BMJ Open

Machine learning models for predicting pre-eclampsia: a systematic review protocol

Amene Ranjbar [1], Elham Taeidi,[2] Vahid Mehrnoush,[2] Nasibeh Roozbeh,[2] Fatemeh Darsareh [2]

To cite: Ranjbar A, Taeidi E, Mehrnoush V, *et al*. Machine learning models for predicting pre-eclampsia: a systematic review protocol. *BMJ Open* 2023;**13**:e074705. doi:10.1136/bmjopen-2023-074705

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2023-074705).

Received 14 April 2023
Accepted 22 August 2023

[1]Fertility and Infertility Research Center, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
[2]Mother and Child Welfare Research Center, Hormozgan University of Medical Sciences, Bandar Abbas, Iran

Correspondence to
Dr Fatemeh Darsareh; famadarsareh@yahoo.com

ABSTRACT

Introduction Pre-eclampsia is one of the most serious clinical problems of pregnancy that contribute significantly to maternal mortality worldwide. This systematic review aims to identify and summarise the predictive factors of pre-eclampsia using machine learning models and evaluate the diagnostic accuracy of machine learning models in predicting pre-eclampsia.

Methods and analysis This review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. This search strategy includes the search for published studies from inception to January 2023. Databases include the Cochrane Central Register, PubMed, EMBASE, ProQuest, Scopus and Google Scholar. Search terms include 'preeclampsia' AND 'artificial intelligence' OR 'machine learning' OR 'deep learning'. All studies that used machine learning-based analysis for predicting pre-eclampsia in pregnant women will be considered. Non-English articles and those that are unrelated to the topic will be excluded. PROBAST (Prediction model Risk Of Bias ASsessment Tool) will be used to assess the risk of bias and the applicability of each included study.

Ethics and dissemination Ethical approval is not required, as our review will include published and publicly accessible data. Findings from this review will be disseminated via publication in a peer-review journal.

PROSPERO registration number This review is registered with PROSPERO (ID: CRD42023432415).

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ A systematic review will provide most of the evidence for developing the predictive model for pre-eclampsia.
⇒ This review will be thoroughly, independently double-checked at each stage and follows best practice guidelines.
⇒ Our review will adhere to the most rigorous methodological guidelines for scoping review to ensure a high-quality review of the evidence.
⇒ The exclusion of non-English language papers may limit results.

INTRODUCTION

Pre-eclampsia is a hypertensive disorder that usually manifests itself after 20 weeks of pregnancy.[1] It can potentially cause severe morbidity, chronic disability and even the death of mothers and babies. With an estimated incidence of 2%–8%, pre-eclampsia is a significant burden on pregnant women.[2] In developing countries, the prevalence of pre-eclampsia ranges from 1.8% to 16.7%.[3] Globally, about 12% of mothers die only from pre-eclampsia.[4] Because of the poorly understood causes, various risk factors and likely multiple pathogenic phenotypes of pre-eclampsia, early prediction of pre-eclampsia is difficult. Statistical learning methods are well-equipped to deal with many variables, such as clinical and laboratory data from patients, and automatically select the most informative features.[5] Artificial intelligence has been increasingly used in health and medicine in recent years. The use of artificial intelligence in obstetrics and gynaecology has piqued the scientific community's interest.[6 7]

Artificial intelligence has been propelled forward by recent advances in computer science. Conventional general programming algorithms generate outputs based on the input data and the rules provided, whereas artificial intelligence can generate rules and patterns based on the input and output data.[8] Artificial intelligence's pattern recognition and prediction performance have been demonstrated in a variety of medical fields.[9] A systematic review of existing prognostic models was deemed necessary to advance efforts to identify women at risk of pre-eclampsia as early and accurately as possible. This would allow existing models to be evaluated for their suitability for immediate use, or to identify those that perform well internally but require external validation on an independent cohort before being considered for clinical use. This approach has the potential to be more efficient than adding a new model to aid in pre-eclampsia prevention. This systematic review aims to identify pre-eclampsia

predictors using machine learning approaches that have been reported in previous studies in this field.

METHODS/DESIGN

The protocol is reported according to the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols[10] (online supplemental file 1). The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines[11] will be used to report this study. This review is registered with PROSPERO (ID: CRD42023432415). We intended to begin the study in September 2023 and complete it by the end of December 2023.

Patient and public involvement

Patients and/or the public were not involved in this research.

Objectives

To identify and summarise the predictive factors of pre-eclampsia using machine learning models and to evaluate the accuracy of machine learning models in predicting pre-eclampsia.

Review questions

1. According to machine learning analysis, what are the predictive factors for pre-eclampsia?
2. What is the accuracy of machine learning models for predicting pre-eclampsia?

Eligibility criteria

All studies that used machine learning-based analysis for predicting pre-eclampsia in pregnant women will be considered. Non-English articles and those that are unrelated to the topic were excluded. Letters to the editor and reviews were excluded as well.

Search strategy and selection criteria

This strategy will include the search for published studies from inception to January 2023. Databases include the Cochrane Central Register, PubMed, EMBASE (Via Ovid), ProQuest, Scopus and Google Scholar. Search terms include 'preeclampsia' AND 'artificial intelligence' OR 'machine learning' OR 'deep learning'. Words and phrases will be selected from a controlled vocabulary (Medical Subject Heading (MeSH) and others) and a free-text search for each database. In addition, the reference lists of the identified articles will also be searched along with hand-searching to ensure that all documents are retrieved, which are combined using Boolean 'OR' and 'AND' operators (online supplemental file 2). All the databases will be searched by an experienced researcher. Two researchers will independently screen the titles and abstracts and then the full texts of potentially eligible studies against the predefined eligibility criteria after eliminating duplicates. Consensus or an appeal to a senior researcher will be used to resolve disagreements. A PRISMA flow diagram will be used to document the study selection process.

Data collection and risk of bias assessment

Data will be extracted independently by two investigators. Disagreements will be solved by a third party. The following items will be extracted: (1) demographic information (eg, the country where the data were collected, the setting, the data source, the study design, the prediction time and the outcome definition); (2) the method of data partitioning, the algorithms used to select the features, the features used to train the model, the type of predictive model machine learning and the validation and application of the model; and (3) the prediction results (eg, accuracy, sensitivity, specificity and area under the recurrence curve).

Two researchers will independently assess the quality of all included studies and discussed discrepancies until a consensus is achieved. PROBAST,[12] which consists of 20 signalling questions divided into four domains (participants, predictors, outcome and analysis), will be used as a tool to assess the risk of bias and applicability of each included study. PROBAST assists in assessing the outcomes studied by considering how it was determined, how objective it is, whether it incorporates any predictor data, how consistently it was determined across individuals, the timing of determination, whether it was independent of predictor information knowledge and whether it matches the review question.

Data synthesis and analysis

Measures of discriminative ability, calibration and classification accuracy will be used to describe model performance. In tabular form, key findings on study design, data sources, prediction model types, sample size, participant characteristics, model objectives, methods, presentation of the final prediction model and outcome measures will be summarised.

Handling missing data

If studies have missing data, authors will be contacted to avoid inappropriate descriptions of study results and to reduce the risk of bias. We will contact the corresponding author via email at the address listed in the published manuscript. If the corresponding author has not responded within 2 weeks, we will resend the email one more time. If all contact attempts fail, the author will be marked as unable to contact and no further attempts will be made. If the author responds after the final 2 weeks but before the final analysis of this review, the information obtained will be incorporated into the analysis.

ETHICS AND DISSEMINATION

Ethical approval is not required, as our review will include published and publicly accessible data. Findings from this review will be disseminated via publication in a peer-review journal.

Twitter Fatemeh Darsareh @famadarsareh

Contributors AR and FD were in charge of protocol design and manuscript conception. VM is in charge of determining study eligibility and reviewing collected data. The full text of papers and data collection are the responsibility of ET and NR. The authors also read the manuscript, provided significant revisions and approved the final version.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Amene Ranjbar http://orcid.org/0000-0003-2390-702X
Fatemeh Darsareh http://orcid.org/0000-0003-2644-792X

REFERENCES

1 Hypertension in pregnancy. Report of the American College of Obstetricians and Gynecologists' task force on hypertension in pregnancy. *Obstet Gynecol* 2013;122:1122–31.
2 Gestational hypertension and preeclampsia: ACOG practice bulletin, number 222. *Obstet Gynecol* 2020;135:e237–60.
3 Osungbade KO, Ige OK. Public health perspectives of preeclampsia in developing countries: implication for health system strengthening. *J Pregnancy* 2011;2011:481095.
4 Maternal mortality. n.d. Available: https://www.who.int/news-room/fact-sheets/detail/maternal-mortality
5 Marić I, Tsur A, Aghaeepour N, *et al*. Early prediction of preeclampsia via machine learning. *Am J Obstet Gynecol MFM* 2020;2:100100.
6 Boujarzadeh B, Ranjbar A, Banihashemi F, *et al*. Machine learning approach to predict postpartum haemorrhage: a systematic review protocol. *BMJ Open* 2023;13:e067661.
7 Mehrnoush V, Ranjbar A, Farashah MV, *et al*. Prediction of postpartum hemorrhage using statistical traditional analysis and machine learning approach. *AJOG Glob Rep* 2023;3:100185.
8 Vahidi Farashah M, Etebarian A, Azmi R, *et al*. An analytics model for telecovas customers' basket clustering using ensemble learning approach. *J Big Data* 2021;8:1–24.
9 Akazawa M, Hashimoto K, Katsuhiko N, *et al*. Machine learning approach for the prediction of postpartum hemorrhage in vaginal birth. *Sci Rep* 2021;11:22620.
10 Shamseer L, Moher D, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;350:g7647.
11 Page MJ, McKenzie J, Bossuyt P, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *MetaArXiv* [Preprint] 2020.
12 Moons KGM, Wolff RF, Riley RD, *et al*. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med* 2019;170:W1–33.

PRISMA-P 2015 Checklist

This checklist has been adapted for use with systematic review protocol submissions to BioMed Central journals from Table 3 in Moher D et al: Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews* 2015 **4**:1

An Editorial from the Editors-in-Chief of *Systematic Reviews* details why this checklist was adapted - Moher D, Stewart L & Shekelle P: Implementing PRISMA-P: recommendations for prospective authors. *Systematic Reviews* 2016 **5**:15

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
ADMINISTRATIVE INFORMATION					
Title					
Identification	1a	Identify the report as a protocol of a systematic review	☒	☐	1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	☐	☒	
Registration	2	If registered, provide the name of the registry (e.g., PROSPERO) and registration number in the Abstract	☒	☐	37
Authors					
Contact	3a	Provide name, institutional affiliation, and e-mail address of all protocol authors; provide physical mailing address of corresponding author	☒	☐	7-19
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	☒	☐	135-139
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	☐	☒	
Support					
Sources	5a	Indicate sources of financial or other support for the review	☒	☐	189
Sponsor	5b	Provide name for the review funder and/or sponsor	☒	☐	189
Role of sponsor/funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	☒	☐	189
INTRODUCTION					
Rationale	6	Describe the rationale for the review in the context of what is already known	☒	☐	50-70
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to	☒	☐	79-81

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
		participants, interventions, comparators, and outcomes (PICO)			
METHODS					
Eligibility criteria	8	Specify the study characteristics (e.g., PICO, study design, setting, time frame) and report characteristics (e.g., years considered, language, publication status) to be used as criteria for eligibility for the review	☒	☐	85-88
Information sources	9	Describe all intended information sources (e.g., electronic databases, contact with study authors, trial registers, or other grey literature sources) with planned dates of coverage	☒	☐	90-95
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	☒	☐	96-99
STUDY RECORDS					
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	☒	☐	100-102
Selection process	11b	State the process that will be used for selecting studies (e.g., two independent reviewers) through each phase of the review (i.e., screening, eligibility, and inclusion in meta-analysis)	☒	☐	103-109
Data collection process	11c	Describe planned method of extracting data from reports (e.g., piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	☒	☐	103-109
Data items	12	List and define all variables for which data will be sought (e.g., PICO items, funding sources), any pre-planned data assumptions and simplifications	☒	☐	103-109
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	☒	☐	104-109
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	☒	☐	110-115
DATA					
Synthesis	15a	Describe criteria under which study data will be quantitatively synthesized	☒	☐	119-122
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data, and methods of combining data from studies, including any planned exploration of consistency (e.g., I^2, Kendall's tau)	☒	☐	120
	15c	Describe any proposed additional analyses (e.g., sensitivity or subgroup analyses, meta-regression)	☐	☒	NA

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	☐	☒	NA
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (e.g., publication bias across studies, selective reporting within studies)	☐	☒	NA
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (e.g., GRADE)	☐	☒	NA

Search strategy

PubMed

(("preeclampsia"[Title/Abstract] AND ("artificial intelligence "[Title/Abstract] OR "machine learning"[Title/Abstract] OR "deep learning"[Title/Abstract]))

Embase

(preeclampsia:ti,ab,kw) AND (artificial intelligence:ti,ab,kw OR machine learning:ti,ab,kw OR deep learning:ti,ab,kw OR)

Scopus

(TITLE-ABS-KEY((preeclampsia)) AND (TITLE-ABS-KEY((“artificial intelligence") OR ("machine learning") OR ("deep learning"))

Web of Science

(TS=((preeclampsia)) AND (TS=((“artificial intelligence ") OR ("machine learning") OR ("deep learning"))

ProQuest

(ab((preeclampsia)) AND (ab((“artificial intelligence") OR ("machine learning") OR ("deep learning"))

Google Scholar

("preeclampsia") AND ("artificial intelligence "OR "machine learning" OR "deep learning")

Cochrane library

((preeclampsia)) AND ((“artificial intelligence") OR ("machine learning") OR ("deep learning")) in Title Abstract Keyword