0.014-0.508)
	Confirmed Covid-19	13	63/1807	0.033 (0.013; 0.059)	75.3%	(0.000-0.222)
	Admitted	11	283/1762	0.060 (0.000; 0.192)	98.2%	(0.000-0.508)
	All	2	13/488	0.024 (0.011; 0.040)	NE	(0.018-0.060)
	Selected	2	19/98	0.192 (0.118; 0.278)	NE	(0.147-0.219)
	High risk	3	35/767	0.080 (0.010; 0.201)	92.5%	(0.025-0.219)
	Any risk	12	280/1581	0.065 (0.000; 0.196)	98.0%	(0.000-0.508)
	HIC	11	62/1393	0.056 (0.027; 0.094)	77.4%	(0.014-0.219)
	LMIC	3	246/567	0.095 (0.000; 0.618)	98.7%	(0.000-0.508)
Pneumonia	All studies	36	1257/7198	0.353 (0.265; 0.446)	97.9%	(0.000-1.000)
	Risk based NHCC	12	305/436	0.809 (0.567; 0.971)	96.5%	(0.000-1.000)
	Universal	10	112/955	0.092 (0.026; 0.188)	93.7%	(0.016-0.440)
	Symptom based	8	665/4646	0.268 (0.142; 0.416)	98.2%	(0.043-1.000)
	Not known	6	175/1161	0.145 (0.086; 0.215)	85.0%	(0.033-0.322)
	Confirmed Covid-19	31	1054/7114	0.252 (0.188; 0.320)	96.4%	(0.000-1.000)
	Admitted	20	535/2754	0.238 (0.141; 0.351)	97.3%	(0.000-1.000)
	All	6	471/3850	0.268 (0.100; 0.480)	97.8%	(0.024-1.000)
	Selected	9	207/550	0.604 (0.306; 0.867)	97.6%	(0.095-1.000)
	High risk	1	75/598	0.125 (0.100; 0.155)	NE	(0.125-0.125)
	Any risk	35	1182/6600	0.362 (0.266; 0.463)	97.9%	(0.000-1.000)
	HIC	19	501/3081	0.132 (0.083; 0.188)	93.6%	(0.016-0.440)
	LMIC	17	756/4117	0.655 (0.400; 0.872)	98.8%	(0.000-1.000)
Acute cardiac, renal or hepatic **IIIJUry**	All studies	13	15/2046	0.002 (0.000; 0.008)	28.3%	(0.000-0.130)
	Risk based NHCC	2	0/48	0.000 (0.000; 0.074)	NE	NA
	Universal	1	2/241	0.008 (0.001; 0.030)	NE	(0.008-0.008)
	Symptom based	4	5/540	0.010 (0.000; 0.049)	72.1%	(0.000-0.130)

	Not known	6	8/1217	0.002 (0.000; 0.009)	12.8%	(0.003-0.059)
	Confirmed Covid-19	12	14/2027	0.002 (0.000; 0.007)	21.2%	(0.000-0.130)
	Admitted	9	10/1161	0.004 (0.000; 0.016)	44.1%	(0.000-0.130)
	All	3	4/821	0.004 (0.000; 0.011)	0.0%	(0.003-0.008)
	Selected	1	1/64	0.016 (0.000; 0.084)	NE	(0.016-0.016)
	High risk	3	4/767	0.004 (0.000; 0.013)	18.6%	(0.003-0.016)
	Any risk	10	11/1279	0.003 (0.000; 0.011)	34.6%	(0.000-0.130)
	HIC	6	9/1091	0.006 (0.000; 0.021)	57.0%	(0.000-0.130)
	LMIC	6	5/567	0.002 (0.000; 0.010)	0.0%	(0.000-0.059)
Pregnancy related maternal outcomes						
Preterm birth <37 weeks	All studies	70	1406/9396	0.165 (0.143; 0.189)	79.6%	(0.000-0.571)
	Risk based NHCC	14	80/424	0.174 (0.128; 0.225)	32.0%	(0.037-0.357)
	Universal	26	247/1671	0.125 (0.095; 0.158)	63.3%	(0.000-0.375)
	Symptom based	12	345/1354	0.230 (0.182; 0.281)	73.3%	(0.000-0.400)
	Not known	18	734/5947	0.168 (0.131; 0.207)	78.2%	(0.058; 0.571)
	Confirmed Covid-19	59	1357/9064	0.168 (0.144; 0.194)	82.5%	(0.000-0.606)
	Admitted	48	644/3645	0.161 (0.134; 0.189)	73.6%	(0.000-0.571)
	All	5	136/558	0.223 (0.123; 0.340)	83.7%	(0.082-0.467)
	Selected	17	626/5193	0.158 (0.118; 0.202)	70.5%	(0.038-0.545)
	High risk	2	65/474	0.126 (0.096; 0.159)	NE	(0.123-0.450)
	Any risk	67	1333/8884	0.163 (0.140; 0.188)	79.8%	(0.000-0.571)
	HIC	40	1113/8043	0.145 (0.121; 0.171)	79.4%	(0.000-0.545)
	LMIC	29	223/1096	0.200 (0.153; 0.250)	71.0%	(0.031-0.571)
PPROM <37 weeks	All studies	18	58/993	0.050 (0.032; 0.072)	29.7%	(0.000-0.211)
	Risk based NHCC	4	15/183	0.079 (0.035; 0.137)	24.2%	(0.037-0.172)
	Universal	7	25/529	0.037 (0.016; 0.062)	21.5%	(0.000-0.100)
	Symptom based	3	7/72	0.105 (0.003; 0.285)	67.4%	(0.024-0.211)
	Not known	4	11/209	0.041 (0.014; 0.078)	0.0%	(0.031-0.133)
	Confirmed Covid-19	17	49/895	0.043 (0.026; 0.064)	20.6%	(0.000-0.211)
	Admitted	12	39/778	0.040 (0.023; 0.061)	20.0%	(0.000-0.211)
	All	2	7/44	0.157 (0.058; 0.286)	NE	(0.133-0.172)
	Selected	4	12/171	0.060 (0.025; 0.105)	0.0%	(0.031-0.167)
	High risk	1	1/32	0.031 (0.001; 0.162)	NE	(0.031-0.031)
	Any risk	17	57/961	0.052 (0.032; 0.076)	33.4%	(0.000-0.211)
	HIC	13	42/794	0.043 (0.023; 0.068)	32.4%	(0.000-0.211)
	LMIC	5	16/199	0.072 (0.037; 0.116)	0.0%	(0.037-0.172)
Spontaneous preterm birth	**All** studies	17	104/1629	0.061 (0.038; 0.088)	67.2%	(0.000-0.321)
	Risk based NHCC	3	8/153	0.049 (0.017; 0.092)	0.0%	(0.036-0.061)

	Universal	5	38/429	0.099 (0.042; 0.174)	69.5%	(0.056-0.312)
	Symptom based	5	50/866	0.058 (0.031; 0.093)	67.7%	(0.018-0.220)
	Not known	4	8/181	0.045 (0.000; 0.161)	80.4%	(0.000-0.132)
	Confirmed Covid-19	16	100/1566	0.062 (0.038; 0.090)	69.1%	(0.000-0.312)
	Admitted	10	75/1052	0.076 (0.044; 0.115)	68.9%	(0.018-0.312)
	All	2	14/354	0.032 (0.015; 0.053)	NE	(0.000-0.054)
	Selected	5	15/223	0.063 (0.032; 0.101)	0.0%	(0.036-0.132)
	High risk	1	2/32	0.062 (0.008; 0.208)	NE	(0.062-0.062)
	Any risk	15	97/1559	0.058 (0.034; 0.086)	69.5%	(0.000-0.312)
	HIC	9	71/1054	0.068 (0.041; 0.100)	62.2%	(0.018-0.220)
	LMIC	7	19/318	0.061 (0.010; 0.142)	77.1%	(0.000-0.312)
Caesarean section	**All** studies	75	3760/9725	0.535 (0.486; 0.583)	93.4%	(0.000-1.000)
	Risk based NHCC	13	344/426	0.824 (0.753; 0.885)	64.5%	(0.650-1.000)
	Universal	28	694/1900	0.399 (0.336; 0.463)	84.6%	(0.000-1.000)
	Symptom based	13	765/1367	0.562 (0.479; 0.644)	87.0%	(0.333-0.941)
	Not known	21	1957/6032	0.475 (0.411; 0.540)	88.7%	(0.260-0.850)
	Confirmed Covid-19	66	3485/9402	0.492 (0.446; 0.537)	91.6%	(0.000-1.000)
	Admitted	48	1568/3719	0.489 (0.429; 0.550)	91.2%	(0.000-1.000)
	All	8	321/602	0.580 (0.468; 0.688)	76.4%	(0.375-0.931)
	Selected	19	1871/5404	0.637 (0.517; 0.749)	96.0%	(0.270-1.000)
	High risk	3	192/506	0.641 (0.268; 0.939)	95.2%	(0.333-0.850)
	Any risk	71	3545/9181	0.530 (0.479; 0.580)	93.5%	(0.000-1.000)
	HIC	42	2776/8212	0.380 (0.339; 0.421)	87.4%	(0.000-1.000)
	LMIC	32	848/1256	0.735 (0.664; 0.801)	84.8%	(0.382-1.000)
Vaginal delivery	**All** studies	74	5410/9708	0.461 (0.417; 0.505)	91.8%	(0.000-1.000)
	Risk based NHCC	12	70/409	0.163 (0.119; 0.211)	29.1%	(0.000-0.279)
	Universal	28	1199/1900	0.593 (0.524; 0.660)	86.3%	(0.000-1.000)
	Symptom based	13	590/1367	0.429 (0.350; 0.510)	86.2%	(0.059-0.667)
	Not known	21	3551/6032	0.518 (0.460; 0.577)	85.5%	(0.150-0.740)
	Confirmed Covid-19	65	5361/9385	0.507 (0.466; 0.548)	89.4%	(0.000-1.000)
	Admitted	48	2130/3719	0.497 (0.435; 0.560)	91.8%	(0.000-1.000)
	All	8	272/602	0.409 (0.302; 0.521)	76.4%	(0.069-0.625)
	Selected	17	3008/5387	0.374 (0.285; 0.469)	93.3%	(0.118-0.730)
	High risk	3	313/506	0.358 (0.061; 0.730)	95.2%	(0.150-0.665)
	Any risk	70	5082/9164	0.465 (0.420; 0.511)	91.7%	(0.000-1.000)
	HIC	42	4913/8212	0.615 (0.577; 0.653)	85.3%	(0.000-1.000)
	LMIC	31	382/1239	0.252 (0.189; 0.321)	83.5%	(0.000-0.616)

Postpartum haemorrhage	**All** studies	15	91/908	0.078 (0.029; 0.143)	86.7%	(0.000-0.303)
	Risk based NHCC	1	0/29	0.000 (0.000; 0.119)	NE	(0.000-0.000)
	Universal	8	75/616	0.105 (0.030; 0.211)	90.1%	(0.031-0.303)
	Symptom based	3	9/74	0.047 (0.006; 0.114)	0.0%	(0.036-0.105)
	Not known	3	7/189	0.058 (0.000; 0.197)	76.0%	(0.014-0.182)
	Confirmed Covid-19	12	84/836	0.090 (0.032; 0.167)	89.0%	(0.014-0.303)
	Admitted	12	86/836	0.083 (0.028; 0.160)	89.1%	(0.014-0.303)
	All	1	0.29	0.000 (0.000; 0.119)	NE	(0.000-0.000)
	Selected	2	5/43	0.108 (0.024; 0.228)	NE	(0.094-0.182)
	High risk	1	3/32	0.094 (0.020; 0.250)	NE	(0.094-0.094)
	Any risk	14	88/876	0.077 (0.026; 0.146)	87.7%	(0.000-0.303)
	HIC	12	89/831	0.096 (0.035; 0.177)	89.2%	(0.014-0.303)
	LMIC	3	2/77	0.017 (0.000; 0.068)	0.0%	(0.000-0.062)
Perinatal outcomes						
Stillbirth	All studies	47	72/9020	0.002 (0.000; 0.005)	31.2%	(0.000-0.235)
	Risk based NHCC	11	0/394	0.000 (0.000; 0.009)	NE	NA
	Universal	17	13/1241	0.001 (0.000; 0.006)	0.0%	(0.000-0.125)
	Symptom based	9	25/2240	0.007 (0.003; 0.013)	14.8%	(0.000-0.058)
	Not known	10	34/5145	0.012 (0.000; 0.033)	75.5%	(0.000-0.235)
	Confirmed Covid-19	43	67/8878	0.0012 (0.0000; 0.0036)	19.2%	(0.000-0.125)
	Admitted	29	27/2621	0.002 (0.000; 0.005)	0.0%	(0.000-0.235)
	All	5	18/1191	0.008 (0.003; 0.016)	0.0%	(0.000-0.235)
	Selected	13	27/5208	0.000 (0.000; 0.004)	38.0%	(0.000-0.125)
	High risk	2	2/84	0.005 (0.000; 0.042)	NE	(0.000-0.100)
	Any risk	44	69/8898	0.001 (0.000; 0.004)	27.5%	(0.000-0.235)
	HIC	26	48/7719	0.000 (0.000; 0.001)	2.6%	(0.000-0.125)
	LMIC	20	18/913	0.007 (0.000; 0.019)	29.9%	(0.000-0.235)
Neonatal death	All studies	51	41/8263	0.000 (0.000; 0.002)	33.0%	(0.000-0.125)
	Risk based NHCC	12	1/371	0.000 (0.000; 0.007)	0.0%	(0.000-0.010)
	Universal	15	4/1200	0.000 (0.000; 0.000)	0.0%	(0.000-0.027)
	Symptom based	10	15/1288	0.004 (0.000; 0.010)	0.0%	(0.000-0.080)
	Not known	14	21/5404	0.002 (0.000; 0.0131)	66.0%	(0.000-0.125)
	Confirmed Covid-19	47	38/7973	0.0002 (0.0000; 0.0025)	33.5%	(0.000-0.125)
	Admitted	30	14/2448	0.000 (0.000; 0.001)	6.5%	(0.000-0.118)
	All	6	10/581	0.001 (0.000; 0.014)	26.2%	(0.000-0.125)
	Selected	15	17/5234	0.001 (0.000; 0.008)	36.3%	(0.000-0.048)
	High risk	NA	**NA**	NA	**NA**	**NA**

	Any risk	50	40/8225	0.000 (0.000; 0.002)	32.1%	(0.000-0.125)
	HIC	27	22/6970	0.000 (0.000; 0.000)	23.3%	(0.000-0.125)
	LMIC	23	14/1036	0.003 (0.000; 0.010)	0.0%	(0.000-0.118)
Admission to neonatal unit	**All** studies	41	934/3323	0.328 (0.237; 0.426)	96.8%	(0.000-1.000)
	Risk based NHCC	7	76/240	0.202 (0.031; 0.453)	93.3%	(0.000-0.700)
	Universal	17	301/1033	0.368 (0.192; 0.563)	97.2%	(0.000-0.700)
	Symptom based	6	302/1089	0.379 (0.205; 0.571)	97.5%	(0.188-1.000)
	Not known	11	255/961	0.321 (0.144; 0.528)	97.4%	(0.113-0.980)
	Confirmed Covid-19	37	799/31665	0.264 (0.186; 0.349)	96%	(0.000-1.000)
	Admitted	25	540/2063	0.311 (0.195; 0.439)	97.0%	(0.000-1.000)
	All	6	235/608	0.595 (0.252; 0.895)	98.3%	(0.188-1.000)
	Selected	10	159/652	0.225 (0.120; 0.349)	90.6%	(0.000-0.700)
	High risk	2	37/126	0.278 (0.202; 0.361)	NE	(0.172-0.636)
	Any risk	38	892/3159	0.331 (0.235; 0.434)	96.9%	(0.000-1.000)
	HIC	25	557/2372	0.287 (0.200; 0.382)	95.4%	(0.000-1.000)
	LMIC	15	308/694	0.399 (0.172; 0.651)	97.6%	(0.000-1.000)
Neonatal sepsis	**All** studies	6	9/499	0.014 (0.002; 0.034)	22.6%	(0.008-0.056)
	Risk based NHCC	1	1/33	0.030 (0.001; 0.158)	NE	(0.030-0.030)
	Universal	1	1/18	0.056 (0.001; 0.273)	NE	(0.056-0.056)
	Symptom based	3	6/362	0.018 (0.000; 0.054)	53.0%	(0.008-0.043)
	Not known	1	1/86	0.012 (0.000; 0.063)	NE	(0.012-0.012)
	Confirmed Covid-19	6	9/499	0.014 (0.002; 0.034)	22.6%	(0.008-0.056)
	Admitted	2	4/87	0.041 (0.005; 0.099)	NE	(0.043-0.056)
	All	2	3/293	0.006 (0.000; 0.021)	NE	(0.008-0.028)
	Selected	2	2/119	0.014 (0.000; 0.048)	NE	(0.012-0.030)
	High risk	NA	NA	NA	NA	NA
	Any risk	6	9/499	0.014 (0.002; 0.034)	22.6%	(0.008-0.056)
	HIC	2	2/54	0.034 (0.000; 0.109)	NE	(0,028-0.056)
	LMIC	3	5/188	0.023 (0.004; 0.053)	0.0%	(0.012-0.043)
Abnormal APGAR <5	All studies	31	42/1479	0.007 (0.001; 0.019)	34.8%	(0.000-0.263)
	Risk based NHCC	7	2/225	0.002 (0.000; 0.019)	0.0%	(0.000-0.111)
	Universal	12	12/636	0.000 (0.000; 0.019)	0.0%	(0.000-0.059)
	Symptom based	5	5/149	0.016 (0.000; 0.074)	53.9%	(0.000-0.111)
	Not known	7	23/469	0.033 (0.001; 0.094)	78.9%	(0.000-0.263)
	Confirmed Covid-19	26	40/1297	0.008 (0.001; 0.022)	36.8%	(0.000-0.263)
	Admitted	21	24/975	0.005 (0.000; 0.014)	0.0%	(0.000-0.071)
	All	2	4/47	0.068 (0.006; 0.169)	NE	(0.000-0.111)

	Selected	8	14/457	0.014 (0.000; 0.055)	73.8%	(0.000-0.263)
	High risk	NA	NA	NA	NA	NA
	Any risk	30	32/1441	0.004 (0.000; 0.011)	0.0%	(0.000-0.111)
	HIC	18	29/1072	0.005 (0.000; 0.014)	0.0%	(0.000-0.111)
	LMIC	13	13/407	0.011 (0.000; 0.042)	60.7%	(0.000-0.263)
Fetal distress	All studies	17	65/553	0.111 (0.074; 0.152)	40.2%	(0.043-0.500)
	Risk based NHCC	7	31/284	0.104 (0.069; 0.145)	0.0%	(0.043-0.182)
	Universal	1	1/17	0.059 (0.001; 0.287)	NE	(0.059-0.059)
	Symptom based	3	6/84	0.069 (0.019; 0.340)	74.1%	(0.053-0.500)
	Not known	6	27/168	0.185 (0.065; 0.340)	74.1%	(0.053-0.500)
	Confirmed Covid-19	13	44/373	0.108 (0.065; 0.159)	40.2%	(0.043-0.500)
	Admitted	11	31/247	0.113 (0.065; 0.171)	29.9%	(0.043-0.500)
	All	1	3/30	0.100 (0.021; 0.265)	NE	(0.100-0.100)
	Selected	5	31/276	0.116 (0.054; 0.196)	66.7%	(0.053-0.289)
	High risk	NA	NA	NA	NA	NA
	Any risk	16	54/515	0.096 (0.066; 0.131)	19.9%	(0.043-0.500)
	HIC	4	10/159	0.059 (0.024; 0.105)	0.0%	(0.043-0.100)
	LMIC	13	55/394	0.133 (0.087; 0.186)	39.5%	(0.043-0.500)

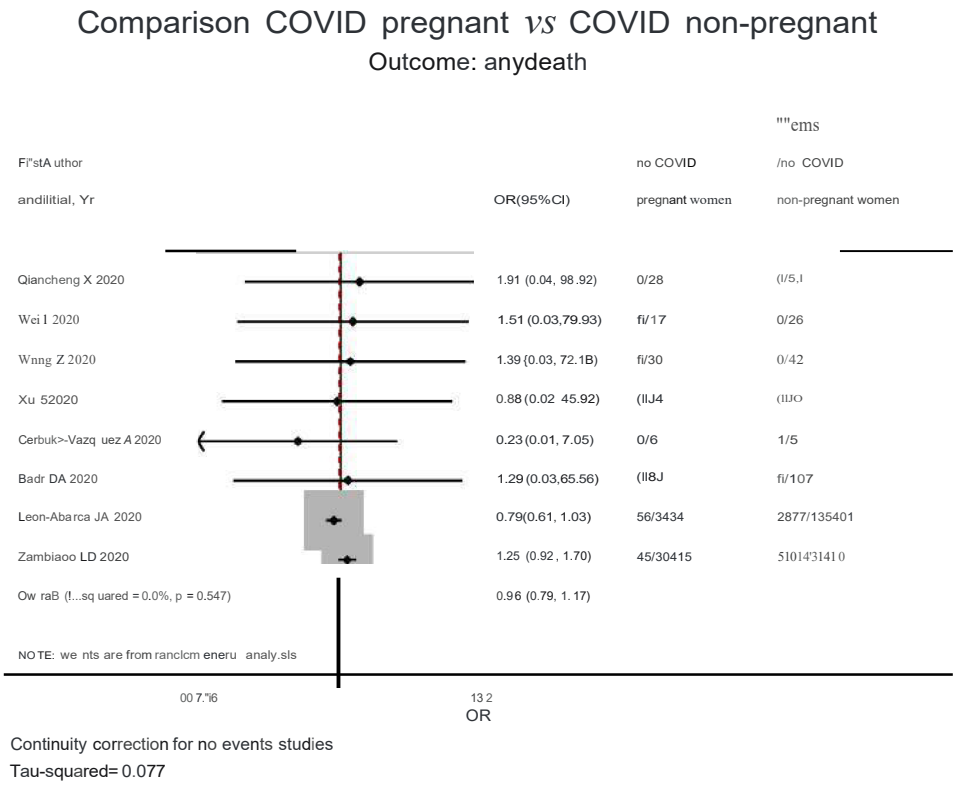
N - Number of pregnant or recently pregnant women/or COVJD-related outcomes andfor preterm birth outcomes; number of women delivered for caesarean section; number of babies born for perinatal outcomes; Cf - Confidence In terval;; ECMO - Extra corporeal membrane oxygenation; NHCC National Health Commission China; NA -Not available; NE - Not estimatable; HIC - High Inco me Countries; LMJC - Low and Middle Income Countries

Risk based NHCC, Universal and Symptom based, Not Known=Sampling frames for detecting COVID -19; Confirmed COVID -l 9=Analys is restricted to women with laborato1y co nfirmation of COVID-19 only; Adm itted, All, Selected = Population types of women in studies; Any risk, High risk = Pregnancy risk status

Appendix 9: Study-level forest plots for COVID-19, pregnancy-related maternal and perinatal outcomes

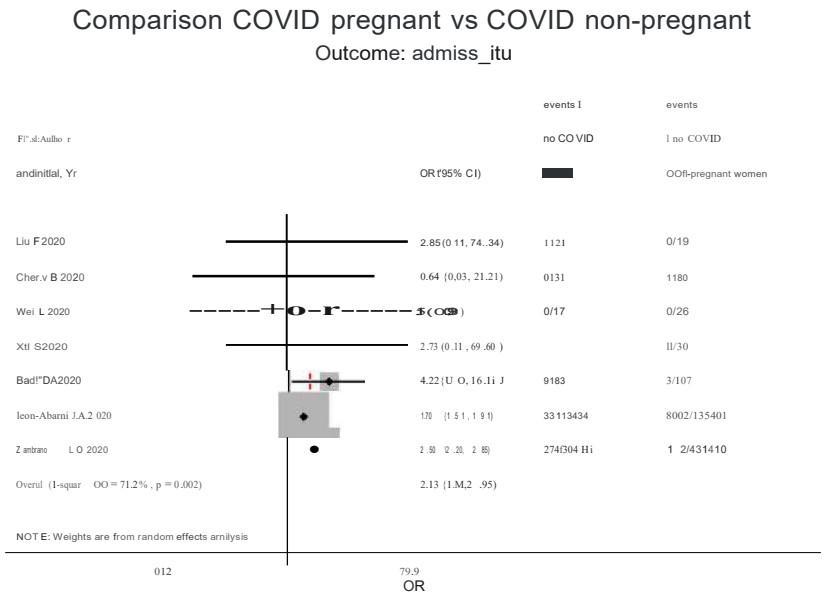
COIVD-19 related outcomes

Any Mortality



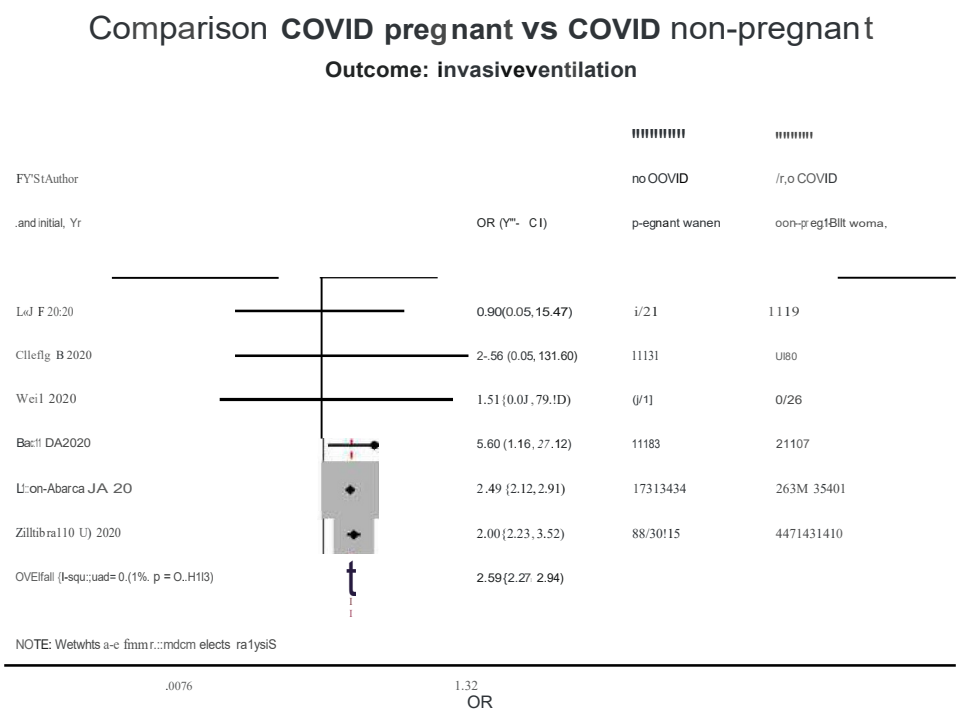
Continuity correction for no events studies
Tau-squared= 0.077

Admission to Intensive Care Unit

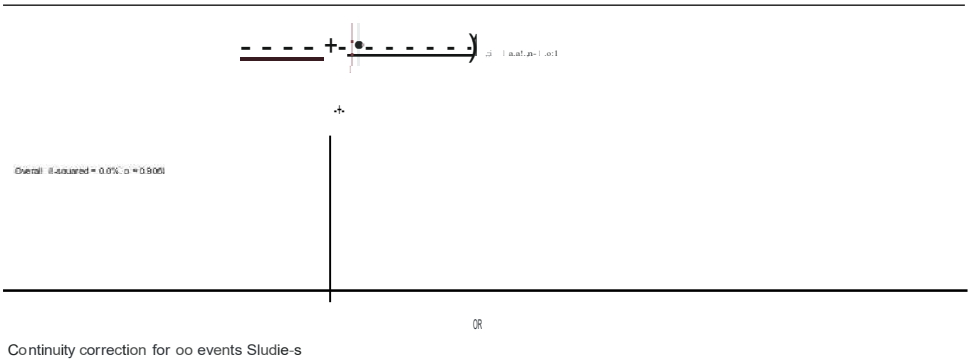


Continuity correction for no events studies
Tau-squared =- 0.061

Invasive vent ilatio n



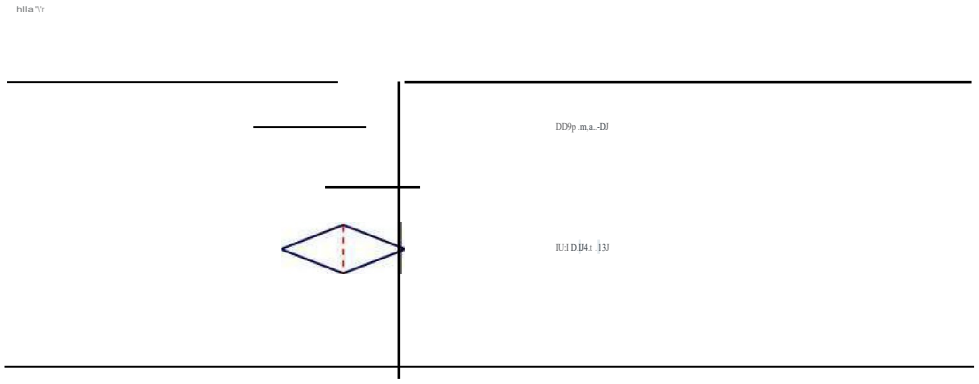
Need for ECMO



Oxygen through nas al canula

Comparison COVID pregnant vs COVID non-pregnant

Outcome: oxyg enthroughcannulaonly



OR

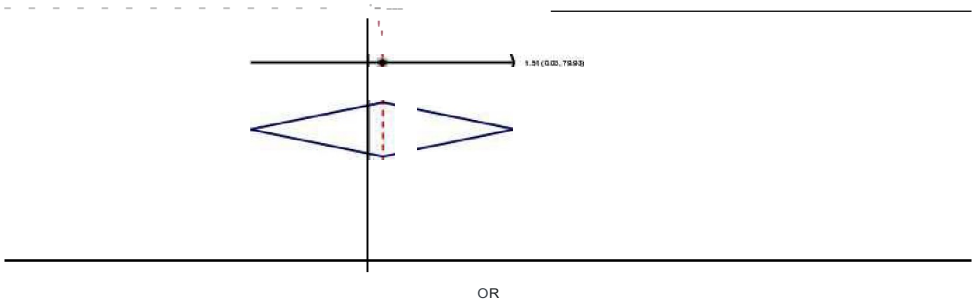
Cont inu it y correction for no events stud ies

Tau-squa red = 0.9

ARDS

Comparison COVID pregnant vs COVID non-pregnant

Outcome: ards



OR

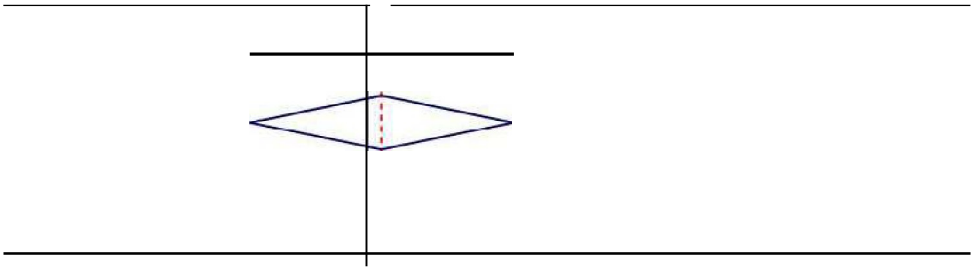
co ntinuity correctJon for no events studies

Tau-squa red:= Not estimable

Cardiac/Live r/Renal failu re

Comparison COVID pregnant vs COVID non-pregnant

Outcome: acutecardiacrenalor hepat



OR

Continuity correc on for no events studies

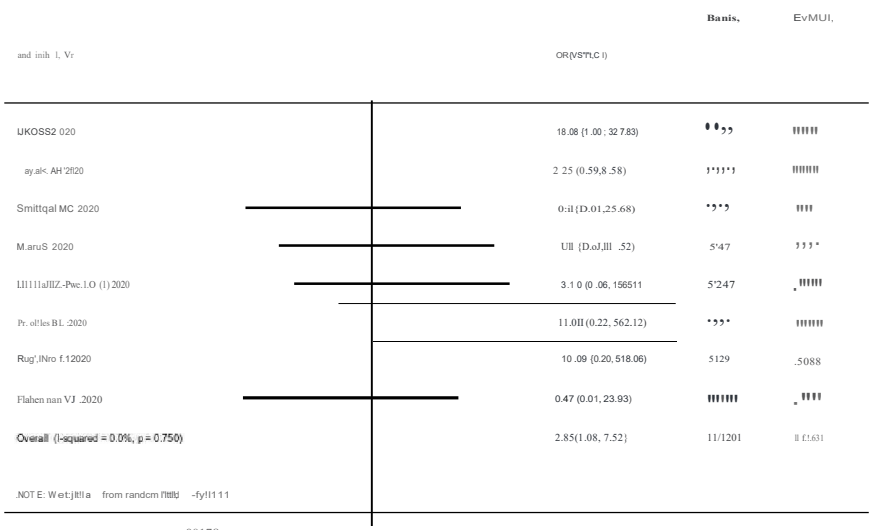
Tau-:square(t::: Not estimable

Pregnancy-related maternal outcomes

Any mortality

Comparison COVID pregnant vs Non-COVID pregnant

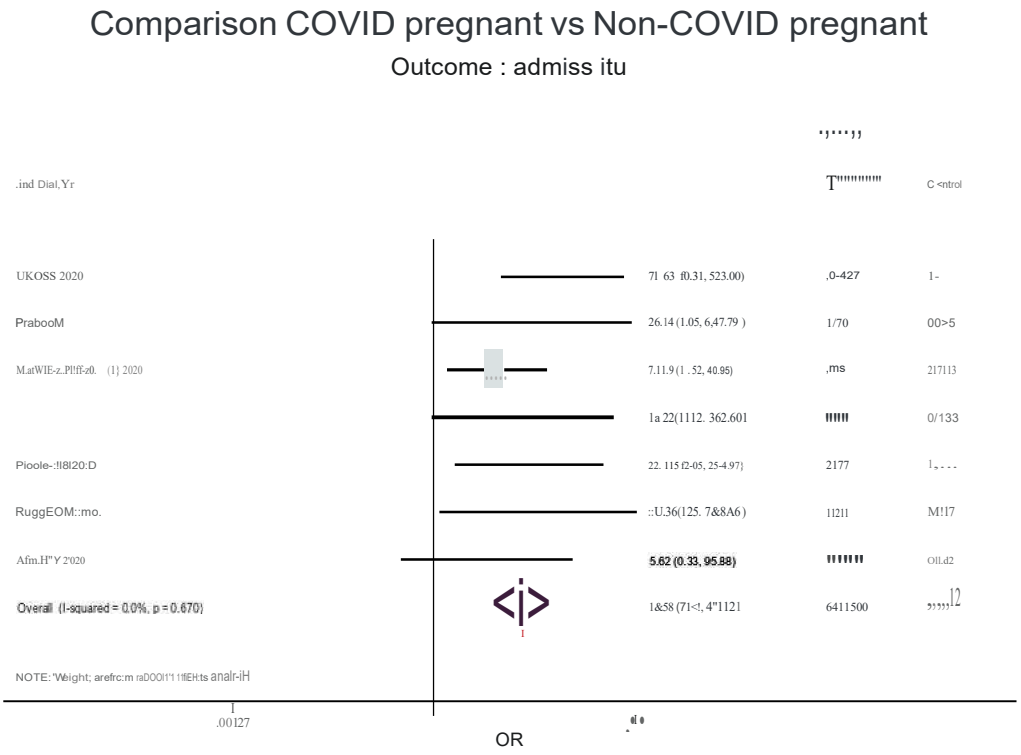
Ou tcome: anydeath



Continuity correctlon for no events studies

Tau-squared= 1.217

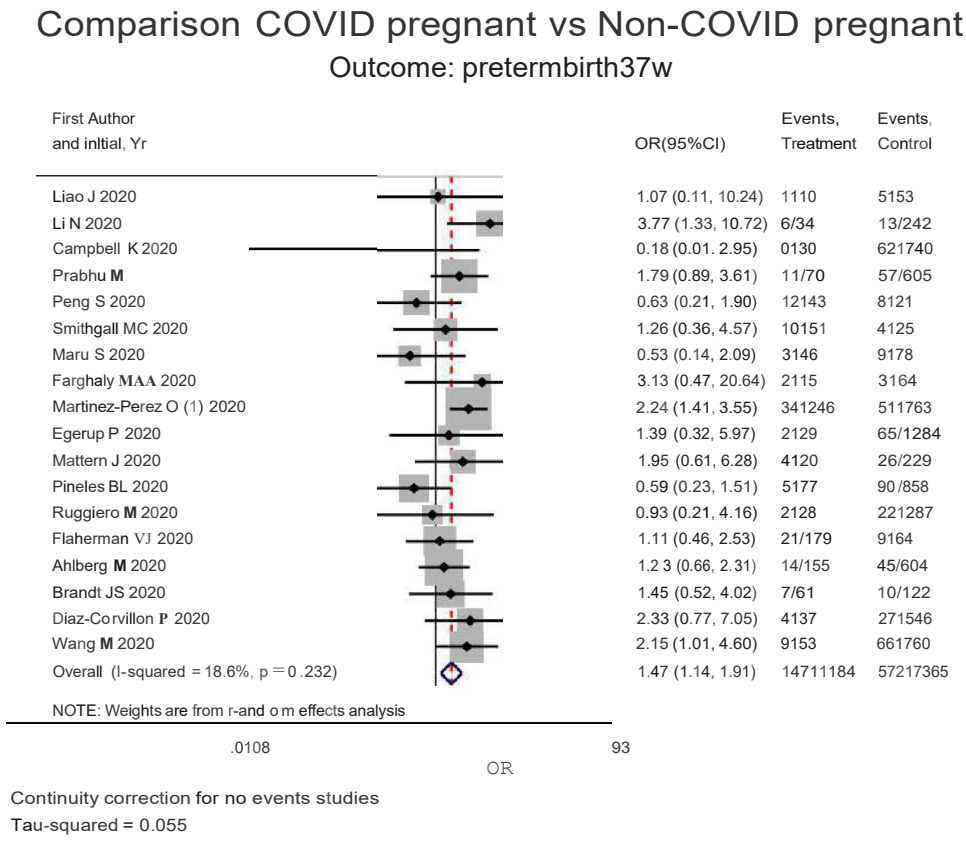
Admission to Intensive Care Unit



Continuity correction for no events studies

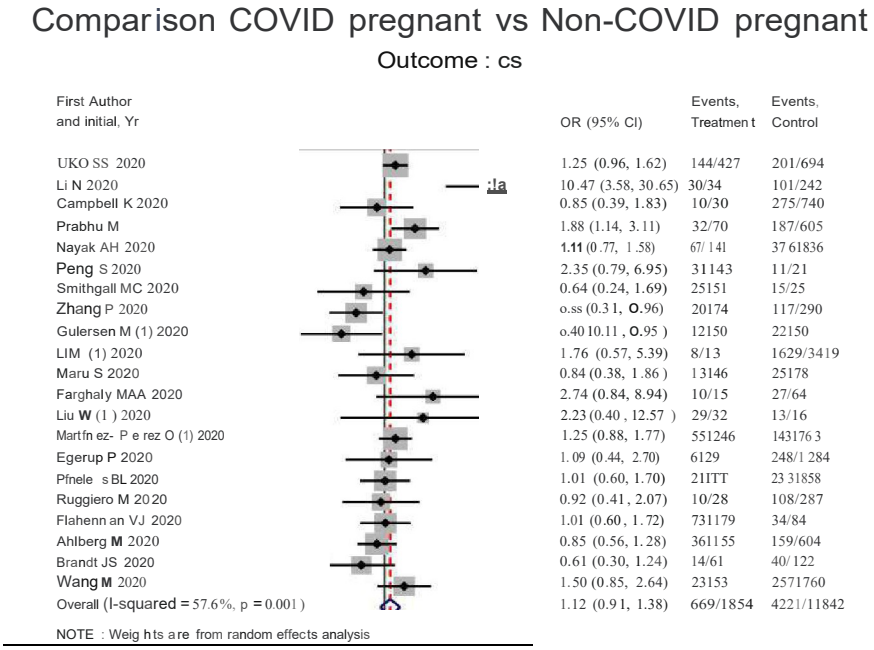
Tau-squared= 0.000

Preterm birth <37 weeks



Continuity correction for no events studies
Tau-squared = 0.055

Caesarean section



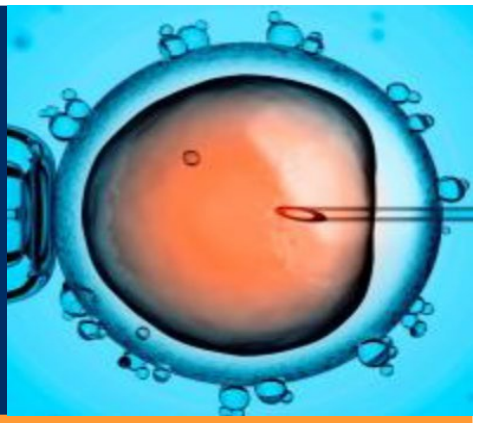
Continuity correction for no events studies
Tau-squared= 0.115

11.3 Cochrane case story

Cochrane
Gynaecology and Fertility

COVID-19
CASE STORY

Collaborating to produce the 'PregCov-19' living systematic review



Pregnant women and their children are an at risk population group for COVID-19. Cochrane Gynaecology and Fertility collaborates with the World Health Organisation (WHO) Collaborating Centre for Global Women's Health at the University of Birmingham to conduct and continuously update a living systematic review on how COVID-19 affects pregnant women and their children.

What we did

Aim

COVID-19 is especially dangerous for at risk populations, such as pregnant women. It is critical to determine how COVID-19 affects pregnant women and their babies. A regular systematic review methodology is not sufficient to synthesise the overwhelming amount of evidence produced daily worldwide. We needed to carry out a living systematic review, meaning the review would be continually updated, incorporating new studies as they become available.

Activities

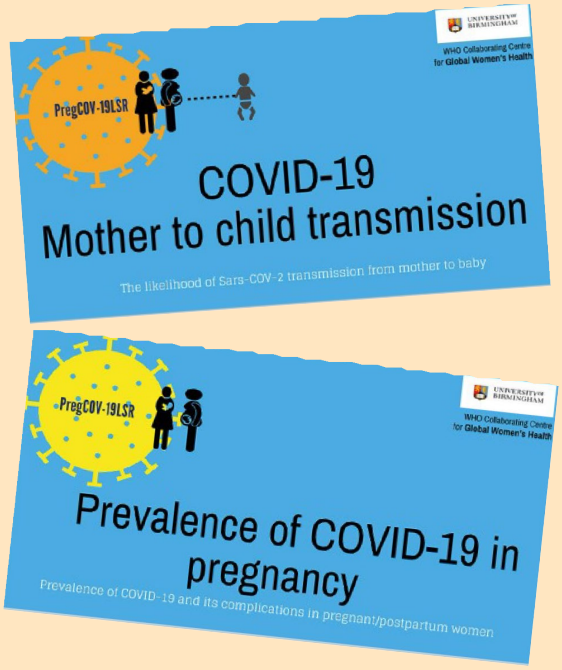
- The 'PregCov-19' living systematic review project commenced at the beginning of April 2020, just as Europe went into full lockdown. Our latest results from the living systematic review were published in September 2020.
- While maintaining the living systematic review, we also sought new members to add to the team and trained them. Training new members involves shadowing other team members for 1-2 weeks before independently carrying out tasks related to the living systematic review.
- We created a webpage on the University of Birmingham website (>14,000 views to date) to highlight the project and make it easily accessible to pregnant women, researchers, and clinicians worldwide. We are currently updating the results on the website every 2 months.

Collaboration

The PregCov-19 living systematic review working group is led by the University of Birmingham and includes the World Health Organization (WHO), Cochrane Gynaecology and Fertility (Netherlands), Cochrane Madrid, CIBER Epidemiology and Public Health, the US Centre for Disease Control, the European Centre for Disease Control, Elizabeth Glaser Pediatric AIDS Foundation, and the EPPI-Centre.

What we achieved

- Up to now, we have included 77 studies (13,118 pregnant women with COVID-19; 83,486 non-pregnant women with COVID-19) in the living systematic review.
- The first publication of the living systematic review *(Clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy: living systematic review and meta-analysis)* was fast tracked by the British Medical Journal. The review took only 5 months from initiation to publication. It currently has an Altmetric score of 943.
- One of our significant findings from the published living systematic review is that pregnant and recently pregnant women may be at increased risk of admission to an intensive care unit. This finding was picked up by various news outlets worldwide such as CNN, the Guardian, and Bloomberg.



See more here:
https://www.birmingham.ac.uk/research/who-collaborating-centre/pregcov/index.aspx
https://doi.org/10.1136/bmj.m3320
https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD202002/full

What we learnt

1. **Creating a solid framework** for a large-scale living systematic review that can answer multiple research questions was crucial. A key lesson is it that we can use the framework and infrastructure for the current living systematic review to respond to other public health issues and future pandemics.
2. We learnt that the way we **organised our team,** i.e. allocating specific tasks to individual team members every week, enabled us to collectively work on the living systematic review without delays. We organised all documents according to the different stages of the review, and for each review question, in one place. The way we organised ourselves is useful for future projects that we work on.

Learn more

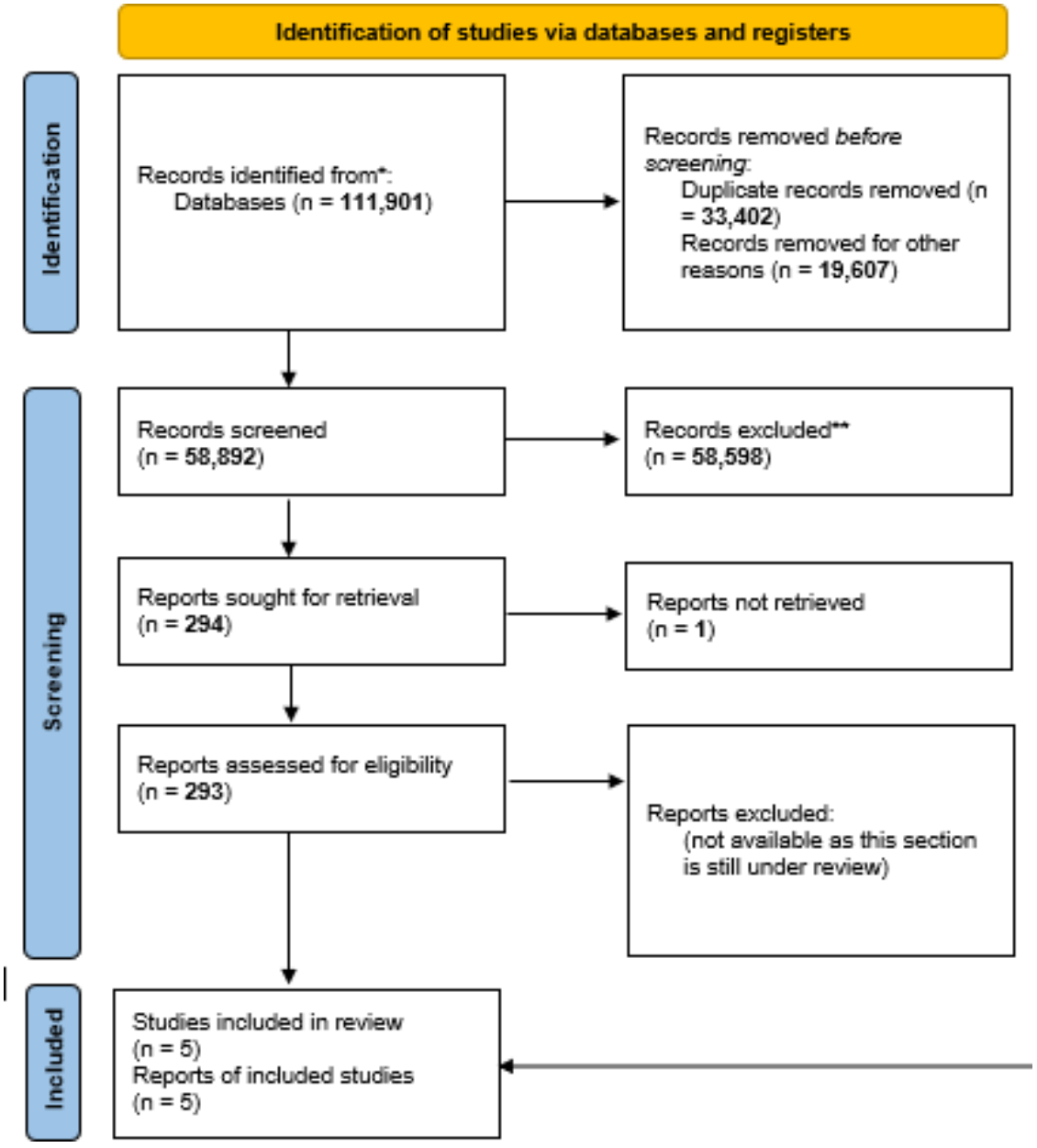
For more information contact:

Javier Zamora (U.zamora.l@bham.ac.uk)
WHO Collaborating Centre for Women's Health, University of Birmingham.
CIBER Epidemiology and Public Health, Madrid, Spain
Madrid Cochrane Associate Center

November 2020

11.4 Cochrane review preliminary result

1. PRISMA Flowchart



2. Preliminary list of included studies and tables of characteristics

Author	Year	Title
Agarwal et al	2015	Gender Disparities in Outcomes and Resource Utilization for Acute Pulmonary Embolism Hospitalizations in the United States
Barrios et al	2017	Sex differences in the characteristics and short-term prognosis of patients presenting with acute symptomatic pulmonary embolism
Borrero et al	2007	Gender differences in 30-day mortality for patients hospitalized with acute pulmonary
Feng et al	2020	Sex Differences in Pulmonary Embolism: Clinical Characteristics, in-Hospital Mortality, and 30-Day Readmissions
Rosovsky et al	2019	Sex differences in risk factors, clinical presentation, treatment and outcomes of patients presenting with acute pulmonary embolism

Table 1. Table of study characteristics: qualitative

Study	Study design	Type of setting	countries	Pulmonary embolism definition	Pulmonary embolism diagnostic criteria	Sex definition	Sex measurement	sex and gender terms adequate use	Terms used	Consistent use	Primary outcome
Agarwal 2015	retrospective cohort	hospital	USA	none	ICD-9-CM codes of 415.11, 415.12, 415.13, and 415.19	none	none	inadequate	sex, gender, female, male, women, men	no	All-cause hospital mortality at 30 days
Barrios 2017	retrospective cohort	hospital	Spain	none	objective testing of an intraluminal filling defect in larger vessels on computerized tomography pulmonary angiography (CTPA)	none	none	inadequate	sex, gender, male female, women, men	no	All-cause hospital mortality at 30 days
Borrero 2007	retrospective cohort	hospital	USA	none	ICD-9-CM codes 415.1, 415.11, 415.19, and 673.20–.24	none	none	inadequate	sex, gender, male female, women, men	no	All-cause hospital mortality at 30 days

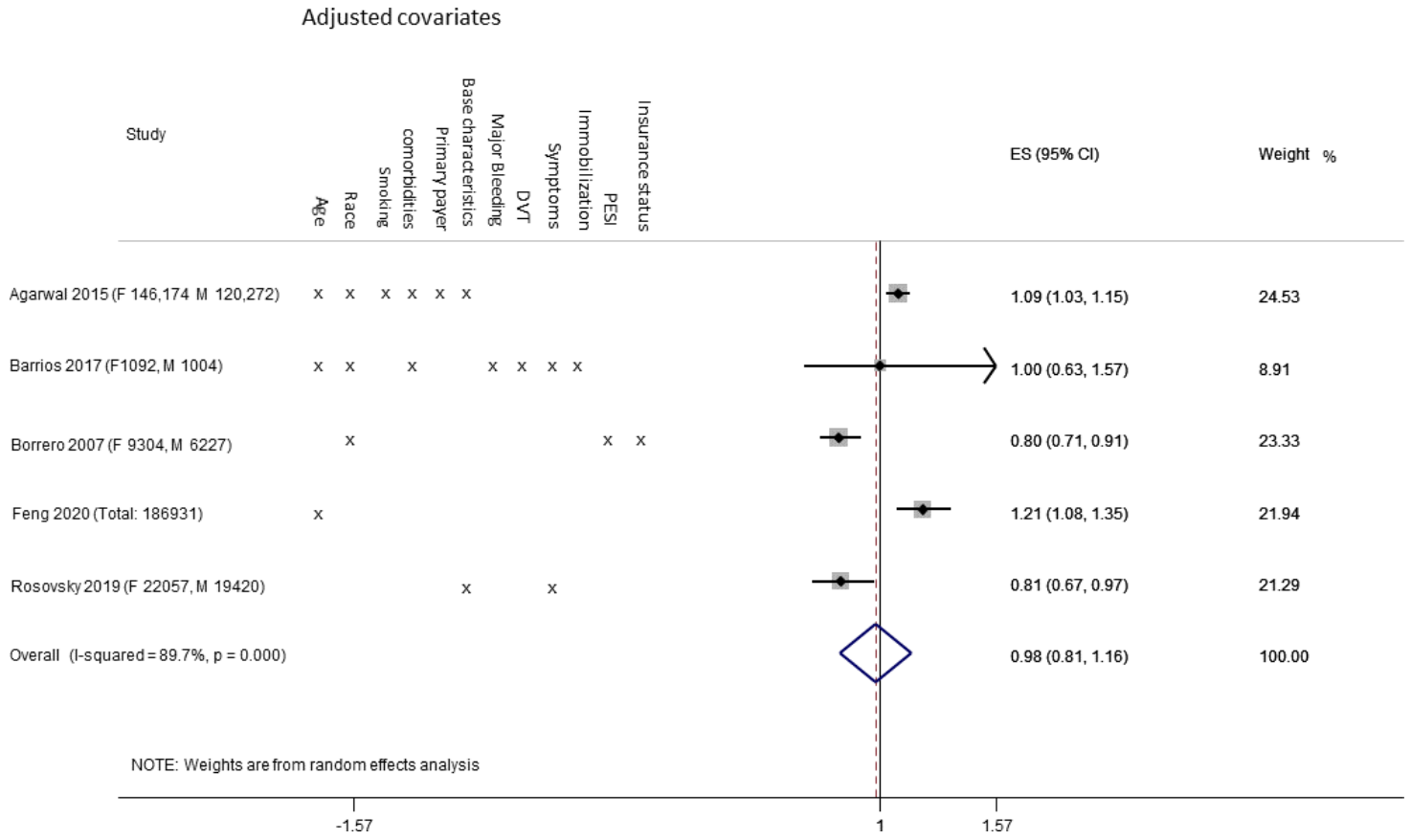
Feng 2020	retrospective cohort	hospital	USA	none	primary ICD-10 diagnosis of PE	none	none	adequate	sex, female, male	yes	All-cause hospital mortality at 30 days
Rosovsky 2019	retrospective cohort	hospital	International registry	none	Not reported	none	none	inadequate	sex, women, men	no	All-cause hospital mortality at 30 days

Table 2: Table of study characteristics: Quantitative

Study	**Sample size**	**Study design cohort vs. CC**	**Outcome measure**	**Study reported effect; variance**	**Unadj. reported effect**	**Unadj_Reported direction of association**	**Adj.reported effect**	**Adj._Reported direction of association**	**Adj. method**	**Covariates**
Agarwal 2015	260,446 (F 146,174 M 120,272)	cohort	All-cause hospital mortality at 30 days	OR;95 % CI	not reported	not reported	1.09 (1.03-1.15)	higher for F	logistic regression	Age, race, SES, Elixhauser co-morbidities, smoking, primary payer, and hospital characteristics
Barrios 2017	2096 (F 1092, M 1004)	cohort	All-cause hospital mortality at 30 days	OR;95 % CI	0.86 (0.61±1.21)	Ø	1 (0.63-1.57)	Ø	logistic regression	Age, COPD, congestive heart failure, major bleeding, DVT, Dyspnoea, Chest pain, syncope, cancer, immobilization

Barrios 2017	2096 (F1092, M 1004)	cohort	PE-related hospital mortality	OR;95 % CI	0.73 (0.43±1.22)	Ø	1.02 (0.5-2.07)	Ø	logistic regression	Age, COPD, congestive heart failure, major bleeding, DVT, Dyspnoea, Chest pain, syncope, cancer, immobilization
Borrero 2007	15531 (F 9304, M 6227)	cohort	All-cause hospital mortality at 30 days	OR;95 % CI	0.90 (0.80, 1.00)	Ø	0.8 (0.71-0.91)	lower for F	random effects logistic regression	PESI, race, insurance status, and hospital volume
Feng 2020	186931 (F na, M na)	cohort	All-cause hospital mortality at 30 days	OR;95 % CI	not reported	not reported	1.21 (1.08-1.35)	higher for F	logistic regression	Age
Rosovsky 2019	41477 (F 22057, M 19420)	cohort	All-cause hospital mortality at 30 days	OR;95 % CI	not reported	not reported	0.81 (0.67-0.97	lower for F	logistic regression	baseline characteristics and clinical presentation
Rosovsky 2019	41478 (F 22057, M 19420)	cohort	All-cause hospital mortality at 90 days	OR;95 % CI	not reported	not reported	0.83 (0.74-0.94)	lower for F	logistic regression	baseline characteristics and clinical presentation

3. Preliminary analysis: Forest plot



4. Kappa results for screening pilot

Agreements with Elena

Elia

Elena	Elia E	Elia I	Total
E	96	3	99
I	1	0	1
Total	97	3	100

Agreement	Expected Agreement	Kappa	Std. Err.	Z	Prob>Z
96.00%	96.06%	-0.0152	0.0862	-0.18	0.5701

Eduardo

Elena	Eduardo E	Eduardo I	Total
E	91	6	97
I	2	1	3
Total	93	7	100

Agreement	Expected Agreement	Kappa	Std. Err.	Z	Prob>Z
92.00%	90.42%	0.1649	0.0909	1.82	0.0348

Marcos

Elena	Marcos E	Marcos I	Total
E	97	1	98
I	1	1	2
Total	98	2	100

Agreement	Expected Agreement	Kappa	Std. Err.	Z	Prob>Z
98.00%	96.08%	0.4898	0.1000	4.90	0.0000

RaquelPR

Elena	Raquel PR E	Raquel PR I	Total
E	98	2	100
I	0	0	0
Total	98	2	100

```
             Expected
Agreement   Agreement     Kappa   Std. Err.          Z       Prob>Z
--------------------------------------------------------------------------------
  98.00%      98.00%     0.0000     0.0000          .            .
```

Alba

```
           |          Alba
     Elena |          E           I |     Total
-----------+------------------------+----------
         E |         98           1 |        99
         I |          1           0 |         1
-----------+------------------------+----------
     Total |         99           1 |       100
             Expected
Agreement   Agreement     Kappa   Std. Err.          Z       Prob>Z
--------------------------------------------------------------------------------
  98.00%      98.02%    -0.0101     0.1000       -0.10       0.5402
```

AndreaG

```
           |       Andrea G
     Elena |          E           I |     Total
-----------+------------------------+----------
         E |         92           6 |        98
         I |          0           2 |         2
-----------+------------------------+----------
     Total |         92           8 |       100
             Expected
Agreement   Agreement     Kappa   Std. Err.          Z       Prob>Z
--------------------------------------------------------------------------------
  94.00%      90.32%     0.3802     0.0785        4.84       0.0000
```

Agreements with Andrea CP

Jesus

```
           |   Jesus
 Andrea CP |          E |     Total
-----------+------------+----------
         E |        100 |       100
-----------+------------+----------
     Total |        100 |       100

             Expected
Agreement   Agreement     Kappa   Std. Err.          Z       Prob>Z
--------------------------------------------------------------------------------
 100.00%     100.00%     0.0000          .          .            .
```

Ray

```
           |          Ray
 Andrea CP |          E           I |     Total
-----------+------------------------+----------
         E |         94           6 |       100
         I |          0           0 |         0
-----------+------------------------+----------
     Total |         94           6 |       100
             Expected
Agreement   Agreement     Kappa   Std. Err.          Z       Prob>Z
```

```
-------------------------------------------------------------------------------
  94.00%      94.00%     0.0000          .          .           .
```

RaquelMG

```
           |       Raquel MG
 Andrea CP |         E          I |     Total
-----------+----------------------+----------
         E |        92          7 |        99
         I |         0          1 |         1
-----------+----------------------+----------
     Total |        92          8 |       100

            Expected
Agreement   Agreement     Kappa   Std. Err.          Z       Prob>Z
-------------------------------------------------------------------------------
  93.00%      91.16%     0.2081     0.0611        3.41       0.0003
```

Aurora

```
           |        Aurora
 Andrea CP |         E          I |     Total
-----------+----------------------+----------
         E |        89         11 |       100
         I |         0          0 |         0
-----------+----------------------+----------
     Total |        89         11 |       100

            Expected
Agreement   Agreement     Kappa   Std. Err.          Z       Prob>Z
-------------------------------------------------------------------------------
  89.00%      89.00%     0.0000     0.0000        0.00       0.5000
```

Miriam

```
           |        Miriam
 Andrea CP |         E          I |     Total
-----------+----------------------+----------
         E |        95          5 |       100
         I |         0          0 |         0
-----------+----------------------+----------
     Total |        95          5 |       100

            Expected
Agreement   Agreement     Kappa   Std. Err.          Z       Prob>Z
-------------------------------------------------------------------------------
  95.00%      95.00%     0.0000          .          .           .
```

Laura

```
           |         Laura
 Andrea CP |         E          I |     Total
-----------+----------------------+----------
         E |        83         17 |       100
         I |         0          0 |         0
-----------+----------------------+----------
     Total |        83         17 |       100

            Expected
Agreement   Agreement     Kappa   Std. Err.          Z       Prob>Z
-------------------------------------------------------------------------------
  83.00%      83.00%     0.0000     0.0000        0.00       0.5000
```

.

11.5 Filter protocol

Protocol for prognostic factor filter

Background

Prognosis research is the investigation of the relationship between future outcomes (endpoints) among individuals with a given baseline health state (start point) (Riley et al. 2019). This research is becoming significantly more important as throughout the world, people are living longer, but with more chronic health conditions and diseases. Prognosis research can be classified into four different themes or areas of research: fundamental prognostics, prognostic models, stratified medicine, and prognostic factors (Riley et al. 2019). Fundamental prognosis research signifies the course of health-related conditions in the context of the type and quality of current care (Hemingway et al. 2013). Prognostic model research refers to the development, validation, and impact of statistical models which predict an individual's risk of a future outcome (Steyerberg et al. 2013). Stratified medicine uses prognostic information to aid in adapting treatment decisions to an individual or a group of individuals with similar characteristics (Hingorani et al. 2013). Prognostic factor research studies specific factors that are associated with prognosis (Hemingway et al. 2013, Riley et al. 2013).

A prognostic factor is a variable associated with the risk of a subsequent health outcome among people with a specific health condition. Prognostic factor research is a relatively new field of research. For example, if you search "prognostic factor" and "systematic review" in title in PubMed only 70 items are retrieved. In comparison if you search for systematic reviews of interventions in PubMed the search will retrieve 998 records (as of September 2019). Systematic reviews are valuable resources as they summarize all relevant studies on specific clinical questions. However finding all information to include in systematic reviews can be challenging with an overwhelming amount of info and studies available (82).

Retrieving information from bibliographic databases such as PubMed or Embase can be very time consuming and the screening process in systematic reviews can be quite exhaustive when more than 30,000 references are retrieved (Riley et al. 2019). This is especially true

when many articles are retrieved which are irrelevant to the research question. To combat this, methodological search filters have been developed to find articles related to various types of clinical questions (Geersing et al. 2012, Beynon et al. 2013). A search filter is a predefined combination of search terms designed to retrieve information on a particular topic. The job of a search filter is to retrieve information related to a specific concept. This concept could be a topic, a study design or a methodology (Harbour et al. 2014).

Search filters attempt to increase the precision of searches and to reduce the resources that are required to screen the results. At the moment, there are various filters publicly available on medical research databases such as PubMed in the form of "clinical queries". Methodological search filters can often be referred to as hedges, optimal search strategies, optimal search filters, search filters or clinical queries (Jenkins 2004). The word methodological search filter was originally created by Wilczynski et al. who defined it as "a search term or terms (such as "random allocation" for sound studies of medical intervention) that select studies that are at the most advanced stages of testing for clinical application' (Wilczynski et al. 1993). Search filters are usually combined into a search strategy using the "AND" Boolean operator.

The various methods used for developing search filters are characterized by generation (59). First generation filters use non-objective methods as these are developed by librarians based solely on their expertise in bibliographic searches, but with no validation process. Second generation filters are developed in a similar manner, however in this development process the filters are validated using a gold standard or reference set. Third generation filters are developed based on statistical approaches and then validated against the gold standard which makes them the most objective type of filter (59).

Generic filters exist for finding prediction and prognosis studies such as the Haynes broad filter, Ingui filter and the Yale prognosis and natural history filter (Geersing et al. 2012). Published prognostic search filters have lower sensitivity and precision than other types of search filters such as those for interventions (Chatterley and Dennett 2012). Due to this half of the systematic reviews being carried out on prognosis studies are using prognosis related terms to limit the search, but only a minority of these are published prognosis search filters.

Ingui filter	(Validat$ OR Predict$.ti. OR Rule$) OR (Predict$ AND (Outcome$ OR Risk$ OR Model$)) OR ((History OR Variable$ OR Criteria OR Scor$ OR Characteristic$ OR Finding$ OR Factor$) AND (Predict$ OR Model$ OR Decision$ OR Identif$ OR Prognos$)) OR (Decision$ AND (Model$ OR Clinical$ OR Logistic Models/)) OR (Prognostic AND (History OR Variable$ OR Criteria OR Scor$ OR Characteristic$ OR Finding$ OR Factor$ OR Model$))
Haynes broad filter	(Predict*[tiab] OR Predictive value of tests[mh] OR Scor*[tiab] OR Observ*[tiab] OR Observer variation[mh])
Haynes narrow filter	(prognos*[Title/Abstract] OR (first[Title/Abstract] AND episode[Title/Abstract]) OR cohort[Title/Abstract])
Yale prognosis and natural history filter	cohort studies[mh] OR prognosis[mh] OR mortality[mh] OR morbidity[mh] OR "natural history" OR prognost*[tiab] OR course[tiab] OR predict*[tiab] OR outcome assessment[mh] OR outcome*[tiab] OR inception cohort* OR disease progression[mh] OR survival analysis[mh]

There are 4 common methodological key elements within search filter design which will be discussed in detail in the methods section: identification of a 'gold standard'; search term selection; evaluation of the search filter and validation (Jenkins 2004). Beynon et al identified 3 sources of possible bias in filter development. These sources are: 1) If systematic reviews were used to compile the reference set, these SRs must not have used a filter (for example in our case a prognosis filter); 2) The choice of the gold standard records

are subject to high bias if they are topic specific and not generalizable; 3) The validation of the filter should be carried out using a separate reference set (validation set) (60).

Objectives

To develop, evaluate and validate a search filter for prognostic factor studies. The main objective is to achieve maximum sensitivity so as not to lose any relevant studies when using the filter, while maintaining specificity. Generally, sensitivity and specificity are double edged swords, increasing sensitivity usually leads to decreased specificity, i.e. the identification of less relevant studies.

The filter details will be fine-tuned in order to optimize sensitivity /(broadness) and/or specificity (narrowness).

Methods

1. ***Identification of "gold standard":***

The first step of search filter development is to create the reference set list, which is most often referred to as the gold standard (Jenkins 2004). The reference set is a known set of studies that are relevant to, in our case prognostic factors studies. There are various methods used in creating a reference standard. These methods include hand searching, relative recall and database searching (Harbour et al. 2014, Beynon et al. 2013). Relative recall can be defined as "the proportion that any specific system retrieves of the total or pooled relevant documents retrieved by all the systems" (Sampson et al. 2006). We will use the relative recall method, which involves replicating the searches of systematic reviews and using the included studies in these reviews as the reference standard. The number of systematic reviews used in relative recall varies from 1- 27 reviews (Harbour et al. 2014). Relative recall is useful as it allows for the inclusion of a broader range of journals and publication years than would be otherwise included practically by handsearching (60, 83). This approach is also more generalizable to topics which is important for our filter as the literature is spread across a broad range of journals.

How to choose systematic reviews to use:

- We will use a validated search filter for systematic reviews and then search for "systematic reviews" and "prognostic factor" [title] in PubMed.

- We will import these reviews to EPPI-reviewer. In EPPI reviewer we will revise the results and put them into categories depending on the topics (gastro, cardio, cancer etc).

- We will select a proportion of the results to include various topics in the selection. Out of the selected SR´s we will review if they used a prognosis filter-if they did, they will be eliminated, so as not to cause bias. The selected SR´s search strategies will also be checked to see if they searched OVID medline, if they did not then they will be eliminated from our list.

- Then we should have our lists of included studies from these reviews. However, these included studies in the reference set are only the studies that were retrieved form OVID Medline, not from other sources searched in the original systematic review (60).

The reference standard will be divided into three subsets: term identification set (TIS), filter development set (FDS) and the filter validation set (FVS). The reference set will consist of studies from the same systematic review (84). For example, if we have 7 SRs included in our reference set the TIS will include all the included studies from 3 SR´s, the FDS all the studies from 2 SR´s and the FVS all the studies from 2 SRs. Using more than one reference set provides the best evidence of the performance of filters outside of the original development and test environment. It also proves the consistency of the a filters performance across different sets of records (60).

2. Search term selection:

1. Carry out frequency analysis

2. Calculate chi square values/analysis

3. Panel of SR methodologists

Search term selection will be based on the objective method used by Rietjens et al, 2019. A frequency analysis of PF articles will be carried out. The free online software systematic review accelerator (http://sr-accelerator.com/#/) will be used to carry out a word frequency analysis. We will enter the negative articles into the word analyser and then separately enter the positive articles to be able to make the comparison.

Chi square values will be calculated for terms generated from the word frequency analysis to determine the significance of the difference in relative frequencies of the terms between positive studies (the studies that are included in the review) and negative studies (studies not included in the review) in the TIS.

We will then organize a Delphi panel consisting of systematic review methodologists and information retrieval specialists to evaluate the appropriateness of including each term in the filter. The Delphi method will consist of three rounds, the first two being individual and the last round will be a panel meeting where a discussion can take place. The terms scoring between 7-9 on the Delphi will be eligible for inclusion in the filter.

3. Evaluation of filter:

An essential component of the search filter development process is the evaluation of how well the search filter performs in retrieving relevant records in a systematic review.

Each filter (sensitive and specific) will be combined with the broad search strategy for OVID Medline that was used in each included SR in the FDS. The performance of the filter will be tested against the included studies in the FDS.

During the evaluation we will test the sensitivity, specificity, precision, and number needed to read (NNR) of the filter. We will use the table below to guide us in the evaluation:

	Gold standard articles	Non-gold standard articles	
Retrieved	A (TP)	B (FP)	
Not retrieved	C (FN)	D (TN)	
	A+C		

Table 1. Table to calculate sensitivity, specificity, precision and NNR of the filter.

Sensitivity is the number of relevant references in the gold standard retrieved by the filter as a proportion of the total number of references in the gold standard (van de Glind et al. 2012, Kok et al. 2015). Sensitivity of the search filter is the extent to which the search was able to pick up, and not miss relevant articles. This reflects how many of the 'included' studies in a systematic review were potentially identifiable using a particular search filter or

database. If the search had low sensitivity, it would miss a large proportion of relevant articles. In contrast, a highly sensitive search is constructed so that it can pick up most of the relevant articles.

It will be calculated by: $\frac{A}{(A + C)} \times 100$

Specificity is the number of references that are not relevant and are not retrieved as a proportion of the total number of non-relevant references (van de Glind et al. 2012).

It will be calculated by: $\frac{D}{(B + D)} \times 100$

Precision (positive predictive value PPV) is the number of relevant records retrieved as a proportion of the total number of records retrieved by the filter (van de Glind et al. 2012).

It will be calculated by:

$$\frac{A}{(A + B)} \times 100$$

The number needed-to read (NNR) is a measure of the usability of the filter, because it indicates how many records a searcher must screen for each relevant record retrieved (van de Glind et al. 2012, Kok et al. 2015). In the context of searching, NNR refers to number of references that have to be checked to find one additional relevant article that is. Typically, in a systematic review, this would be the number of titles or abstracts retrieved from the electronic search that would have to be manually checked and considered to pick up one additional relevant article from the set of retrieved citations.

It will be calculated by:

(1/precision)

If necessary, we will fine tune the filters, eliminating different words to increase sensitivity and/or specificity.

11.6 Other publications related to thesis

11.6.1 Consideration of sex and gender in Cochrane reviews of interventions for preventing healthcare-associated infections: a methodology study

BMC Health Services Research

RESEARCH ARTICLE Open Access

Consideration of sex and gender in Cochrane reviews of interventions for preventing healthcare-associated infections: a methodology study

Jesús López-Alcalde[1,2,3,4*] Elena Stallings[3], Sheila Cabir Nunes[5], Abelardo Fernández Chávez[6], Mathilde Daheron[5], Xavier Bonfill Cosp[7,8] and Javier Zamora[3,8]

Abstract

Background: Healthcare-associated infections (HAIs) are common and increase morbidity, mortality, and healthcare costs. Their control continues to be an unresolved issue worldwide. HAIs epidemiology shows sex/gender differences. Thus the lack of consideration of sex/gender in Cochrane reviews will limit their applicability and capacity to support informed decisions. This study aims to describe the extent to which Cochrane reviews of interventions for preventing HAIs consider sex and gender.

Methods: Methodology study appraising Cochrane reviews of interventions to prevent HAIs. Search methods: *Cochrane Database of Systematic Reviews* from 1995 (launch of the journal) to 31 December 2016. Two authors independently extracted data with *EPPI-Reviewer 4* software, and independently appraised the sex/gender content of the reviews with the *Sex and Gender Appraisal Tool for Systematic Reviews (SGAT-SR)*.

Results: This study included 113 reviews assessing the effects of interventions for preventing HAIs. 100 reviews (88%) used at least one sex or gender-related term. The terminology used was heterogeneous, being "sex" the term used in more reviews (51%). No review defined neither sex nor gender. Thus we could not assess the definitions provided. Consideration of sex and gender was practically absent in the included reviews; in fact, no review met all the applicable items of the SGAT-SR, and 51 reviews (50%) fulfilled no item. No review provided a complete description of the sex and the gender of the samples of the included studies. Only ten reviews (10%) planned to perform sex- and gender-based analysis and only three (3%) could complete the analysis. The method chosen was always the subgroup analysis based on sex (one review) or gender (two reviews). Three reviews (3%) considered sex or gender-related findings in the conclusions.

Conclusion: Consideration of sex and gender in Cochrane reviews of interventions for preventing HAIs was practically absent. This lack of attention to sex and gender reduces the quality of Cochrane reviews, and their applicability for all people: women and men, boys and girls, and people of diverse gender identities. Cochrane should attempt to address the shortfalls detected.

Keywords: Systematic reviews, Data extraction, Sex, Gender, Sex/gender, Equity, Cochrane, Gender bias, Healthcare-associated infection

* Correspondence: cochrane.madrid@ufv.es
[1]Department of Paediatrics, Obstetrics & Gynaecology and Preventative Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain
[2]Faculty of Health Sciences, Universidad Francisco de Vitoria (UFV)-Madrid, Madrid, Spain
Full list of author information is available at the end of the article

Background

Health inequality and health inequity

'Health inequality' and 'health inequity' are commonly confused terms, although they have different meanings. Health inequalities refer to the differences in health status or in the distribution of health determinants between different populations (e.g., racial, ethnic, sex, gender, sexual orientation, or socioeconomic groups) [1]. On the other hand, 'health inequities,' also known as 'health disparities' [2], are avoidable and unfair differences in health across socioeconomic, demographic and geographic factors [1–6]. According to the World Health Organization (WHO) *Commission on Social Determinants of Health*, health inequity is caused by the following interacting factors: a) the socioeconomic and political context, b) the social position, b) the material circumstances, and d) the health system [7].

To reduce health inequities both within and between countries remains a priority on the agenda of international organisations, such as the WHO, and local, regional and national governments [8, 9]. The design and implementation of health care interventions and health programmes should apply an "equity lens" to ensure that benefits reach the most hard-to-reach segments of the population and to avoid intervention-generated inequalities [10, 11]. See Additional file 1 for definitions of key terms.

The relevance of sex and gender in health

Sex, gender, or sexual orientation are characteristics that may contribute to health inequalities and health inequities [5, 10, 12, 13]. The concepts of sex and gender are distinct but interrelated [14]. According to the *Canadian Institutes of Health Research,* every cell is sexed, and every person is gendered [15]. Sex, usually defined as female or male, refers to a number of biological characteristics in humans and animals [16]. Sex is linked with physical and physiological features, such as chromosomes, gene expression, hormone function and reproductive/sexual anatomy [16, 17].

On the other hand, gender refers to the social roles, behaviours, expressions and identities of girls, women, boys, men, and gender diverse people [16, 17]. Consequently, gender influences how people perceive themselves and each other, how they behave and interact, and how power and resources distribute in society [16, 17].

Sex and gender are usually conceptualised as binary factors. Thus, analyses often consider male/female for sex, as well as masculine/feminine for gender [16, 17]. However, this may not reflect the reality, as the attributes of gender are multidimensional, dynamic, and interactive [18]. The term 'sex/gender' highlights this 'entanglement' of the biological and the social [17, 19, 20].

Biological and gender-based differences result in differential health risks, disease incidence, and health service needs [10]. Consequently, sex and gender interactions can influence health and well-being in a variety of ways [16]. First, pharmacokinetics and pharmacodynamics of drugs differ between sexes, resulting in differential adverse event profiles and further affecting treatment outcomes [21–23]. Secondly, sex and gender both affect environmental and occupational risks, risk-taking behaviours, access to health care, health care-seeking behaviour, health care utilisation, and perceived experience with health care, and thus, disease prevalence and treatment outcomes [16, 24].

Consideration of sex and gender in research

The consideration of sex and gender in research is relevant for many reasons, such as for warranting scientific rigour, for reducing and enhancing the effectiveness of healthcare interventions, for promoting an informed-decision making, and for addressing inequities in health [17, 25–27]. The absence of consideration of sex and gender in research limits the external validity of research findings and their applicability for women, but also for men [16].

Various stakeholders (e.g., journal editors, research funders, policymakers) agree that sex and gender matter to health outcomes [16]. As an example, the National Institutes of Health (NIH) Revitalization Act of 1993 in the United States of America (USA) required NIH-funded clinical trials to include women and minorities as participants and to assess outcomes by sex and race or ethnicity [28]. Also, other relevant stakeholders are asking systematic reviews (SRs) to determine the evidence of differential effects across age, sex and socioeconomic status [29]; this is the case of NICE (*National Institute for Health and Care Excellence*) [30], or the *PRISMA* statement [31].

However, research design, reporting, and implementation, and general science communication often neglect sex and gender differences [14, 16, 17, 25, 26, 32–35] and policies attempting to solve this problem, such as the NIH policies cited above, have not resulted in significant increases in reporting results by sex, race, or ethnicity [36].

Methods to consider sex and gender in systematic reviews

A SR is a review that departs from a clear question and follows rigorous and explicit methods in all its stages, that is, from the identification of the studies to the analysis of the data [37].

SRs are essential tools to transfer research knowledge into evidence-informed policy and practice [29, 38, 39]. Also, SRs are crucial in the promotion of health equity as they help to determine the effects of interventions across studies conducted in a variety of settings and populations, which allows for the exploration of both prognostic factors and treatment-covariate interactions [4, 29, 40–42]. Decision-makers are interested in health equity as one of the considerations for decision-making, as they need to know the effects of interventions in the overall population and across population groups [5, 43]. Considering sex and

gender in SRs is a significant step forward in determining to whom the evidence applies, which is critical to make sound clinical and policy decisions [33].

Sex- and gender-based analysis (SGBA) is the analytical approach that incoporates the sex and gender perspective into health research, policies, and programmes, and in health planning and decision-making [16]. SGBA systemically inquiries about biological (sex-based) and socio-cultural (gender-based) differences between women and men, and boys and girls, without presuming that there are disparities [44]. In the context of SRs, SGBA is any analytical framework aiming to promote the consideration of sex and gender properly within SRs, so they have the potential to expand their findings for all people: women and men, boys, girls, and people of diverse gender identities.

There are several methodological approaches for addressing factors related to equity (sex and gender among them) in SRs, such as performing subgroup analysis or performing targeted analyses of sex/gender populations [4].

Consideration of sex and gender in Cochrane reviews

Cochrane is an international organisation that prepares SRs to support people in making well-informed decisions about health care (36). Although the extent to which current Cochrane reviews consider sex and gender is not well known, some studies suggest that there is much room for improvement. To our knowledge, only the study by Doull et al. [14] has evaluated the consideration of sex and gender in Cochrane reviews. This study concluded that SGBA was generally absent in a random sample of 38 Cochrane reviews published before April 2007 in cardiovascular health.

Moreover, Cochrane reviews seem to rarely assess whether interventions have intended or unintended effects on health equity [29]; according to another study, only 1% of a random sample of Cochrane reviews assessed differences in the effectiveness of interventions across socioeconomic or demographic factors [45]. This shortfall is also present in non-Cochrane SRs. In fact, the analysis or report of equity issues, [38, 46–49], and sex/gender in particular [17, 25, 26, 33, 35], is infrequent in SRs.

There are several obstacles to consider sex/gender in any type of SRs, such as how the included studies defined sex and gender, the methodological difficulties in measuring and analysing sex and gender, the availability of data to perform sex and gender analysis, and also the quality of this data [33]. As an example, the studies included in SRs usually do not report about the inclusion of specific populations or, if they do, they may not assess variation in effects across critical characteristics, such as sex, age, ethnicity or socioeconomic status [5]. Moreover, sex and gender are highly interrelated, and it is sometimes difficult to attribute particular male-female differences to either sex or gender alone [24]. Consequently, SRs do not clearly identify to whom the research results apply and do not present adequate data and analyses about health equity factors, including sex and gender [34, 40, 50–52]. A proper SGBA framework can help in determining external validity—to whom a particular body of evidence applies and to what degree there is sufficient evidence to generalise results [14].

The Sex/Gender Methods Group, a subgroup of the *Campbell and Cochrane Equity Methods Group* [53], was established in December 2005. One of its aims is to develop methods and tools to integrate sex and gender in the development and reporting of research synthesis [54]. Cochrane Madrid is making this a priority, and this article describes our first step to promote the application of an "equity lens" to Cochrane reviews.

Healthcare-associated infections: a public health problem

Healthcare-associated infections (HAIs) are infections acquired as a result of the delivery of health care [55]. HAIs represent a public health problem worldwide, as about 6 and 4% of the hospitalised patients in Europe and the U. S, respectively, have at least one HAI [56, 57]. Rates of HAIs seem to be even higher in low or middle-income countries [58]. HAIs increase morbidity (prolonged hospital stay and worse prognosis) [59], mortality [59, 60], and healthcare costs [59, 61, 62]. Finally, there can be limited options for treating HAIs caused by certain drug-resistant organisms, such as methicillin-resistant *Staphylococcus aureus* (MRSA) [56, 57, 63].

Role of sex and gender in infectious diseases transmission and outcomes

Traditionally, little attention has been paid to sex and gender differences in infectious diseases [24]. However, both sex and gender can affect infectious diseases incidence, duration, severity, and mortality through several pathways [24].

There are biological differences between sexes that affect infectious diseases. For example, pregnant and lactating women represent a high-risk group for many infectious diseases [24], or females have an increased risk of catheter-associated urinary tract infection (CAUTI) due to anatomy that facilitates the bacterial contamination of the catheter. On the other hand, the vulnerability to infections differs between sexes due to differences in their immune systems; in this line, pre-menopausal females seem to have a natural advantage under septic conditions [64–69], which may be explained by the role of sex steroids, that change the host immune function, alter genes and modify behaviours that influence susceptibility and resistance to infection [69]. However, a recent SR concluded that the impact of sex on sepsis outcomes remains equivocal [70]. Finally, males may present higher overall HAIs prevalence, thirty-day mortality, and one-year mortality [71].

Gender differences in behaviours, activities, exposures, and access to resources and decision-making affect transmission and outcomes for different HAIs [24]. For example, women are less likely to receive antibiotics within 3 hours of the diagnostic of sepsis, as compared to men [72].

Why this study is important

Sex and gender are necessary to understand the transmission of infectious diseases. The integration of a sex/gender perspective into SRs of interventions to control the transmission of HAIs is a new and challenging area but critical to defining successful infection control programmes [24].

Although we need more research to understand better how sex and gender interact with HAIs, there is enough knowledge available to justify the inclusion of a sex/gender perspective in research and programmes for HAIs. However, infection control strategies in the healthcare setting do not often consider sex and gender, and thus, they are generally the same for males or females [24, 73].

To consider sex and gender in Cochrane reviews of interventions to prevent HAIs is important. First, it will allow for the identification of the most effective and safest interventions for women and men. Second, it will contribute to the reduction of health inequities between men and women, and thereby promote human rights [24]. Third, the consideration of sex and gender in Cochrane reviews will help an informed-decision making for women and men. Fourth, the findings of our study will contribute to promoting the incorporation of a sex/gender perspective into Cochrane reviews of any topic.

We chose to focus our study on Cochrane reviews for several reasons. First, we foresaw that a high number of Cochrane reviews had evaluated infection control interventions [74]. Second, a wide range of Cochrane Review Groups publishes Cochrane reviews of infection control interventions, which gives an overview of the general approach of Cochrane as an organisation towards sex and gender. Third, Cochrane reviews are recognised as a reliable source of evidence worldwide and have a high impact on decision making, for example, through the consideration of Cochrane reviews in clinical guidelines [75–77], or in the medical policy documents of private health insurers [78]. As an example, the percentage of Cochrane reviews used in WHO guidelines have been steadily rising, and so far for 2016, Cochrane reviews have been included in 90% of the WHO guidelines [77]. However, if Cochrane reviews do not consider sex and gender, they will not be able to generate evidence that applies to all the people that can benefit from the research findings. Thus, Cochrane reviews may not be useful for policy-makers [29] who seek information on the distribution of effects in the population [48, 79], or may even lead to the implementation of policies and programs which inadvertently increase health inequities [80, 81].

Study aims

The general aim of this study was to describe the extent to which Cochrane reviews of interventions for preventing HAIs consider sex and gender. The specific objectives were the following.

Objective 1. To describe and assess the terminology and definitions used for sex and gender.
Objective 2. To determine the content of the reviews about sex and gender.
Objective 3. To describe the SGBA of the reviews.
Objective 4. To assess whether the review conclusions considered the sex- and gender-related findings.

Methods

Study design

Methodology study, that is, a study that assesses the methods used in randomised trials, other healthcare evaluations or systematic reviews [82].

We did not register this study in PROSPERO database because it did not meet the inclusion criteria, mainly because this is not a systematic review [83]. Our research does not adhere to any reporting statement as to our knowledge there is no guidance for reporting methodology studies.

Criteria for considering reviews for this study

Types of reviews that were eligible

Cochrane reviews published from 1995 (launch of the journal) until 31st December 2016 that evaluated the effects of interventions for preventing HAIs. The review had to be defined as a 'published review' (not at protocol nor title stage), an 'active review' (not withdrawn), and as an 'intervention review', that is, a review assessing the effects (benefits, harms or both) of health care or health policy interventions [84]. Thus, we excluded the remaining types of Cochrane reviews: methodology reviews, diagnostic reviews, overviews of reviews, prognosis reviews, and qualitative reviews.

Participants: the review must have considered any healthcare consumer in risk of healthcare-associated colonisation or HAI, except for reviews focusing on neonates, pre-terms, low birth weight or immunocompromised patients, due to the epidemiological peculiarities of these participants. We also excluded reviews focusing on the prevention of infections in healthcare professionals. All healthcare settings were eligible.

Interventions: any strategy, pharmacological or not, aimed at preventing any healthcare-associated colonisation or infection. Additional file 2 details the eligible interventions. The review could consider any comparator, that

is, an inactive comparator (such as doing nothing, use of placebo, or use of a sham intervention), or an active one (such as a pharmacological intervention, or a non-pharmacological intervention, for example, an educational or organisational one).

Outcomes: the review must have planned to assess at least one of the following outcomes: (a) occurrence of HAIs; (b) occurrence of colonisations; (c) mortality due to HAI; (d) total mortality; (e) resistance to antimicrobials; or (f) any surrogate measure of HAI, such as fever, positive culture, or antibiotic use. The review had to consider at least one of the previous outcomes in the "Types of outcome measures" section (as a primary or secondary outcome) or had to present results about any of these outcomes ("Results" section).

Search methods for identification of reviews

We searched the *Cochrane Database of Systematic Reviews (CDSR)* looking for all the Cochrane interventions reviews active and published from 1995 (launch of the journal) until 31st December 2016. Additional file 3 details the full search strategy. We used *EPPI-Reviewer 4* software [85] to create the database of reviews.

Selection of reviews

Two authors (JLA and SCN or ES) screened each title and abstract independently to select potentially eligible reviews. If there was any uncertainty based on this information, we obtained the full-text review for further assessment. Two authors (JLA and SCN or ES or AFC) independently assessed the eligibility of the retrieved full texts and resolved disagreements by discussion. If there was no consensus, we consulted a third author (JZ). We used *EPPI-Reviewer 4* software [85] to implement the selection process. We piloted the selection process with 100 records. We created a PRISMA flowchart [86] describing the results of the selection process.

Data extraction

We designed a data extraction template and piloted the form on ten reviews. Two of the piloted reviews were not eligible for this study, but we used them because they had been defined by the *Canadian Institutes of Health Research* [87] as exemplar Cochrane reviews regarding the consideration of sex and gender. The data extraction template is available upon request.

We extracted data with the *EPPI-Reviewer 4* software [85]. At least two authors (JLA, ES, AF or SCN) extracted the data for each item of the form. For critical items, two authors extracted data independently. For other items, one author extracted the data, and another author cross-checked the information extracted. We resolved discrepancies by consensus. In the case of no consensus, a third author intervened. We did not contact the review authors to obtain missing information or clarification.

Next, we detail the methods used to complete each specific objective.

Analysis methods

Objective 1. To describe and assess the terminology and definitions used for sex and gender

We described the terms used for sex and gender and the sections in which these terms appeared. Moreover, we assessed the appropriateness of the terms and definitions used by comparing them with the proposals of the *SAGER* (*Reporting of Sex and Gender Equity in Research)* guidelines [16]. We followed the classifications detailed in Tables 1 and 2. See also Additional file 1 for sex and gender definitions.

Objective 2. To determine the content of the reviews about sex and gender

We determined the sex and gender content of each review according to the domains proposed by the *Sex and Gender Appraisal Tool for Systematic Reviews (SGAT-SR)* [14]. This 21-item tool assesses how a Cochrane review has considered sex and gender. It appraises seven review

Table 1 Classification of the sex and gender terms

Classification	
Correct	Any of the following. 1. Male or female for sex and an adequate definition of sex (biological). 2. Men or women for gender and an adequate definition of gender (cultural and socially determined roles).
Incorrect	Any of the following. 1. Male or female for gender. 2. Men or women for sex.
Unclear	Any of the following. 1. Terms for sex or gender used without defining sex or gender. For example, if a review stated that the sex of the participants was male (56%) and female (64%), but there was no definition for sex, we judged the terminology as unclear because we could not know that the review was referring to sex or gender. 2. Terms inconsistently used in the review. For example, the review used the terms 'male' and 'men' for the same concept. 3. Abbreviations used without the full term provided.
Not applicable	No mention to sex or gender

Table 2 Classification of the sex and gender definitions

Classification	
Correct	Definitions like SAGER guidelines
Incorrect	Definitions different to SAGER guidelines
Unclear	Terms for sex or gender used without a definition for sex or gender
Not applicable	There was no mention of sex or gender in the review

sections: (a) Background, (b) Objectives, (c) Criteria for inclusion/exclusion, (d) Methods, (e) Results and analysis, (f) Discussion and conclusions, and (g) Table of included studies. Each question of the tool has four answers: (a) "Yes, review met criteria"; (b) "No, the review did not meet criteria"; (c) "Item was not applicable to review"; or (d) "Unable to determine". The tool allows adding free text comments to each response in case of need. We made minor wording changes to the SGAT-SR. Additional file 4 shows the domains of the tool and the guidance on which we based our judgements.

We tabulated the responses to the tool by simple counts, and summarised the results numerically to provide an indication of overall responses (as done by Doull et al. [14] in a methodology study that used the *SGAT-SR* to examine the consideration of sex and gender in a sample of Cochrane reviews in the area of cardiovascular health). We calculated the percentage of reviews meeting each item only when this was applicable (Number of reviews meeting item × 100/Total number of included reviews in which the item was applicable); thus, when an item was not applicable for a review, that review was considered neither in the numerator nor in the denominator for that item. This omission applied to reviews focusing on females only, that is, those reviews addressing pregnancy and delivery. Breast cancer was not the case, as it can also affect males.

Two authors (SCN, JLA, ES, or AFC) independently answered each item of the tool not masked to the review details. We resolved disagreements through discussion and by consulting a third author if there was no consensus. We tabulated the judgements of the SGAT-SR and used *Powerpoint 2016* [88] and *Review Manager 5.3* [89] to summarise our judgements graphically.

Objective 3. To describe the SGBA of the reviews

Objective 3a. To describe the reporting of the sex and the gender characteristics of the study participants We described if the review had attempted to report the sex and gender characteristics of the participants recruited for each included study. For a review to describe the study samples accurately, we agreed that it should have attempted to report at least the following information for each included study (based on Clayton et al. [90]). We will report the following key sex and gender characteristics.

· Sex measurement (ascertained by genotyping of blood sample)
· Number of female and male participants
· Gender measurement (ascertained by self-report)
· Number of women and men participants

We focused on the information that the review reported in the table of included studies. We classified the review reporting of the sex and gender characteristics of the study participants according to the following categories.

· *Correct with a complete description:* the review attempted to describe all the key sex and gender characteristics, and this information was available for all the included studies.

· *Correct with an incomplete description:* the review attempted to describe all the key sex and gender characteristics, but this information was not available for all the included studies, which was highlighted by the review authors.

· *Incorrect:* the review did not attempt to describe all the key sex and gender characteristics for all the included studies.

Objective 3b. To describe the SGBA in the reviews First, we identified the SRs that had planned or used any SGBA. Second, we described the SGBA methods planned and finally used. We considered Welch et al. [4] to describe the SGBA methods used:

· Subgroup analysis-pooled results: SRs that assessed impacts of health interventions on the outcome using subgroup analysis with pooling.
· Subgroup analysis-descriptive: SRs that described within-study differences without pooling.
· Targeted analyses of sex or gender populations
· Other methods used for SGBA

Objective 4. To assess whether the review conclusions considered the sex- and gender-related findings

We calculated the percentage of reviews considering the sex- and gender-related findings in the conclusions, in particular:

- % of reviews considering sex or gender findings in the "Implications for clinical practice" section

- % of reviews considering sex or gender findings in the "Implications for research" section
- % of reviews performing SGBA that considered sex or gender findings in the "Implications for clinical practice" section
- % of reviews performing SGBA that considered sex or gender findings in the "Implications for research" section

Results

Results of the search

The search strategy in *CDSR* generated 7156 records. First, we screened their titles and abstracts, and we excluded 6836 records because they were not eligible. Second, we subsequently retrieved 320 full texts for further examination. We excluded 207 full texts (see reasons in Fig. 1), and we finally included 113 reviews.

Description of the included reviews

This study included 113 reviews (see Additional file 5, also available in RIS format upon demand). The reviews were published between 2003 and 2016 within 23 different Cochrane Review Groups. The Cochrane Wounds Group was the review group with the most reviews included in this study (35/113 reviews [31%]), followed by the Anaesthesia, Critical, and Emergency Care Group (13/113 reviews [12%]), and the Incontinence Group (10/113 reviews [9%]). Each of the remaining Cochrane Review Groups published less than ten of the included reviews.

All the reviews evaluated the effects of interventions for preventing HAIs. The interventions most frequently evaluated were those aiming to prevent HAIs associated to surgery (50/113 [44%] reviews), followed by interventions to prevent infections associated to vascular accesses (21/113 [19%] reviews), and interventions based on patient and healthcare personnel hygiene (14/113 [12%] reviews). Other interventions evaluated were, for example, those to prevent urinary catheter-associated infection, education and training to prevent HAIs, or interventions to prevent infection associated with dental procedures.

All the reviews planned to assess at least one of our study outcomes. HAI was defined as eligible in 105/113 (93%) reviews, followed by total mortality (66/113 reviews [58%]), surrogate measures of HAI (31/113 reviews [27%]), colonisation (19/113 reviews [17%]), mortality due to HAI (15/113 reviews [13%]), and resistance to antimicrobials (14/113 reviews [12%]).

The most common study design was the randomised controlled trial (RCT), which was eligible in all the included reviews (100%). The RCT was the only study design eligible in 57 reviews (50%), but another 55 reviews (49%) admitted the inclusion of at least one type of non-randomised study (NRS) as well. For one review (1%), it was unclear if NRS were eligible.

Objective 1. To describe and assess the terminology and definitions used for sex and gender

100/113 reviews (88%) used at least one sex or gender-related term. The terms used varied, and no review made it explicit that sex and gender were different concepts. 'Sex' was the term used in the most reviews (58/113 reviews, 51%), followed by 'male' (54/113 reviews, 48%), 'woman' or 'women' (53/113 reviews, 47%), 'female' (50/113 reviews, 44%), 'gender' (42/113 reviews, 37%), 'men' or 'man' (28/113 reviews, 25%), or other terms, in particular 'boys' and 'girls' (3/113 reviews, 3%).

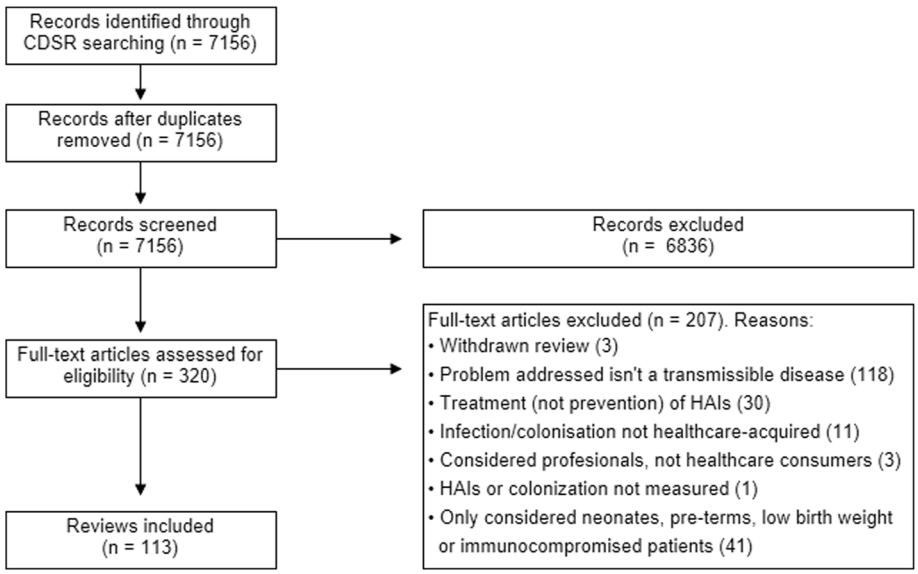


Fig. 1 Flow diagram of the selection process

13/113 reviews (12%) used no term related to sex nor gender.

The terms appeared mainly in the 'Characteristics of studies' section (84/113 reviews, 74%). On the other hand, only 10/113 (9%) reviews used these terms in the conclusions section. Table 3 details the review sections that used these terms.

The reviews defined neither sex nor gender. Thus we could not assess the definitions provided. Moreover, the absence of definitions hindered our attempt to assess the appropriateness of the terms used. For thirteen reviews it was not applicable to assess the terminology because they did not use sex or gender terms. Another 11/100 (11%) reviews used incorrect terms: 'male' or 'female' for gender (8 reviews), and 'men' or 'women' for sex (3 reviews). Also, 89/100 reviews (89%) used unclear terminology because there was no definition for sex or gender (80 reviews), because both sex and gender terms were used inconsistently apparently for the same concepts (eight reviews), or because the abbreviation "M" was used without detailing the full term (one review). Reasons to judge the terms as unclear or incorrect are shown in Fig. 2.

Objective 2. To determine the content of the reviews about sex and gender

We did not determine the sex and gender content for eleven reviews [91–101] as they included females only due to the topics addressed, that is, interventions to prevent HAIs in caesarean section, operative vaginal delivery, abortion, amniotomy, or prelabour rupture of membranes. Table 4 details the overall responses to the SGAT-SR questions, and Additional files 6 and 7 detail our responses to the tool for each review, which can be seen graphically in Fig. 3.

No review met all the applicable items of the SGAT-SR. In fact, 51/102 reviews (50%) fulfilled none of the applicable items. The remaining reviews fulfilled one (38/102 reviews [37%]), two (9/102 [9%]), three (3/102 [3%]), or four (1/102 [1%]) of the applicable items.

Review section: background

12/102 reviews (12%) used sex or gender-related terms in the background. However, only 2/102 reviews (2%) defined the relevance of sex or gender to the review question, and this was unclear for 8/102 reviews (8%). No review (0/102) discussed in its background why sex or gender differences might be expected.

Review section: objectives

No review (0/102 reviews) used the terms sex, gender, male, or female in the objectives.

Review section: criteria for inclusion/exclusion

No review used sex or gender as criteria for deciding on study eligibility (0/102 reviews) or explained why to consider sex or gender differences for study eligibility (0/102 reviews).

Review section: methods

No review (0/102 reviews) planned to examine or finally examined whether outcome measures were different for males and females. No review planned to extract or extracted data by sex (0/102 reviews). No review extracted withdrawals and dropouts by sex (0/102 reviews). No review used sex/gender as a proxy for other measures (0/102 reviews).

Table 3 Percentage of reviews using sex and gender terms in each review section

Review section	Sex	Gender	Male[a]	Female[b]	Men[c]	Women[d]	Other
Abstract	2%	–	–	1%	2%	8%	–
Background	3%	1%	4%	5%	3%	15%	–
Plain language summary		2%	1%	2%	–	8%	–
Eligibility criteria	4%	8%	1%	1%	4%	13%	–
Search methods	–	–	–	–	–	1%	–
Data collection and analysis	12%	12%	4%	4%	5%	9%	–
Results	9%	10%	12%	12%	7%	18%	1%
Discussion	4%	2%	2%	1%	4%	15%	–
Authors' conclusions	1%	–	–	1%	1%	7%	–
Characteristics of studies	31%	22%	42%	35%	13%	35%	2%
Other sections	5%	3%	–	–	–	3%	–
Not reported	49%	63%	52%	56%	75%	53%	97%

[a]'Male', 'males' or 'm'
[b]'Female', 'females' or 'f
[c]'Men' or 'man' or 'm'
[d]'Women' or woman

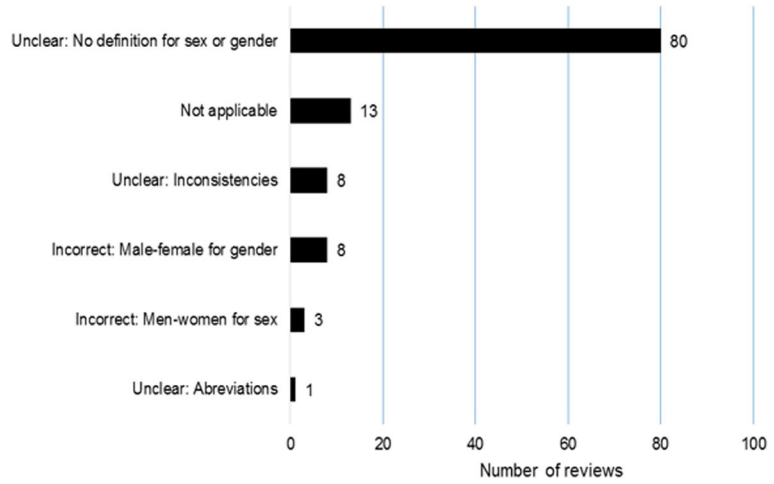


Fig. 2 Appropriateness of the sex and gender terminology

Table 4 Responses to the Sex and Gender Appraisal Tool for Systematic Reviews (SGAT-SR)

Reviews meeting the criteria (n)	Tool question	Reviews meeting the criteria (n)			
		Yes	No	Unable to determine	Not applicable
1.Background	1.1. Are the terms sex and gender used in the background?	12	90	0	11
	1.2. Are sex/gender identified as relevant or not to review question?	2	92	8	11
	1.3. Does background discuss why sex/gender differences may be expected?	0	102	0	11
2. Objectives	2.1. Are the terms sex, gender, male, or female used in objectives?	0	102	0	11
3.Criteria for inclusion-exclusion	3.1. Do the review's inclusion-exclusion criteria consider sex-gender differences?	0	102	0	11
	3.2. Was there justification or explanation for the exclusion of some groups?	0	102	0	11
4. Methods	4.1. Does the review examine whether outcome measures are different for males and females?	0	102	0	11
	4.2. Did the review extract data by sex?	0	102	0	11
	4.3. Did the review extract data on sex of withdrawals and dropouts?	0	102	0	11
	4.4. In cases where sex/gender is used as a proxy, is there an explanation?	0	0	0	113
	4.5. Were any subgroup analyses completed?	41	61	0	11
	4.6. Were subgroup analyses by sex completed?	3	99	0	11
5.Results and analysis	5.1. Do results distinguish between findings for males/females?	0	96	1	16
	5.2. Does the review report conclusions that are different for men and women?	0	96	1	16
	5.3. If adverse effects are reported, is information sex-disaggregated?	0	60	0	53
	5.4. Does review note that subgroup analyses by sex could not be done?	7	3	1	102
6. Discussion and conclusion	6.1. Does the review report that primary studies analysed or failed to analyse results by sex?	0	97	0	16
	6.2. Does the review address sex/gender implications for clinical practice?	1	101	0	11
	6.3. Does the review address sex/gender implications for policy and regulation?	0	102	0	11
	6.4. Does the review address sex/gender implications for research?	2	100	0	11
7. Table of included studies	7.1. Detailed information on sex/gender of the study samples?	0	97	0	16

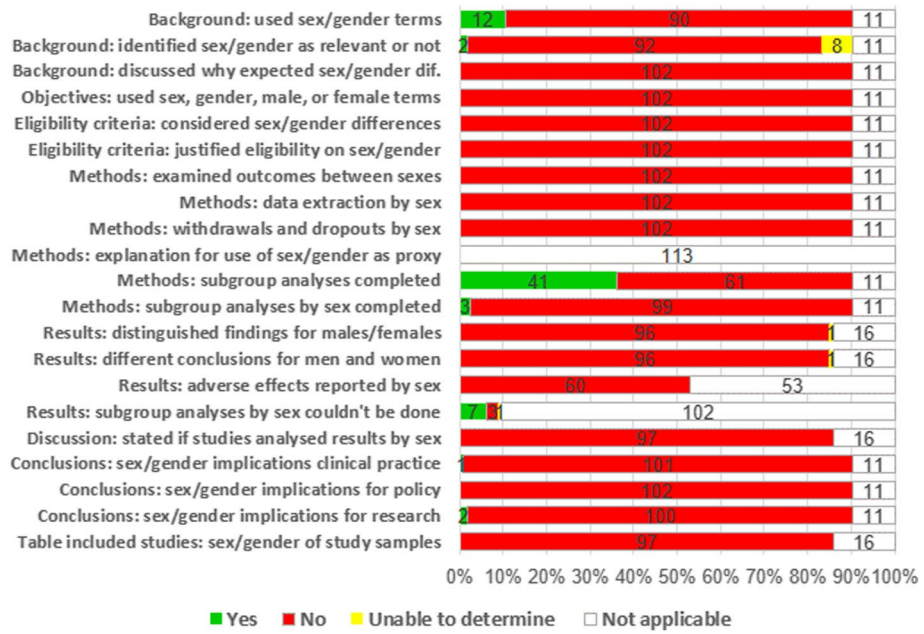


Fig. 3 Sex and gender appraisal graph. Judgements across reviews for each item of the tool

86/102 reviews (84%) planned to perform subgroup analysis, but only 41/102 reviews (40%) could complete at least one subgroup analysis. 10/102 reviews (10%) chose sex as the factor to analyse, but only three (3/102 [3%]) could complete the analysis.

Review section: results and analysis

No review in which it was applicable distinguished between findings for males and females (0/97), or reported conclusions (of effectiveness, efficacy, safety) that were different for men and women (0/97 reviews). For one review [102] we judged that these two items of the tool were unclear because the results and conclusions were reported separately only for some review outcomes. For sixteen reviews it was not applicable to assess these items because they focused on one sex only (11 reviews), or they included no study (6 reviews).

No review (0/60 reviews) reported adverse effects disaggregated by sex. For the remaining 53 reviews, it was not applicable to assess this item of the tool due to the following reasons: the reviews focused on one sex only (11 reviews), they included no study (6 reviews), they did not plan to assess adverse effects (33 reviews), or although they planned to assess safety the included studies did not report adverse effects (9 reviews).

Seven out of the ten reviews (70%) that planned but could not perform subgroup analysis by sex explained why this analysis could not be completed.

Review section: discussion and conclusions

No review in which it was applicable (0/97 reviews) reported if the included studies had analysed or failed to analyse results by sex. This aspect was not applicable in 16 reviews because they had no included studies (6 reviews) or because they focused on one sex only (11 reviews).

A total of 3/102 reviews (3%) considered sex/gender implications in their conclusions: 1/102 review (1%) addressed the implications for clinical practice, and the other 2/102 reviews (2%) considered the implications for research. However, no review addressed the implications of sex/gender for policy and regulation.

Review section: table of included studies

No review (0/97 reviews [0%]) provided detailed information on the sex and the gender of the samples of all the included studies. For sixteen reviews, this was not applicable as they did not include any study (six reviews) or they focused on one sex only (eleven reviews).

Regarding the sex of the participants recruited, 78/97 reviews (80%) provided no information at all. Another 19/97 reviews (20%) provided unclear information. Regarding the gender, 87/97 reviews (90%) provided no information, and in the other 10/97 reviews (10%) the information was unclear. We considered that the information about the sex or gender of the recruited samples was unclear for several reasons: there was no definition for sex or gender, the authors used sex or gender terms but they did not state if they referred to sex or gender, or the authors misused the terms (men or women for sex, or male or female for gender).

Objective 3. To describe the SGBA of the included reviews

Objective 3a. To describe the reporting of the sex and the gender characteristics of the study participants The review reporting of the sex and gender characteristics of the study participants was always incorrect, as no review attempted to report the sex or the gender of the participants of all the included studies. The method to ascertain the sex or gender of the recruited participants was never reported in the reviews.

Objective 3b. Description of the SGBA in the included reviews Ten reviews (10/102 [10%]) planned to perform SGBA, but only three (3/102 [3%]) could complete the analysis. The method chosen for SGBA was always the use of subgroup analysis based on sex (one review) or gender (two reviews). Two reviews performed subgroup analysis by pooling results of studies to assess the effect of sex or gender on the outcome, and the other review performed a descriptive subgroup analysis, that is, described within-study differences by gender without pooling.

Objective 4. To assess whether the review conclusions considered the sex- and gender-related findings

Only 3/102 reviews (3%) [102–104] considered the sex or gender-related findings in the conclusions. One of them mentioned the sex/gender implications for clinical practice by stating that "Siliconised catheters may be less likely to cause urethral side effects in men" [102]. The other two reviews considered the implications of sex/gender for research by stating that "sub-group analysis would give valuable data as to whether certain policies are more effective in subgroups such as females" [103], or "Future trials comparing suprapubic and intermittent urethral catheterisation for short-term use in hospitalised men should be conducted [...]" [104]. However, no review addressed the implications of sex/gender for policy and regulation. Table 5 details how the reviews considered sex or gender findings in their conclusions.

Discussion

Summary of main results

One hundred thirteen Cochrane reviews assessed the effects of interventions to prevent HAIs. Consideration of sex and gender in these reviews was practically absent. Several reasons may explain this inattention.

First, SRs may replicate limitations or gaps in primary studies regarding their research question, their data analysis, and the interpretation of their results [14]. Primary studies of hospital infection control interventions usually ignore sex, gender or having a gender identity that does not match one's biological sex. Several reasons can explain this inattention. For example, researchers may consider that making conclusions about these factors is challenging because hospital-based studies are often based on small sample sizes. Moreover, it is probably unfeasible for retrospective large primary studies, or those based on electronic records, to capture gender and sex differences because they are based on data sources that have not collected vital information, such as the assessment of biological sex by genotyping or the gender measurement by self-report. As a consequence, the included studies may not have reported the sex and gender distributions of the recruited samples. Alternatively, the authors may have just detailed that the study groups had equal numbers of males and females at baseline as a proxy for the randomisation success. Thus, sex was "controlled for" in the primary studies rather than considered to assess how the study outcomes vary across sex or gender groups [5]. Our study did not assess the data provided by the primary studies, so we cannot confirm this. Nevertheless, other studies have highlighted that RCTs do not usually provide sex or gender disaggregated data [14, 36, 105–109], and that sex/gender policies have not resulted in significant increases in reporting results by sex [36]. On the other hand, the SRs included in our study did not even plan to perform any SGBA, so we think the lack of data in the included studies cannot entirely explain their lack of attention to sex and gender.

Second, Cochrane launched the SGSR-AT [110] in 2011, that is, the Cochrane guidance to integrate sex and gender in SRs is quite recent. Thus, the guidance was not available at the time of writing the protocols of the included reviews. In fact, 31 included reviews (27%) were

Table 5 Reviews considering sex or gender findings in the conclusions

	% reviews (n/total number of reviews)
Reviews considering sex or gender findings in the "Implications for clinical practice" section	1% (1/102)
Reviews considering sex or gender findings in the "Implications for research" section	2% (2/102)
Reviews performing SGBA that considered sex or gender findings in the "Implications for clinical practice" section	33% (1/3)
Reviews performing SGBA that considered sex or gender findings in the "Implications for research" section	33% (1/3)

SGBA sex- and gender-based analysis

published in 2011 or before, and the 43 SRs published from 2012 to 2014 probably did not consider this guidance at their protocol stage.

Third, the authors of the included SRs may have thought that sex/gender was not a relevant factor to consider when evaluating the effects of interventions to prevent HAIs. Again, this is possible, but we cannot confirm that this was the case as the review authors did not make this assumption explicit in the reviews.

Fourth, the review authors did not even consider the role of sex/gender while planning the review. This situation is, in our opinion, the most plausible explanation and the most worrying one, as it denotes a knowledge gap about the potential relevance of sex and gender in these reviews.

Overall completeness and applicability of evidence

This study aimed to describe the extent to which Cochrane reviews of interventions for preventing HAIs consider sex and gender. We are quite confident that we have achieved our goal. Firstly, we have identified all the Cochrane reviews of interventions to prevent HAIs (we did not rely on a sample of Cochrane reviews), and secondly, we have demonstrated that SGBA is practically absent in these reviews.

Our study findings apply only to Cochrane reviews of interventions to prevent HAIs. Thus, we cannot infer that our study findings can be applied to Cochrane reviews of other health topics, or to non-Cochrane reviews. First, other Cochrane reviews may have incorporated the guidance of the Sex/Gender Methods Group [111]. Second, many interventions to prevent HAIs are non-pharmacological, for example, the use of gloves or hand washing. The methods used to evaluate non-pharmacological interventions may differ from those applied to evaluate drugs, and this may imply a different approach to consider sex and gender. Third, there are other health areas where the consideration of the relevance of sex and gender may be more common than in infection control research.

The findings of this study are relevant and confirm that the reviews did not consider the sex and the gender of the body of the evidence synthesised. Thus, these reviews do not provide critical information to judge the applicability of the results to the target population [112]. Also, the sex and gender characteristics should have been considered to judge the "indirectness" of the evidence with the GRADE (*Grading of Recommendations Assessment, Development and Evaluation*) system, and therefore, to rate the quality of that evidence [113]; however, we have demonstrated that this was not the case.

The tables of characteristics of the included studies are essential to describe the sex and the gender of the subjects considered within each study. In our opinion, the relevant information is the sex and gender characteristics of the sample from which the study results were obtained (and not the inclusion criteria of the study). However, specific guidance to collect and report this information in Cochrane reviews is still lacking.

Strengths and limitations of this study

Strengths

We made efforts to identify all the Cochrane reviews assessing the effects of interventions to prevent HAIs. We obtained all the Cochrane reviews published in *CDSR* until December 31st, 2016 and screened all these records. To minimise bias, at least two authors independently participated in the selection process, and, in case of disagreement, a third author was consulted. We also defined and applied explicit exclusion criteria which made the process even more rigorous.

We determined the sex/gender content of each review with the SGAT-SR. Again, to minimise bias, at least two reviewers (JLA and ES or SCN) independently participated in the assessment. In case of disagreement, a third reviewer (JZ) was consulted. Furthermore, to improve the consistency we piloted the tool with several reviews, and we prepared a user guide.

The SGAT-SR helped us to assess how the different Cochrane review sections had considered sex and gender. We chose this tool because it is recommended by the Sex/Gender Cochrane Working Group [54]. Moreover, we performed an extensive search up to April 2017 that revealed a lack of additional sex- and gender-specific appraisal tools for SRs. As the tool had already been used to assess SRs on cardiovascular diseases, we were able to better understand the tool's content by comparing our judgments with those presented in the published article [14]. Independent subject experts had also reviewed the tool to ensure consistency with common understandings of the concepts of sex and gender and to ensure compatibility with Cochrane review format and style [14].

Limitations

This study aimed to assess how the reviews had been conducted, and not how they had been reported. However, we did not write to the review authors to obtain any missing, incomplete, or unclear information. Thus, we made assumptions for information that was not clear by reading the review: we generally considered that the lack of reporting of a particular aspect meant that this was not done. However, this may not represent what the reviewers did.

During the selection and data extraction processes, we were not masked to the review team or institution. Moreover, JLA and SCN were the authors of this methodology study and three of the included reviews [114–116]. However, we prevented that this fact influenced the decisions, as the selection and extraction processes were done by at least two authors independently, and, in case of disagreement, we consulted a third author.

We did not measure the reliability of the SGAT-SR judgements (for example, by obtaining the kappa statistic), so we cannot confirm that our decisions were reliable.

We encountered some challenges with the use of the SGAT-SR. First, the tool needs to be more operative and manageable, that is, it should provide specific guidance with examples taken from other reviews. Second, it is not clear if some items of the tool refer to the planning of the review or to what the review finally did. Third, the tool should suggest when an item is not applicable; for example, for reviews with no included studies, it is not clear if the items related to the results must be answered as "not meeting the criteria" or as "not applicable". Fourth, the tool does not allow explicit assessment on the dimension of sex and the dimension of gender separately. It gathers both dimensions in the same questions. We think that each of these two domains warrants a focused appraisal. Fifth, the tool does not assess relevant sections of a Cochrane review, such as the abstract, the plain language summary, the discussion, or the summary of findings tables.

As stated in the methods section, we considered that it was not applicable to use the SGAT-ST for reviews of topics focused on one sex, such as pregnancy or delivery. We decided this because we felt that the SGAT-SR was not developed to assess both features, sex and gender, separately. However, although it did not apply to the assessment of the sex content for these reviews, it was still relevant to assess how the gender was addressed.

Finally, it is noteworthy that although sex and gender are important factors in clinical research [90], more work is needed to standardise the way sex and gender are measured and reported, and the methods to determine how these factors influence health and health care [90]). There is no consensus on how to disaggregate demographic and outcome data by sex, gender, or both, or on how to report this information. Therefore, the report of crucial information on sex, gender, or both, is incomplete in primary studies and systematic reviews. In the end, this lack of consensus moves away from a personalised medicine approach.

Agreements and disagreements with other studies or reviews

Our results are consistent with another study describing the consideration of SGBA in Cochrane reviews in the area of cardiovascular health [14], which also concluded that SGBA was practically absent in this sample of Cochrane reviews. In this line, our study supports the idea that current Cochrane reviews do not consider sex and gender, and that there is much room for improvement in this aspect. It is also noteworthy that, to our knowledge, only one study has appraised the SGBA done in Cochrane reviews [14].

This gap does not only affect Cochrane reviews, as it is also present in non-Cochrane ones. In fact, SRs do not often analyse or report equity issues in general [38, 46–49], and sex/gender in particular [17, 25, 26, 33, 35].

Conclusions

Main conclusions of this study

Consideration of sex and gender in Cochrane reviews of interventions for preventing HAIs was practically absent. This lack of attention to sex and gender reduces the quality of Cochrane reviews, and the applicability of their results for all people: women and men, boys and girls, and people of diverse gender identities.

Recommendations derived from this study

Cochrane should map the consideration of sex and gender in all Cochrane reviews and, if necessary, plan how to address the shortfalls detected efficiently.

Cochrane should continue encouraging review authors to consider sex and gender in their reviews.

The SGAT-SR helps to assess how sex and gender have been considered in a Cochrane review, although this tool has some room for improvement.

Cochrane should provide review authors with more operative guidance to consider sex and gender in Cochrane reviews.

Cochrane guidance to consider sex and gender in Cochrane reviews should be updated, validated, and required to meet the *Methodological Expectations of Cochrane Intervention Reviews (MECIR) standards.*

Primary studies should consider sex and gender differences in their research questions, data, analyses, interpretation and reporting of the study results. This approach will facilitate the consideration of sex and gender in systematic reviews and meta-analyses.

Additional files

Additional file 1: Glossary of terms. (DOCX 28 kb)
Additional file 2: Interventions defined as eligible. (DOCX 13 kb)
Additional file 3: Search strategy. (DOCX 15 kb)
Additional file 4: SGAT-SR template and guidance. (DOCX 30 kb)
Additional file 5: List of included reviews. (DOCX 21 kb)
Additional file 6: Responses to the Sex and Gender Appraisal Tool for Systematic Reviews (SGAT-SR). (DOCX 96 kb)
Additional file 7: Sex and gender appraisal summary. (PDF 643 kb)

Abbreviations

CAUTI: Catheter-associated urinary tract infection; CDSR: *Cochrane Database of Systematic Reviews*; GRADE: Grading of Recommendations Assessment, Development and Evaluation; HAI: Healthcare-associated infection; MECIR: Methodological Expectations of Cochrane Intervention Reviews; MRSA: Meticillin- (or methicillin-) resistant *Staphylococcus aureus*; NICE: National Institute for Health and Care Excellence; NIH: National Institutes of Health (NIH); NRS: Non-randomised study; RCT: Randomised controlled trial; RevMan: Review Manager; SAGER: Sex and Gender Equity in Research; SGAT-SR: Sex and gender appraisal tool for systematic reviews;

SGBA: Sex- and gender-based analysis; SR: Systematic review; UCL: University College London; USA: United States of America; WHO: World Health Organization

Acknowledgements
Rebeca Rees (EPPI Centre, Institute of Education, University of London) for her methodological support and guidance.
Purificación González Villanueva for her support in the data extraction.
EPPI reviewer technical support team for the technical support.
SEXCOMPLEX Working Group: Héctor F. Escobar-Morreale, Rosa del Campo, Agustín Albillos, Luisa María Villar, David Jiménez, Jesús López-Alcalde, Javier Zamora (Hospital Universitario Ramón y Cajal).

Funding
This study is part of the SEXCOMPLEX project ("Influence of sex and sex hormones on human chronic disorders of complex etiology"), a two-year project (2017–2019) coordinated by Hospital Ramón y Cajal (Madrid, Spain). The SEXCOMPLEX project was supported by *Instituto de Salud Carlos III* (Plan Estatal de I + D + i 2013–2016) and co-financed by the *European Development Regional Fund* "A way to achieve Europe" (ERDF) grant number PIE16/00050. These funding sources had no role in the design of this study, its execution, analyses, interpretation of the data, or decision to submit results.

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions
JLA conceived, coordinated, designed, and conducted the study; designed and executed the search strategy; selected reviews for inclusion; designed the data extraction template; piloted the data extraction; extracted data; checked the quality of data extraction; analysed and interpreted data and wrote the conclusions; wrote and edited the manuscript; approved the final report prior to submission; and is the guarantor of the study. ES, SCN and AFC selected reviews for inclusion; designed the data extraction template; piloted the data extraction; extracted data; checked the quality of data extraction; analysed and interpreted data; edited the manuscript; revised the manuscript critically for important intellectual content; and approved the final report before submission. MD extracted data; checked the quality of data extraction; interpreted data; edited the manuscript; revised the manuscript critically for important intellectual content, and approved the final report before submission. XB interpreted data; edited the manuscript; revised the manuscript critically for important intellectual content; and approved the final report before submission. JZ conceived, coordinated, and designed the study; interpreted data; edited the manuscript; revised the manuscript critically for important intellectual content, and approved the final report before submission. All authors have made substantive intellectual contributions to the study, and have read and given final approval of the manuscript.

Authors' informations
JLA is a PhD candidate in Methodology of Biomedical Research and Public (Department of Paediatrics, Obstetrics & Gynaecology and Preventative Medicine at the Universitat Autònoma de Barcelona, Spain). This study is part of his PhD.
SCN, as part of her dissertation for the *MSc Research for Public Policy and Practice* at the *Institute of Education (IOE), University College London (UCL),* analysed a subgroup of the included reviews. The dissertation, presented September 1st, 2017, was titled "Sex and gender analysis in Cochrane reviews of interventions for preventing methicillin-resistant *Staphylococcus aureus* (MRSA)", MSc, Institute of Education, University College London, London UK. Thus, the dissertation had common elements with our study, in particular, the selection process of the reviews (screening for titles/abstracts and full texts assessment), and the development of the data extraction templates. From the total number of reviews included in the study, SCN identified those reviews evaluating interventions to prevent MRSA transmission. Moreover, the data extraction templates and processes used in the dissertation represented a pilot of the whole study.
The SEXCOMPLEX Working Group leads the two-year project (2017–2019) titled "Influence of sex and sex hormones on human chronic disorders of complex etiology" (SEXCOMPLEX). The project is coordinated by *Hospital Ramón y Cajal* (Madrid, Spain) and funded by the *Institute of Health Carlos III* (*Ministry of Economy, Industry, and Competitiveness*, Spain). SEXCOMPLEX aims to assess the influence of sex/gender differences and sex hormones on the pathogenesis, clinical presentation, course and prognosis of chronic disorders of complex multifactorial aetiology. This manuscript is the deliverable of one of the work packages of the SEXCOMPLEX project.
We published a prior abstract (less than 400 words) and poster presenting the preliminary results of this study as part of Global Evidence Summit, Cape Town, South Africa (2017) [117].

Ethics approval and consent to participate
Not applicable. The Ethics Committee of *Hospital Universitario Ramón y Cajal* (Madrid, Spain) did not require ethical approval for this study, as it does not report on or involve the use of any animal or human data or tissue.

Consent for publication
Not applicable. The manuscript does not contain data from any individual person.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Department of Paediatrics, Obstetrics & Gynaecology and Preventative Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain. [2]Faculty of Health Sciences, Universidad Francisco de Vitoria (UFV)-Madrid, Madrid, Spain. [3]Clinical Biostatistics Unit, Hospital Universitario Ramón y Cajal (IRYCIS), Madrid, Spain. [4]Cochrane Associate Centre of Madrid, Madrid, Spain. [5]Independent Researcher, Madrid, Spain. [6]Preventive medicine Unit, Hospital Universitario Ramón y Cajal, Madrid, Spain. [7]Iberoamerican Cochrane Centre, IIB Sant Pau, Universitat Autònoma de Barcelona, Madrid, Spain. [8]CIBER Epidemiology and Public Health (CIBERESP), Madrid, Spain.

Received: 14 August 2018 Accepted: 7 March 2019
Published online: 15 March 2019

References

1. Health Impact Assessment (HIA) - Glossary of terms used. Available from: http://www.who.int/hia/about/glos/en/index1.html.
2. Inequity and inequality in health. Available from: http://www.globalhealtheurope.org/index.php/resources/glossary/values/179-inequity-and-inequality-in-health.html.
3. Whitehead M. The concepts and principles of equity and health. Int J Health Serv. 1992;22:429–45.
4. Welch V, Petticrew M, Ueffing E, Benkhalti Jandu M, Brand K, Dhaliwal B, Kristjansson E, Smylie J, Wells GA, Tugwell P. Does consideration and assessment of effects on health equity affect the conclusions of systematic reviews? A methodology study. PLoS One. 2012;7:e31360.
5. Petkovic J, Welch V, Jull J, Petticrew M, Kristjansson E, Rader T, Yoganathan M, McGowan J, Lyddiatt A, Grimshaw JM, et al. How health equity is reported and analyzed in randomized trials (Protocol). Cochrane Database Syst Rev. 2017:MR000046. https://doi.org/10.1002/14651858.MR000046.
6. Nelson A. Unequal treatment: confronting racial and ethnic disparities in health care. J Natl Med Assoc. 2002;94:666–8. https://www.ncbi.nlm.nih.gov/pubmed/25032386.
7. Kelly MP, Morgan A, Exworthy M, Popay J, Tugwell P, Robinson V, Simpson S, Narayan T, Myer L, Houweling T, et al. The social determinants of health: developing an evidence base for political action. Geneva: Final report to the World Health Organization Commission on the social determinants of health; 2007.
8. Petticrew M, Roberts H. How to find the studies: the literature search. In: Systematic reviews in the social sciences: a practical guide. Oxford: Blackwell; 2006. p. 79–124.
9. WHO Europe: Addressing the social determinants of health: the urban dimension and the role of local government. 2012.
10. O'Neill J, Tabish H, Welch V, Petticrew M, Pottie K, Clarke M, Evans T, Pardo Pardo J, Waters E, White H, Tugwell P. Applying an equity lens to

interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. J Clin Epidemiol. 2014;67:56–64.
11. Lorenc T, Petticrew M, Welch V, Tugwell P. What types of interventions generate inequalities? Evidence from systematic reviews. J Epidemiol Community Health. 2013;67:190–3.
12. Evans T, Brown H. Road traffic crashes: operationalizing equity in the context of health sector reform. Inj Control Saf Promot. 2003;10:11–2.
13. Oliver S. Health promotion, inequalities and young people's health: a systematic review of research. London: Evidence for Policy and Practice Information and Co-ordinating Centre; 2008.
14. Doull M, Runnels VE, Tudiver S, Boscoe M. Appraising the evidence: applying sex- and gender-based analysis (SGBA) to Cochrane systematic reviews on cardiovascular diseases. J Womens Health (Larchmt). 2010;19:997–1003.
15. CIHR Institute of Gender and Health launches NEW Training Modules: Integrating Sex and Gender into Health Research | Canadian Association of Research Ethics Boards (CAREB-ACCER). Available from: https://www.careb-accer.org/newsroom/cihr-institute-gender-and-health-launches-new-training-modules-integrating-sex-and-gender.
16. Heidari S, Babor TF, Castro PD, Tort S, Curno M. Sex and Gender Equity in Research: rationale for the SAGER guidelines and recommended use. Res Integr Peer Rev. 2016;1:2.
17. Canadian Institutes of Health Research Institute of Gender and Health: What a difference sex and gender make: a gender, sex and health research casebook - CIHR. Vancouver: British Columbia; 2012.
18. Doull M, Shea B, Puil L, Runnels V, Welch V, Tudiver S, Boscoe M. for the Sex/Gender Methods Group: Addressing Sex/Gender in Systematic Reviews: Cochrane HIV/AIDS Group Briefing Note. 2014-01.
19. Springer KW, Mager Stellman J, Jordan-Young RM. Beyond a catalogue of differences: a theoretical frame and good practice guidelines for researching sex/gender in human health. Soc Sci Med. 2012;74:1817–24.
20. Kaiser A, Haller S, Schmitz S, Nitsch C. On sex/gender related similarities and differences in fMRI language research. Brain Res Rev. 2009;61:49–59.
21. Stangier J, Stähle H, Rathgen K, Fuhr R. Pharmacokinetics and pharmacodynamics of the direct oral thrombin inhibitor dabigatran in healthy elderly subjects. Clin Pharmacokinet. 2008;47:47–59.
22. Reilly PA, Lehr T, Haertter S, Connolly SJ, Yusuf S, Eikelboom JW, Ezekowitz MD, Nehmiz G, Wang S, Wallentin L, Investigators R-L. The effect of dabigatran plasma concentrations and patient characteristics on the frequency of ischemic stroke and major bleeding in atrial fibrillation patients: the RE-LY Trial (Randomized Evaluation of Long-Term Anticoagulation Therapy). J Am Coll Cardiol. 2014;63:321–8.
23. Soldin OP, Mattison DR. Sex differences in pharmacokinetics and pharmacodynamics. Clin Pharmacokinet. 2009;48:143–57.
24. World Health Organization. Taking sex and gender into account in emerging infectious disease programmes: an analytical framework. Manila: WHO Regional Office for the Western Pacific; 2011. p. 85.
25. Ovseiko PV, Greenhalgh T, Adam P, Grant J, Hinrichs-Krapels S, Graham KE, Valentine PA, Sued O, Boukhris OF, Al Olaqi NM, et al. A global call for action to include gender in research impact assessment. Health Res Policy Syst. 2016;14:50.
26. Day S, Mason R, Lagosky S, Rochon PA. Integrating and evaluating sex and gender in health research. Health Res Policy Syst. 2016;14:75.
27. Doull M, Shea B, Puil L, Runnels V, Welch V, Tudiver S, Boscoe M, for the Sex Gender Methods Group. Addressing Sex/Gender in Systematic Reviews: Cochrane HIV/AIDS Group Briefing Note. Version 2014-01. 2014.
28. National Institutes of Health. NIH Guidelines on the inclusion of women and minorities as subjects in clinical research. Federal Register. 1994;59:14508–13. https://grants.nih.gov/grants/funding/women_min/guidelines.htm.
29. Welch V, Brand K, Kristjansson E, Smylie J, Wells G, Tugwell P. Systematic reviews need to consider applicability to disadvantaged populations: inter-rater agreement for a health equity plausibility algorithm. BMC Med Res Methodol. 2012;12:187.
30. Methods for the development of NICE public health guidance (third edition). Available from: https://www.nice.org.uk/process/pmg4/chapter/reviewing-the-scientific-evidence.
31. Welch V, Petticrew M, Tugwell P, Moher D, O'Neill J, Waters E, White H, group P-EB. PRISMA-Equity 2012 extension: reporting guidelines for systematic reviews with a focus on health equity. PLoS Med. 2012;9:e1001333.
32. Doull M, Welch V, Puil L, Runnels V, Coen SE, Shea B, O'Neill J, Borkhoff C, Tudiver S, Boscoe M. Development and evaluation of 'briefing notes' as a novel knowledge translation tool to aid the implementation of sex/gender analysis in systematic reviews: a pilot study. PLoS One. 2014;9:e110786.
33. Runnels V, Tudiver S, Doull M, Boscoe M. The challenges of including sex/gender analysis in systematic reviews: a qualitative survey. Syst Rev. 2014;3:33.
34. Johnson SM, Karvonen CA, Phelps CL, Nader S, Sanborn BM. Assessment of analysis by gender in the Cochrane reviews as related to treatment of cardiovascular disease. J Womens Health (Larchmt). 2003;12:449–57.
35. Phillips SP, Hamberg K. Doubly blind: a systematic review of gender in randomised controlled trials. Glob Health Action. 2016;9:29597.
36. Geller SE, Koch AR, Roesch P, Filut A, Hallgren E, Carnes M. The more things change, the more they stay the same: a study to evaluate compliance with inclusion and assessment of women and minorities in randomized controlled trials. Acad Med. 2017. https://doi.org/10.1097/ACM.0000000000002027.
37. Glossary | Cochrane Community. http://community.cochrane.org/glossary.
38. Lavis JN, Davies HT, Gruen RL, Walshe K, Farquhar CM. Working within and beyond the Cochrane Collaboration to make systematic reviews more useful to healthcare managers and policy makers. Healthc Policy. 2006;1:21–33. https://www.ncbi.nlm.nih.gov/pubmed/19305650.
39. World Health Organization. World Health Assembly: Resolution 5834 on the ministerial summit on health research. 2005.
40. Mackenbach JP. Tackling inequalities in health: the need for building a systematic evidence base. J Epidemiol Community Health. 2003;57:162.
41. Thomas S, Fayter D, Misso K, Ogilvie D, Petticrew M, Sowden A, Whitehead M, Worthy G. Population tobacco control interventions and their effects on social inequalities in smoking: systematic review. Tob Control. 2008;17:230–7.
42. Chopra M, Munro S, Lavis JN, Vist G, Bennett S. Effects of policy options for human resources for health: an analysis of systematic reviews. Lancet. 2008;371:668–74.
43. Vogel JP, Oxman AD, Glenton C, Rosenbaum S, Lewin S, Gulmezoglu AM, Souza JP. Policymakers' and other stakeholders' perceptions of key considerations for health system decisions and the presentation of evidence to inform those considerations: an international survey. Health Res Policy Syst. 2013;11:19.
44. Canadian Institutes of Health Research. Gender and sex-based analysis in health research: a guide for CIHR researchers and reviewers. Ottawa: CIHR; 2006.
45. Tsikata S, Robinson V, Petticrew M, Kristjansson B, Moher D, McGowan J, Shea B, Wells G, Tugwell P. Do cochrane systematic reviews contain useful information about health equity? [abstract]. XI cochrane colloquium: evidence, health care and culture 2003.
46. Maden M. Consideration of health inequalities in systematic reviews: a mapping review of guidance. Syst Rev. 2016;5:202.
47. Tugwell P, Maxwell L, Welch V, Kristjansson E, Petticrew M, Wells G, Buchbinder R, Suarez-Almazor ME, Nowlan MA, Ueffing E, et al. Is health equity considered in systematic reviews of the Cochrane Musculoskeletal Group? Arthritis Rheum. 2008;59:1603–10.
48. Petticrew M, Whitehead M, Macintyre SJ, Graham H, Egan M. Evidence for public health policy on inequalities: 1: the reality according to policymakers. J Epidemiol Community Health. 2004;58:811–6.
49. Welch V, Tugwell P, Petticrew M, de Montigny J, Ueffing E, Kristjansson B, McGowan J, Benkhalti Jandu M, Wells GA, Brand K, Smylie J: How effects on health equity are assessed in systematic reviews of interventions. Cochrane Database Syst Rev. 2010:Mr000028. https://doi.org/10.1002/14651858.MR000028.pub2.
50. Tugwell P, de Savigny D, Hawker G, Robinson V. Applying clinical epidemiological methods to health equity: the equity effectiveness loop. BMJ. 2006;332:358–61.
51. Hankivsky O, Blackwood E, Hunt R, Pigg S, Morrow M, Reid C, Patton C. Gender, diversity and evidence-based decision-making. Health Law Can. 2007;28:1-15. https://www.ncbi.nlm.nih.gov/pubmed/17982982.
52. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". Lancet. 2005;365:82–93.
53. Cochrane Methods Equity. Available from: http://methods.cochrane.org/equity/welcome.
54. Sex/Gender Methods Group | Cochrane Equity. Available from: http://methods.cochrane.org/equity/sex-and-gender-analysis.
55. Pratt RJ, Pellowe C, Loveday HP, Robinson N, Smith GW, Barrett S, et al. The epic project: developing national evidence-based guidelines for preventing healthcare associated infections. Phase I: Guidelines for preventing hospital-acquired infections. Department of Health (England). J Hospital Infect. 2001; (47 Suppl):S3-82.

56. European Centre for Disease Prevention and Control. Point prevalence survey of healthcare associated infections and antimicrobial use in European acute care hospitals. Stockholm: ECDC; 2013.
57. Magill SS, Edwards JR, Bamberg W, Beldavs ZG, Dumyati G, Kainer MA, et al. Multistate point-prevalence survey of health care-associated infections. N Engl J Med. 2014;370:1198–208.
58. World Health Organization. Report on the burden of endemic health care-associated Infection worldwide. A systematic review of the literature. Geneva: World Health Organisation; 2011.
59. World Health Organization. Antimicrobial resistance: global report on surveillance, 2014. France; 2014.
60. Klevens RM, Edwards JR, Richards CL Jr, Horan TC, Gaynes RP, Pollock DA, et al. Estimating health care-associated infections and deaths in U.S. hospitals, 2002. Public Health Reports. 2007;122:160–6.
61. Gould IM, Reilly J, Bunyan D, Walker A. Costs of healthcare-associated methicillin-resistant Staphylococcus aureus and its control. Clin Microbiol Infect. 2010;16:1721–8.
62. Marchetti A, Rossiter R. Economic burden of healthcare-associated infection in US acute care hospitals: societal perspective. J Med Econ. 2013;16:1399–404.
63. Siegel JD, Rhinehart E, Jackson M, Chiarello L, and the Healthcare Infection Control Practices Advisory Committe: Management of multidrug-resistant organisms in healthcare settings. 2006.
64. Al-Hasan MN, Eckel-Passow JE, Baddour LM. Bacteremia complicating gram-negative urinary tract infections: a population-based study. J Infect. 2010;60:278–85.
65. Angele MK, Pratschke S, Hubbard WJ, Chaudry IH. Gender differences in sepsis: Cardiovascular and immunological aspects. Virulence. 2014;5:12–9.
66. Fabbian F, De Giorgi A, Lopez-Soto PJ, Pala M, Tiseo R, Cultrera R, Gallerani M, Manfredini R. Is Female Gender as Harmful as Bacteria? Analysis of Hospital Admissions for Urinary Tract Infections in Elderly Patients. J Womens Health. 2015;24:587–92.
67. Cohen B, Choi YJ, Hyman S, Furuya EY, Neidell M, Larson E. Gender differences in risk of bloodstream and surgical site infections. J Gen Intern Med. 2013;28:1318–25.
68. Trentzsch H, Nienaber U, Behnke M, Lefering R, Piltz S. Female sex protects from organ failure and sepsis after major trauma haemorrhage. Injury. 2014; 45(Suppl 3):S20–8.
69. Klein SL. The effects of hormones on sex differences in infection: from genes to behavior. Neurosci Biobehav Rev. 2000;24:627–38.
70. Papathanassoglou E, Middleton N, Benbenishty J, Williams G, Christofi MD, Hegadoren K. Systematic review of gender- dependent outcomes in sepsis. Nurs Crit Care. 2017;22:284–92.
71. Koch AM, Nilsen RM, Eriksen HM, Cox RJ, Harthug S. Mortality related to hospital-associated infections in a tertiary hospital; repeated cross-sectional studies between 2004-2011; 2015.
72. Madsen TE, Simmons J, Choo EK, Portelli D, McGregor AJ, Napoli AM. The DISPARITY Study: do gender differences exist in Surviving Sepsis Campaign resuscitation bundle completion, completion of individual bundle elements, or sepsis mortality? J Crit Care. 2014;29:473.e477–11.
73. World Health Organization: Addressing sex and gender in epidemic-prone infectious diseases. Geneva; 2007.
74. Special Collection - Hospital acquired infections. Available from: https://www.cochranelibrary.com/collections/doi/10.1002/14651858.SC000004/full.
75. Tovey D. The role of The Cochrane Collaboration in support of the WHO Nutrition Guidelines. Adv Nutr. 2014;5:35–9.
76. Alderson PTT. The use of Cochrane Reviews in NICE clinical guidelines [editorial]. Cochrane Database Syst Rev. 2011;8. https://doi.org/10.1002/14651858.ED000032.
77. Use of Cochrane Reviews to inform WHO guidelines. Available from: http://www.cochrane.org/news/use-cochrane-reviews-inform-who-guidelines.
78. Singh A, Hussain S, Najmi AK. Role of Cochrane Reviews in informing US private payers' policies. J Evid Based Med. 2017;10:293–331.
79. Lavis J, Davies H, Oxman A, Denis JL, Golden-Biddle K, Ferlie E. Towards systematic reviews that inform health care management and policy-making. J Health Serv Res Policy. 2005;10(Suppl 1):35–48.
80. Tugwell P, Petticrew M, Robinson V, Kristjansson E, Maxwell L, Team CEFE. Cochrane and Campbell Collaborations, and health equity. Lancet. 2006;367: 1128–30.
81. Petticrew M, Tugwell P, Welch V, Ueffing E, Kristjansson E, Armstrong R, Doyle J, Waters E. Better evidence about wicked issues in tackling health inequities. J Public Health (Oxf). 2009;31:453–6.
82. The Editorial Team: Cochrane Methodology Review Group. About The Cochrane Collaboration (Cochrane Review Groups (CRGs)). The Cochrane Library. 2012:Art. No.: METHOD.
83. PROSPERO Inclusion criteria. Available from: http://www.crd.york.ac.uk/prospero/#aboutregpage.
84. About Cochrane systematic reviews. Available from: http://www.cochranelibrary.com/about/about-cochrane-systematic-reviews.html.
85. Thomas JBJ, Graziosi S. EPPI-Reviewer 4.0: software for research synthesis. Version 4.0. London: Institute of Education; 2010.
86. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med. 2009;151:W65–94.
87. Featured Cochrane reviews. Available from: http://www.cihr-irsc.gc.ca/e/193.html.
88. Microsoft: Microsoft PowerPoint. 2016.
89. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration; 2014.
90. Clayton JA, Tannenbaum C. Reporting sex, gender, or both in clinical research? JAMA. 2016;316:1863–4.
91. Gyte Gillian M, Dou L, Vazquez Juan C. Different classes of antibiotics given to women routinely for preventing infection at caesarean section. Cochrane Database Syst Rev. 2014. https://doi.org/10.1002/14651858.CD008726.pub2.
92. Haas David M, Morgan S, Contreras K. Vaginal preparation with antiseptic solution before cesarean section for preventing postoperative infections. Cochrane Database Syst Rev. 2014;9:CD007892. https://doi.org/10.1002/14651858.CD007892.pub4.
93. Hadiati Diah R, Hakimi M, Nurdiati Detty S, Ota E. Skin preparation for preventing infection following caesarean section. Cochrane Database Syst Rev. 2014;17(9):CD007462. https://doi.org/10.1002/14651858.CD007462.pub3.
94. Liabsuetrakul T, Choobun T, Peeyananjarassri K, Islam QM. Antibiotic prophylaxis for operative vaginal delivery. Cochrane Database Syst Rev. 2014;13(10):CD004455. https://doi.org/10.1002/14651858.CD004455.pub3.
95. Low N, Mueller M, Van Vliet HA, Kapp N. Perioperative antibiotics to prevent infection after first-trimester abortion. Cochrane Database Syst Rev. 2012; 14(3):CD005217. https://doi.org/10.1002/14651858.CD005217.pub2.
96. Mackeen AD, Packard Roger E, Ota E, Berghella V, Baxter Jason K. Timing of intravenous prophylactic antibiotics for preventing postpartum infectious morbidity in women undergoing cesarean delivery. Cochrane Database Syst Rev. 2014;5(12):CD009516. https://doi.org/10.1002/14651858.CD009516.pub2.
97. May W, Gülmezoglu AM, Ba-Thike K. Antibiotics for incomplete abortion. Cochrane Database Syst Rev. 2007;17(4):CD001779. https://doi.org/10.1002/14651858.CD001779.pub2.
98. Nabhan Ashraf F, Allam Nahed E, Hamed Abdel-Aziz Salama M. Routes of administration of antibiotic prophylaxis for preventing infection after caesarean section. Cochrane Database Syst Rev. 2016;17(6):CD011876. https://doi.org/10.1002/14651858.CD011876.pub2.
99. Ray A, Ray S: Antibiotics prior to amniotomy for reducing infectious morbidity in mother and infant. Cochrane Database Syst Rev. 2014;2014(10):CD010626. https://doi.org/10.1002/14651858.CD010626.pub2.
100. Smaill Fiona M, Grivell Rosalie M. Antibiotic prophylaxis versus no prophylaxis for preventing infection after cesarean section. Cochrane Database Syst Rev. 2014;10:CD007482. https://doi.org/10.1002/14651858.CD007482.pub3.
101. Wojcieszek Aleena M, Stock Owen M, Flenady V. Antibiotics for prelabour rupture of membranes at or near term. Cochrane Database Syst Rev. 2014;29(10):CD001807. https://doi.org/10.1002/14651858.CD001807.pub2.
102. Lam Thomas B, Omar Muhammad I, Fisher E, Gillies K, MacLennan S. Types of indwelling urethral catheters for short-term catheterisation in hospitalised adults. Cochrane Database Syst Rev. 2014;23(9):CD004013. https://doi.org/10.1002/14651858.CD004013.pub4.
103. Cooper Fergus P, Alexander Cameron E, Sinha S, Omar Muhammad I. Policies for replacing long-term indwelling urinary catheters in adults. Cochrane Database Syst Rev. 2016;26(7):CD011115. https://doi.org/10.1002/14651858.CD011115.pub2.
104. Kidd Emily A, Stewart F, Kassis Nadine C, Hom E, Omar Muhammad I. Urethral (indwelling or intermittent) or suprapubic routes for short-term catheterisation in hospitalised adults. Cochrane Database Syst Rev. 2015; 10(12):CD004203. https://doi.org/10.1002/14651858.CD004203.pub3.
105. Geller SE, Koch A, Pellettieri B, Carnes M. Inclusion, analysis, and reporting of sex and race/ethnicity in clinical trials: have we made progress? J Womens Health (Larchmt). 2011;20:315–20.

106. Geller SE, Adams MG, Carnes M. Adherence to federal guidelines for reporting of sex and race/ethnicity in clinical trials. J Womens Health (Larchmt). 2006;15:1123–31.
107. Gilmore-Bykovskyi A, Johnson R, Walljasper L, Block L, Werner N. Underreporting of gender and race/ethnicity differences in NIH-funded dementia caregiver support interventions. Am J Alzheimers Dis Other Demen. 2018;33(3):145-52. https://doi.org/10.1177/1533317517749465. Epub 2017 Dec 27.
108. Vidaver RM, Lafleur B, Tong C, Bradshaw R, Marts SA. Women subjects in NIH-funded clinical research literature: lack of progress in both representation and analysis by sex. J Womens Health Gend Based Med. 2000;9:495–504.
109. Grady D, Chaput L, Kristof M. Results of systematic review of research on diagnosis and treatment of coronary heart disease in women. Evid Rep Technol Assess (Summ). 2003:1–4. https://www.ncbi.nlm.nih.gov/pubmed/12827897.
110. Doull M, Runnels V, Tudiver S, Boscoe M. Sex and Gender in Systematic Reviews Planning Tool; 2011.
111. The Sex/Gender Methods Group: Guidance for reviews with the HIV/AIDS Review Group. Sex/Gender Methods Group Version 2014-01. 2014.
112. Britton A, McKee M, Black N, McPherson K, Sanderson C, Bain C. Threats to applicability of randomised trials: exclusions and selective participation. J Health Serv Res Policy. 1999;4:112–21.
113. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Falck-Ytter Y, Jaeschke R, Vist G, et al. GRADE guidelines: 8. Rating the quality of evidence--indirectness. J Clin Epidemiol. 2011;64:1303–10.
114. Barajas-Nava Leticia A, López-Alcalde J, Roqué i Figuls M, Solà I, Bonfill Cosp X. Antibiotic prophylaxis for preventing burn wound infection. Cochrane Database Syst Rev. 2013;6(6):CD008738. https://doi.org/10.1002/14651858.CD008738.pub2.
115. Bravo Zuñiga Jessica I, Loza Munárriz C, López-Alcalde J. Isolation as a strategy for controlling the transmission of hepatitis C virus (HCV) infection in haemodialysis units. Cochrane Database Syst Rev. 2016;(8):CD006420. https://doi.org/10.1002/14651858.CD006420.pub2.
116. López-Alcalde J, Mateos-Mazón M, Guevara M, Conterno LO, Solà I, Cabir Nunes S, Bonfill Cosp X. Gloves, gowns and masks for reducing the transmission of meticillin-resistant Staphylococcus aureus (MRSA) in the hospital setting. Cochrane Database Syst Rev. 2015;7:CD007087. https://doi.org/10.1002/14651858.CD007087.pub2.
117. López-Alcalde J, Stallings E, Cabir S, Zamora J. Abstract 1052: Sex/gender analysis in Cochrane reviews of healthcare-associated infections is uncommon. Cochrane Database Syst Rev. 2017;9:160. https://doi.org/10.1002/14651858.CD201702.

11.6.2 ´Give me the C´ poster