

RESEARCH ARTICLE

Risk prediction models for maternal mortality: A systematic review and meta-analysis

Kazuyoshi Aoyama^{1,2*}, Rohan D'Souza^{2,3}, Ruxandra Pinto⁴, Joel G. Ray^{5,6}, Andrea Hill⁴, Damon C. Scales^{2,4}, Stephen E. Lapinsky⁷, Gareth R. Seaward^{2,3}, Michelle Hladunewich⁸, Prakesh S. Shah^{2,9}, Robert A. Fowler^{2,4}

1 Department of Anesthesia and Pain Medicine, The Hospital for Sick Children, Toronto, ON, Canada, **2** Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada, **3** Department of Obstetrics and Gynaecology, Division of Maternal-Fetal Medicine, Mount Sinai Hospital, Toronto, ON, Canada, **4** Department of Critical Care Medicine, Sunnybrook Health Science Center, Toronto, ON, Canada, **5** Keenan Research Centre of the Li Ka Shing Knowledge Institute of St. Michael's Hospital, Toronto, ON, Canada, **6** Department of Obstetrics and Gynecology, St. Michael's Hospital, Toronto, ON, Canada, **7** Department of Critical Care Medicine, Mount Sinai Hospital and University Health Network, Toronto, ON, Canada, **8** Kidney Care Centre, Sunnybrook Health Science Center, Toronto, ON, Canada, **9** Departments of Paediatrics, Mount Sinai Hospital, Toronto, ON, Canada

* kazu.aoyama@utoronto.ca



OPEN ACCESS

Citation: Aoyama K, D'Souza R, Pinto R, Ray JG, Hill A, Scales DC, et al. (2018) Risk prediction models for maternal mortality: A systematic review and meta-analysis. *PLoS ONE* 13(12): e0208563. <https://doi.org/10.1371/journal.pone.0208563>

Editor: Olalekan Uthman, The University of Warwick, UNITED KINGDOM

Received: October 13, 2018

Accepted: November 19, 2018

Published: December 4, 2018

Copyright: © 2018 Aoyama et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: KA, JGR, DCS, SEL, GRS, MH, PSS and RAF have received a Canadian Institutes of Health Research Transitional Open Operating Grant for this project (#342397) (<http://www.cihr-irsc.gc.ca>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Purpose

Pregnancy-related critical illness leads to death for 3–14% of affected women. Although identifying patients at risk could facilitate preventive strategies, guide therapy, and help in clinical research, no prior systematic review of this literature exploring the validity of risk prediction models for maternal mortality exists. Therefore, we have systematically reviewed and meta-analyzed risk prediction models for maternal mortality.

Methods

Search strategy: MEDLINE, EMBASE and Scopus, from inception to May 2017.

Selection criteria: Trials or observational studies evaluating risk prediction models for maternal mortality.

Data collection and analysis: Two reviewers independently assessed studies for eligibility and methodological quality, and extracted data on prediction performance.

Results

Thirty-eight studies that evaluated 12 different mortality prediction models were included. Mortality varied across the studies, with an average rate 10.4%, ranging from 0 to 41.7%.

The Collaborative Integrated Pregnancy High-dependency Estimate of Risk (CIPHER) model and the Maternal Severity Index had the best performance, were developed and validated from studies of obstetric population with a low risk of bias. The CIPHER applies to critically ill obstetric patients (discrimination: area under the receiver operating characteristic curve (AUC) 0.823 (0.811–0.835), calibration: graphic plot [intercept—0.09, slope 0.92]).

The Maternal Severity Index applies to hospitalized obstetric patients (discrimination: AUC 0.826 [0.802–0.851], calibration: standardized mortality ratio 1.02 [0.86–1.20]).

Abbreviations: AUC, area under the receiver operating characteristic curve; SAPS, simplified acute physiology score; APACHE, acute physiology and chronic health evaluation score; MPM, the mortality prediction model; SOFA, sequential organ failure assessment; ICU, intensive care unit; PRISMA, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocols; CI, confidence interval; SE, standard error; SMR, Standardized Mortality Ratio; TRIPOD, Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis; PROBAST, Prediction model study Risk Of Bias Assessment Tool; WHO, World Health Organization.

Conclusions

Despite the high heterogeneity of the study populations and the limited number of studies validating the finally eligible prediction models, the CIPHER and the Maternal Severity Index are recommended for use among critically ill and hospitalized pregnant and postpartum women for risk adjustment in clinical research and quality improvement studies. Neither index has sufficient discrimination to be applicable for clinical decision making at the individual patient level.

Introduction

Pregnancy- and peri-partum-related critical illness occurs at a frequency of 0.7 to 7.6 cases per 1,000 live births in developed countries [1,2], and leads to death for 3–14% of affected women [1,3,4]. Determination of the risk of a woman becoming critically ill or dying is helpful to better anticipate and possibly prevent serious illness and to guide therapeutic decision-making. In clinical research, groups of characteristics that together predict an outcome can be used to help account for differences between patients, when you wish to estimate the influence of some new factor on a clinical outcome such as death [5].

A number of risk prediction models have been developed for outpatients, hospitalized patients, and those who are critically ill. The simplified acute physiology score (SAPS) [6], acute physiology and chronic health evaluation score (APACHE I, II, III, IV) [7], the mortality prediction model (MPM) [8,9], and sequential organ failure assessment (SOFA) scores [10] were originally designed to predict mortality in a general adult intensive care unit (ICU) populations.

These and other prediction models have been applied to pregnant and postpartum women, either in the ICU or in a general ward; however, their performance characteristics have generally not been determined among pregnant and postpartum women [11]. Within obstetrics, a limited number of risk prediction models have been developed for specific obstetric conditions (e.g. preeclampsia, postpartum hemorrhage) [12,13]. Optimal prediction models for unselected, broad cohorts of pregnant and postpartum patients have not been well summarized and previous reviews have concluded that existing *comorbidity* indices have modest predictive ability for obstetric patients [14–16]. While risk prediction models developed from non-pregnant and postpartum populations have been adopted in clinical research for obstetric patients [11], they may have important limitations due to a combination of unique conditions leading to pregnancy-related critical illness and/or death—the typically young age of pregnant patients, and physiological changes specific to pregnancy that may be different from other patient populations. Previously published studies show that non-specific risk prediction models tend to overestimate mortality when applied to pregnant and postpartum women [14,17]. There is no prior systematic review of this literature exploring the validity of risk prediction models for mortality among pregnant and postpartum critically ill women admitted to acute care hospitals.

Therefore, we aimed to systematically review and meta-analyze risk prediction models for maternal mortality in hospitalized and critically ill pregnant and postpartum women.

Methods

This meta-analysis was conducted on the basis of a guideline for the systematic review of prediction models [18]. The results were reported following the Preferred Reporting Items for

Systematic Reviews and Meta-Analyses protocols (PRISMA-P) 2015 statement [19]. This systematic review was registered at PROSPERO (CRD42017070424).

Criteria for considering studies for this review

Type of studies. We included clinical trials, cohort and case-control studies. Case-reports, case-series, reviews and editorials were excluded.

Participants. Participants were hospitalized pregnant and postpartum women (to 6 weeks after delivery) in acute care hospitals. Patients in outpatient clinics or emergency rooms were excluded.

Index models. Prediction models derived from general hospitalized pregnant and postpartum populations or from critically ill patient populations (e.g. SAPS, APACHE, MPM and SOFA) [6–10]. Models focusing on only specific diagnoses (e.g. preeclampsia, postpartum hemorrhage) were excluded due to their limited generalizability to all obstetric patients [12,13]. We excluded indices that focused only upon pre-existing comorbidity indices as we have previously investigated their predictive performance in obstetric populations [16].

Primary outcome of index models. Maternal mortality (death during pregnancy and up to 42 days after delivery or termination of pregnancy).

Search methods for identification of studies

Electronic search. MEDLINE, EMBASE (OvidSP) and Scopus were searched systematically for eligible studies, from their inception to May 2017 (S1 File), containing three sets of terms reflecting the research question: the models (index risk prediction models), the target condition (maternal critical illness or death), and the patient population (pregnant and postpartum women). Known models were included as a key word in a broader search strategy (S1 File) [6–10]. No language restriction was made.

Searching other resources. In addition to identified articles retrieved from electronic databases, a citation search was performed in Web of Science to identify other articles that cited the identified articles above. A manual search was conducted from the reference lists of the Web of Science identified articles. Lastly, experts (SL, JGR) in the field were contacted to identify unpublished studies or studies that may not have been captured in MEDLINE, EMBASE and Scopus.

Data collection and analysis

Selection of studies. Inclusion criteria: 1) study reports performance of a mortality risk prediction model 2) while pregnant or within 42 days of delivery or termination of the pregnancy; and, 3) among patients admitted to an acute care hospital.

Exclusion criteria: 1) study design was a case-report, case-series, reviews or editorials; 2) patient population was pregnant or postpartum women with a specific diagnosis (e.g. only pre-eclamptic women); or, 3) indices including only pre-existing comorbid conditions.

The two independent reviewers (KA, RD) scanned the titles and abstracts of every record retrieved to determine whether the article was relevant, according to the above eligibility criteria. The full text of potentially eligible articles was then retrieved. The reference lists of retrieved articles were also searched for additional citations. Two reviewers (KA, RD) independently assessed and determined the eligibility of studies. Disagreement was resolved by discussion and, when necessary, a third reviewer (RAF) assisted in adjudicating a final decision.

Data extraction and management. Reviewers used standardized, piloted data forms to independently extract data from all eligible studies. Each data element was compared between primary and secondary reviewers. Any discrepancies were resolved by discussion or

adjudication by the third reviewer. Each study was described by general information (title, journal, year, publication status and study design [prospective or retrospective]), descriptors (number of included patients, age, country, subgroups, type of risk prediction model, and stated purpose of the model), reference information (clinical follow-up, mortality rate) and descriptors relevant for assessing the fitness of the model for its intended use: 1) Discrimination—the area under the receiver operating characteristic curve (AUC) or the equivalent c-statistic with 95% confidence interval (CI) or standard error (SE); and, if the AUC or c-statistic was not reported, other operational statistics such as sensitivity and specificity or positive and negative predictive values were recorded when available; 2) Calibration—information on the predicted versus observed mortality ratio is presented as the Standardized Mortality Ratio [SMR] (i.e. observed mortality divided by predicted mortality where $SMR < 1$ reflects an overestimation of the outcome and $SMR > 1$ reflects underestimation of the outcome) and goodness-of-fit statistics (e.g. Hosmer-Lemeshow [H-L] goodness-of-fit test) [20]. The corresponding author of the original study was contacted to provide missing data.

We used a recently developed reporting system of prediction models in systematic reviews (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) [21,22] and extracted 22 data components for each study in the form of a TRIPOD checklist.

Assessment of methodological quality. There is no single standard for the assessment of quality for prediction or prognostic studies [23,24]. However, PROBAST (Prediction model study Risk Of Bias Assessment Tool), a new tool for assessing the methodological quality of risk prediction models was employed [25] [Fig 1]. The usability (an actionable recommendation) of a risk prediction model is determined in the following manner [25]. First, the risk of bias and any concerns of applicability (*whether the model fits the research question: i.e. what is the most reliable and best-validated risk adjustment and outcome prediction tool for hospitalized pregnant and postpartum women?*) of the model to the intended patient population are noted. Second, the model's predictive performance (i.e. discrimination and calibration) is considered. Although PROBAST does not specify how good discrimination and calibration should be, generally AUCs higher than or equal to 0.8 for evaluating discrimination are good and an SMR of approximately 1.0 for evaluating calibration is considered as excellent. Calibration can be described in other forms as well—such as graphic plots or according to the Hosmer-Lemeshow statistic. Third, if a risk prediction model has a low risk of bias and low concern about applicability, and is accompanied by good predictive performance, we conclude that the model is quite “usable” (i.e. usability = “Yes”). If studies lack assessment of either discrimination or calibration but are judged to be at low risk of bias and with minimal concerns of applicability, then we have designated usability as “maybe” (a modified definition from PROBAST). However, even when models with low concerns of applicability are applied, subtleties of the original population for model development should still be considered. For example, a model developed from either an obstetric or a non-obstetric population may perform less well if applied to a different population. As might a model developed primarily for critically ill patients or non-critically ill patients because data elements often differ substantially in each setting.

Statistical analysis, data synthesis and meta-analysis. The performance of each index was reported as per the original study using the AUC or c-statistic with 95% confidence intervals (CI) or standard error (SE) for discrimination, and standardized mortality ratio or the H-L goodness-of-fit test statistic for calibration [21,22,25]. A meta-analysis was performed if at least three studies evaluating a prediction model were available and reported on the AUC with 95% CIs or SE, or if the AUC could be calculated [18]. The AUC was pooled on the logit scale and the standard errors of the logit transformed AUC were derived from equations previously summarized [18]. We then summarized the AUC using the inverse variance method random-

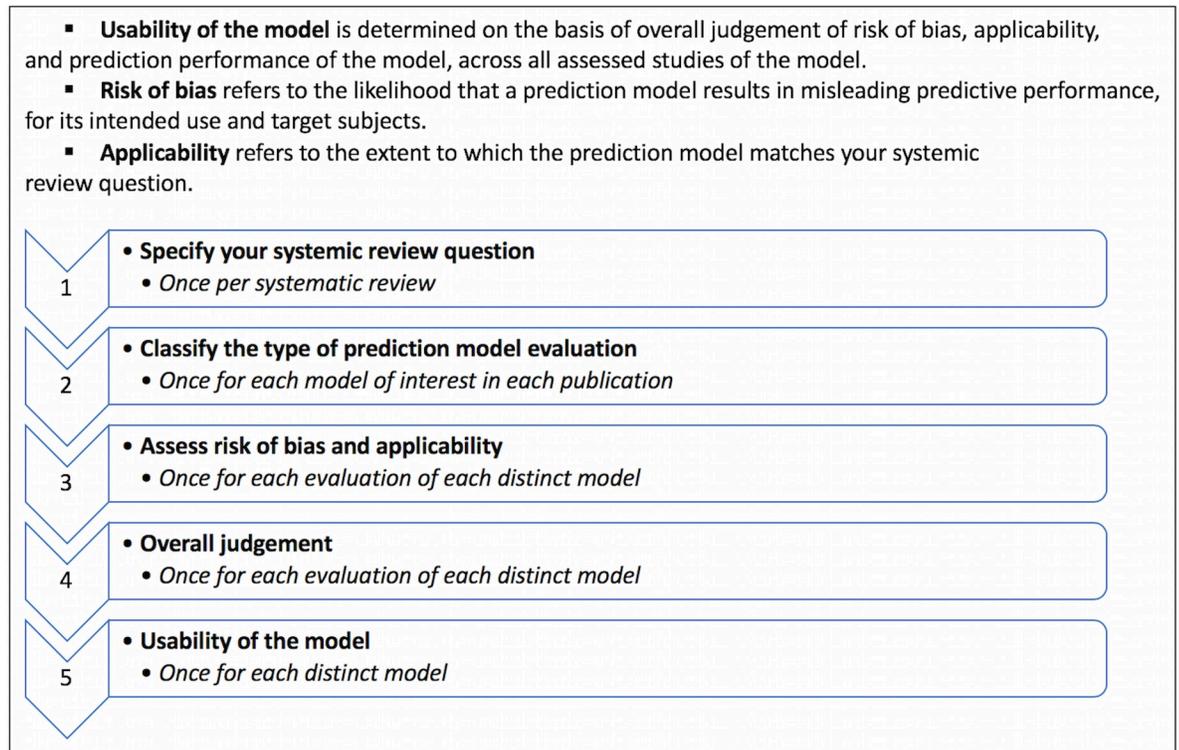


Fig 1. Box summary of PROBABST (Prediction model study Risk Of Bias Assessment Tool).

<https://doi.org/10.1371/journal.pone.0208563.g001>

effects model, estimated with restricted maximum likelihood and Hartung-Knapp-Sidik-Jonkman adjustment to generate 95% CIs [18,26].

Clinical heterogeneity across included studies was assessed by examining details of participants and baseline characteristics. The main sources of heterogeneity we expected to encounter related to differences in patients’ baseline characteristics, care delivery and outcomes. The I^2 statistic was used to explore statistical heterogeneity, defined as moderate when $I^2 = 50-74\%$ and high for $I^2 \geq 75\%$. We planned that if there were more than 10 studies assessing a distinct prediction model for meta-analyses, funnel plots would be drawn to assess the possibility of publication bias [27]. Statistical computations were undertaken using R version 3.4.0 (Free Software Foundation) with R package (meta).

Sensitivity and subgroup analyses. Only one sensitivity analysis was pre-planned and included a sub-set of studies, restricted to include those of higher methodological quality (i.e. low risk of bias). One *post-hoc* sensitivity analysis was conducted according to mortality rate: low (under 1%); moderate (to 10%); high to 20%); and very high (greater than 20%).

Two *post-hoc* subgroup analyses were conducted, to investigate high heterogeneity across the eligible studies. One was to separate the all studies into high-income countries, low-income countries or mixed income-countries. The other was to divide the studies into study setting: ICUs or obstetric general wards.

Results

Our initial search of MEDLINE, EMBASE and Scopus retrieved 9,710 citations, 8,935 of which remained after removal of duplications (Fig 2). Screening of titles and abstracts resulted in 74 relevant articles for which the manuscript full text was assessed for final eligibility. An

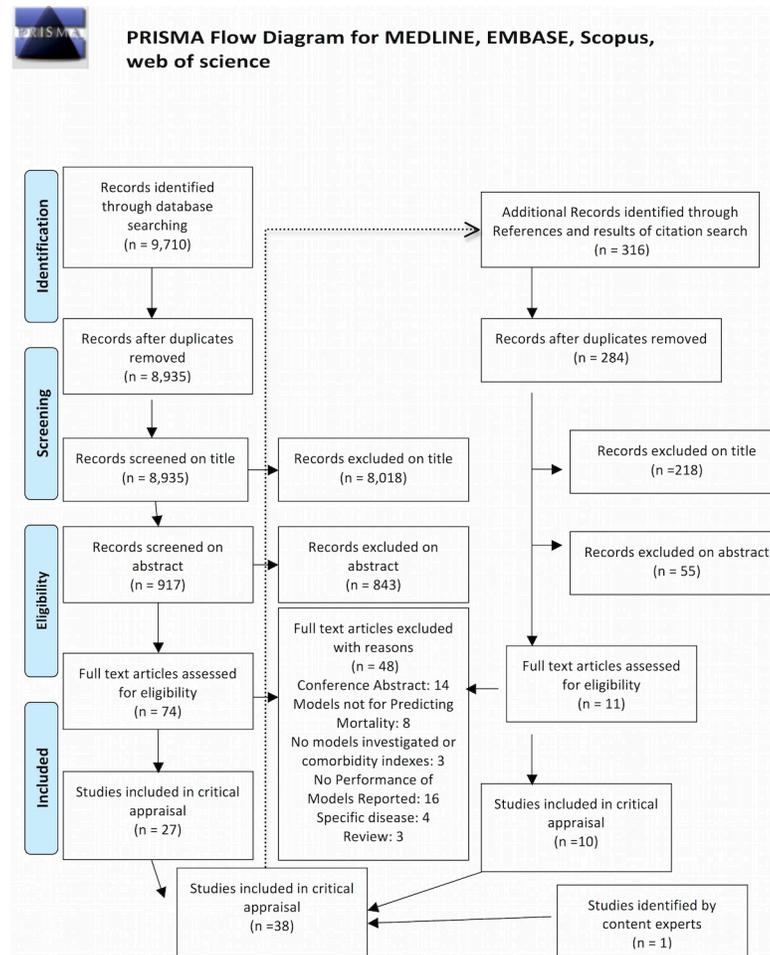


Fig 2. Study selection process (PRISMA flow).

<https://doi.org/10.1371/journal.pone.0208563.g002>

additional 11 articles were identified through the reference lists and citation tracking of 316 relevant articles with use of Web of Science. Finally, 38 studies of 12 prediction models met inclusion criteria (Tables 1 and 2, Figs 3 and 4, and S1 Table) [4,17,36–45,28,46–55,29,56–63,30–35]. The TRIPOD checklist of all 38 eligible articles is available on request (S2 File).

Of these 38, 4 studies both developed *and* validated their model (Table 1) [30,37,54,63], and the remainder (n = 34) were primarily validation studies. Most studies (n = 36) employed a cohort design (prospective in 7, retrospective in 28 and unknown in 1) and 2 used a cross-sectional design (Table 2). Nine studies investigated more than one prediction model in the study. Samples size in included studies ranged from 24 to more than 80,000 subjects. The most commonly reported primary outcome was hospital mortality (n = 14), followed by ICU mortality (n = 13), although eight studies did not specify timing or place of death. Mortality varied across the studies, with an average rate 10.4%, ranging from 0 to 41.7% (Table 2, Fig 3 and Fig 4). Most studies (n = 16) were from a single developed country, 4 were multi-country studies, and 18 were from developing countries.

No uniform measure was reported to quantify the predictive performance (ability to predict outcomes of interest) of eligible models (Fig 3 and Fig 4). Among 20 studies reporting AUC, values ranged from 0.77 to 0.98; 18 studies reported AUCs higher than 0.8, indicating good discriminative performance. SMR was the most commonly reported form of calibration

Table 1. The number and setting of eligible studies in each prediction model, and original population and setting for development of each study.

	Original patient population	Original setting	External Validation in another Obstetric population (number of studies)
Acute Physiology And Chronic Health Evaluation score II & III*	Non-obstetric	ICU	21
Simplified Acute Physiology Score 2 & 3	Non-obstetric	ICU	9
Sepsis-related Organ Failure Assessment	Non-obstetric	ICU	5
Mortality Prediction Model 2 & 3	Non-obstetric	ICU	3
World Health Organization criteria	Obstetric	ICU	1
Obstetric Early Warning Score	Obstetric	ICU	1
Collaborative Integrated Pregnancy High-dependency Estimate of Risk	Obstetric	ICU	0
Maternal Severity Index	Obstetric	General ward	2
Maternal Mortality Score	Obstetric	General ward **	0

ICU = intensive care unit

* include updated tool from original model

** Hospitals in developing countries

<https://doi.org/10.1371/journal.pone.0208563.t001>

(n = 26), ranging from 0 to 1.57. Only 4 studies reported the Hosmer-Lemeshow goodness-of-fit test [50,53,54,61]. Four studies investigated classification measures (e.g. sensitivity and specificity) for a particular cut point of each model with or without reporting discrimination and calibration [4,31,38,43] (Fig 3 and Fig 4).

The methodological quality for each study as determined by PROBAST (Fig 3 and Fig 4, and S3 File) is summarized by a measure of model *Usability* [25] [Fig 1].

Predictive models originally developed from obstetric populations

Five models were developed and validated to identify obstetric patients at risk of death using a combination of comorbid health conditions, clinical characteristics, physiological and laboratory based data: the Collaborative Integrated Pregnancy High-dependency Estimate of Risk (CIPHER), the World Health Organization (WHO) Criteria, Maternal Severity Index, Obstetric Early Warning Score and the Maternal Mortality Score [30,31,37,54,63] (Table 2, Figs 3 and 5).

The Collaborative Integrated Pregnancy High-dependency Estimate of Risk was developed for predicting mortality and prolonged organ support of pregnant and postpartum women. The cohort included individuals in Intensive Care Units of 11 high-, middle- and low-income countries, with an overall mortality rate 9.6% [63]. The final model contains 10 predictors: maternal age, surgery in the preceding 24 hours, systolic blood pressure, Glasgow Coma Scale, serum bilirubin, activated partial thromboplastin time, serum creatinine, potassium, sodium and arterial blood gas pH. Discrimination (AUC: 0.823, 95% CIs 0.811–0.835) and calibration (Graphic plot [intercept=−0.09, slope 0.92]) were internally validated with a bootstrapped sample—in a graphic plot, perfect calibration shows a slope of 1 and intercept of 0. There was no external validation.

The WHO introduced criteria consisting of patient characteristics, vital signs, and physiological and laboratory data to predict maternal severe morbidity. Cecatti et al assessed laboratory- and management-based markers of severity of illness (for example, use of vasopressors, dialysis, ventilation, transfusion, need for hysterectomy, receipt of cardiopulmonary resuscitation) on the risk of maternal death [31]. Using a combined count of all events, sensitivity and

Table 2. Summary of 38 eligible articles.

Prediction models	Country	Study design	Study period	Study population	Sample size	Outcomes
Collaborative Integrated Pregnancy High-dependency Estimate of Risk [Payne]	17 countries	Retrospective cohort*	2000–12	ICU	477	Antepartum and postpartum mortality
Combined WHO criteria: laboratory and management criteria [Cecatti]	Brazil	Retrospective cohort*	2002–7	All hospitalization	673	ICU mortality
Maternal Severity Index [Souza]*****	Brazil	Prospective cohort**	2009–10	All hospitalization	82,388	Hospital mortality
Maternal Severity Index [Souza]	29 countries	Cross-sectional*	2010–11	All hospitalization	314,623	Hospital mortality
Maternal Severity Index [Haddad]*****	Brazil	Cross-sectional*	2009–10	All hospitalization	9,555	Hospital mortality
Obstetric early warning score, Modified Early Obstetric Warning System, the confidential enquiries into maternal death Obstetric EWS, the royal college of physician’s non-obstetric NEWS [Carle]****	United Kingdom	Retrospective cohort**	1995–2008	ICU	4,440	ICU mortality
Obstetric early warning score [Paternina-Caicedo]	Colombia	Retrospective cohort*	2006–11	ICU	702	Antepartum and postpartum mortality
Maternal mortality Score [Huchon]	Senegal and Mali	Prospective cohort**	2007–8	All hospitalization	43,624 for development, 46,328 for validation	Hospital mortality
SOFA [Kallur]	India	Retrospective cohort*	2011–12	ICU	69	Mortality ***
SOFA [Oliveira-Neto]	Brazil	Retrospective cohort*	2002–7	ICU	673	ICU mortality
SOFA [Jain]	India	Prospective cohort*	2010–11	ICU	90	ICU mortality
APACHE II, SOFA [Simsek]	Turkey	Retrospective cohort*	1999–2009	ICU	63	ICU mortality
APACHEII, SOFA [Vasquez]	Argentina	Prospective cohort*	2012	ICU	362	Hospital mortality
APACHE II, SAPS2, MPM2 [el-Solh]	United States	Retrospective cohort*	1989–95	ICU	93	Antepartum and postpartum mortality
MPM2 [Gupta]	India	Retrospective cohort*	2009–10	ICU	24	ICU mortality
SAPS2, 3 MPM2, 3 [Rojas-Suarez]	Colombia	Retrospective cohort*	2006–11	ICU	726	Mortality ***
APACHE II, SAPS2, APACHE III [Hazelgrove]	England	Retrospective cohort*	1994–6	ICU	210	Mortality ***
SAPS2 [Gombar]	India	Retrospective cohort*	2007–12	ICU	151	Mortality ***
APACHE II, SAPS2 [Lapinsky]	6 countries	Retrospective cohort*	1994–8	ICU	332	Hospital mortality
APACHE II SAPS2 [Mjahed]	Morocco	Retrospective cohort*	1995–2002	ICU	364	Hospital mortality
SAPS2 [Gilbert]	United States	Cohort, unknown pro/retro*	1991–1998	ICU	233	Hospital mortality
SAPS2 [Tempe]	India	Retrospective cohort*	2002–04	ICU	57	Hospital mortality
SAPS2 [Togal]	Turkey	Retrospective cohort*	2006–09	ICU	73	Mortality ***

(Continued)

Table 2. (Continued)

Prediction models	Country	Study design	Study period	Study population	Sample size	Outcomes
APACHE II [Afessa]	United States	Retrospective cohort*	1991–1998	ICU	74	Hospital mortality
APACHE II [Aldawood]	Saudi Arabia	Retrospective cohort*	1999–2009	ICU	75	ICU mortality
APACHE II [Bhadade]	India	Prospective cohort*	2009–10	ICU	122	ICU mortality
APACHE II [Harde]	India	Prospective cohort*	2011–12	ICU	61	Mortality ***
APACHE II [Harrison]****	United Kingdom	Retrospective cohort*	1995–2003	ICU	1,902	Hospital mortality
APACHE II [Karnad]	India	Retrospective cohort*	1997–2001	ICU	453	Mortality ***
APACHE II [Lenz]	Austria	Retrospective cohort*	March 1996- Oct 2001, Nov 2004-Jun 2005	ICU	80	Hospital mortality
APACHE II [Lewinsohn]	Israel	Retrospective cohort*	non specific 8 years	ICU	58	Hospital mortality
APACHE II [Mahutte]	Canada	Retrospective cohort*	1992–97	ICU	131	ICU mortality
APACHE II [Muench]	United States	Prospective cohort*	Non specific 2 years	ICU	34	Mortality ***
APACHE II [Tang]	China	Retrospective cohort*	1998–1995	ICU	49	ICU mortality
APACHE II [Thakur]	United States	Retrospective cohort*	2006–12	ICU	69	ICU mortality
APACHE II, updated APACHE II [Paternina-Cacedo]	Colombia	Retrospective cohort*	2006–11	ICU	654	ICU mortality
APACHE II [Vasquez]	Argentina	Retrospective cohort*	1998–2005	ICU	161	ICU mortality
APACHE III [Crozier]	Australia	Retrospective cohort*	2006–8	ICU	60	Hospital mortality

Abbreviations: WHO: World Health Organization, SOFA: Sequential Organ Failure Assessment, APACHE: Acute Physiology And Chronic Health Evaluation, SAPS: Simplified Acute Physiology Score, MPM: Mortality Probability Model

* validation study

** development and validation study

*** without specific time period specified

**** Same cohort from ICNARC (Intensive Care National Audit and Research Center in United Kingdom)

*****Same cohort from The Brazilian Network for Surveillance of Severe Maternal Morbidity

<https://doi.org/10.1371/journal.pone.0208563.t002>

specificity (with at least one criteria of severe morbidity) for predicting death was assessed. Mortality rate of the cohort was 2.67%. Subsequently, Souza et al. validated the WHO criteria's predictive ability using a different dataset in a distinct middle-income country [54]. The Maternal Severity Index was developed on the basis of this validation, incorporating the WHO criteria in addition to other markers of severe illness, for those who were in general wards and also Intensive Care Units [54]. The Maternal Severity Index was subsequently externally validated using multi-national cohorts, demonstrating a good AUC (0.82, 95% CIs 0.78–0.86) and SMR (1.02) [29,49]. In total, 3 studies investigated the Maternal Severity Index, with mortality ranging from 0.15 to 1.47%.

The Obstetric Early Warning Score incorporates vital signs, level of consciousness and oxygen requirements. It was developed and internally validated in the United Kingdom with

Prediction Model	Author	Mortality of entire study cohort (%)	Predictive performance		Risk of Bias	Concern of Applicability	Usability of model for predicting mortality for inpatient pregnant and postpartum women
			Discrimination	Calibration			
Collaborative Integrated Pregnancy High-dependency Estimate of Risk	Payne	9.6	AUC 0.823(0.811-0.835)	Visual assessment with calibration slope (0.92) and intercept (-0.09)	low	low	yes
WHO criteria	Cecatti	2.67	Sensitivity 100% (78.1-100) Specificity 70.4% (66.7-73.8)**	Not reported	high	low	no
Maternal Severity Index	Souza 2012	0.17	AUC 0.954(0.925-0.985)	Hosmer-Lemeshow 0.402, Nagelkerke R square 0.754	low	low	yes
	Souza 2013	0.15	AUC 0.826(0.802-0.851)	Not reported	low	low	
	Haddad	1.47	Not reported	SMR 1.02(0.86-1.20)	low	low	
Obstetric Early Warning Score	Carle	1.6	AUC 0.957(0.923-0.991)	Not reported	high	high	no
	Paternina-Caicedo	4.1	AUC 0.84 (0.75-0.92)	Not reported	low	low	
Maternal Mortality Score	Huchon	0.69 in development, 0.79 in validation	AUC 0.89(0.87-0.91) in development, 0.90(0.89-0.92) in validation	Not reported	low	low	Maybe*

Fig 3. Characteristics of predictive models originally developed from obstetric populations. *PROBAST does not currently provide a “Maybe” option; however, we have added this term for perceived intermediate/potentially usable models for pregnant and post-partum populations. **for at least one of the WHO criteria.

<https://doi.org/10.1371/journal.pone.0208563.g003>

mortality 1.6% and an AUC of 0.957 (95% CIs 0.923–0.991) for predicting death [30]. The study was considered to be at high risk of bias due to a number of excluded participants from the final model. There was high concern of applicability to the obstetric patient population in general wards because the model was developed on the basis of ICU patients and mortality in ICU rather than hospital patients and hospital mortality. Recently, an external validation was performed, demonstrating good predictive ability (AUC: 0.84, 95% CIs 0.75–0.92) for mortality among obstetrics patients, with a mortality rate 4.1% [61]. However, calibration has not yet been evaluated.

The Maternal Mortality Score uses 9 clinical and social conditions, was developed in low- and middle-income countries, and showed good discrimination ability in the development (AUC 0.89, 95% CIs 0.87–0.91, mortality rate: 0.69%) and validation (AUC 0.90, 95% CIs 0.89–0.92, mortality rate: 0.79%) cohorts [37]. The Maternal Mortality Score has only been validated by one study, and calibration has not yet been evaluated.

Of all models developed and validated to identify obstetric patients at risk of death using a combination of comorbid health conditions, clinical characteristics, physiological and laboratory based data, the Collaborative Integrated Pregnancy High-dependency Estimate of Risk and the Maternal Severity Index were developed/validated from studies with a low risk of bias and low concern of applicability of the model to obstetric populations for predicting maternal death, leading to a designation of high “usability” [29,37,54,63] (Fig 3).

Models originally developed primarily from non-obstetric patient populations

Thirty studies explored 7 predictive models developed primarily from non-obstetric patient populations—the Simplified Acute Physiology Score 2 and 3, Acute Physiology and Chronic Health Evaluation Score 2 and 3, Sequential Organ Failure Assessment, and the Mortality Probability Model, versions 2 and 3) [4,17,40–49,28,50–53,55–57,59,60,62,32–36,38,39]

Prediction Model	Author	Mortality of whole study cohort (%)	Predictive performance		Risk of Bias	Concern of Applicability	Usability of model for predicting mortality for inpatient pregnant and postpartum women
			Discrimination	Calibration			
Sequential Organ Failure Assessment	Kallur	5.88	AUC 0.77(0.46-1.00)	Not reported	high	low	no
	Oliveira-Neto	2.67	AUC 0.958(0.914-1.0)	Not reported	high	low	
	Simsek	20.63	Sensitivity 69.2% Specificity 68.0% for SOFA 6.5	Not reported	high	low	
	Vasquez-2	3.59	AUC 0.892(0.831-0.953)	Not reported	low	low	
Mortality Probability Model 2	Jain	33.3	AUC 0.949	Not reported	low	low	no*
	el-Solh	5.88	AUC 0.9147±0.0433	Not reported	low	low	
	Gupta	41.67	AUC 0.74	SMR 1.57	low	low	
Mortality Probability Model 3	Rojas-Suarez	20.63	AUC 0.868 (0.793–0.943)	Hosmer-Lemeshow 0.571 SMR 0.88 (0.60–1.25)	high	low	no
Simplified Acute Physiology Score 2	el-Solh	5.88	AUC 0.9336±0.0216	Not reported	low	low	Maybe**
	Gilbert	3.43	Not reported	SMR 0.42	high	low	
	Gombar	31.13	AUC 0.93(0.89-0.96)	SMR 0.97	low	low	
	Hazलगrove	3.33	AUC 0.92 (0.85–0.99)	SMR 0.43 (0.15–0.93)	high	low	
	Lapinsky	12.05	AUC original 0.784, modified 0.79	SMR 0.886	high	low	
	Mjahed	16.7	Not reported	SMR 0.70	low	low	
	Rojas-Suarez	20.63	AUC 0.865 (0.789–0.941)	Hosmer-Lemeshow 0.030 SMR 0.51 (0.34–0.77)	high	low	
	Tempe	40.35	AUC 0.92	SMR 0.736	low	low	
	Togal	12.33	Not reported	SMR 0.73	high	low	
Simplified Acute Physiology Score 3	Rojas-Suarez	20.63	AUC 0.903 (0.845–0.960)	Hosmer-Lemeshow 0.103 SMR 0.60 (0.39–0.89)	high	low	no***
Acute Physiology And Chronic Health Evaluation Score II	Afessa	2.70	Not reported	SMR 0.15	high	low	no
	Aldawood	8.00	Not reported	SMR 0.36(0.13-0.76)	high	low	
	Bhadade	30.3	Not reported	SMR 0.99	low	low	
	el-Solh	5.88	AUC 0.896±0.0439	Not reported	low	low	
	Harde	6.56	Not reported	SMR 0.42	high	low	
	Harrison	3.05	AUC 0.839(0.820-0.857)	SMR 0.245(0.167-0.346) Hosmer-Lemeshow 166(<.001)	high	low	
	Hazलगrove	3.33	AUC 0.94 (0.86–1.0)	SMR 0.24 (0.09–0.52)	high	low	
	Karnad	21.63	Not reported	SMR 0.78(0.63-0.95)	high	low	
	Lapinsky	12.05	AUC original 0.820, modified 0.847	SMR original 0.432, modified 0.524	high	low	
	Lenz	6.25	Not reported	SMR 0.46	high	low	
	Lewinsohn	6.90	Not reported	SMR (0.416)	high	low	
	Mahutte	2.3	Not reported	SMR 0.23	high	low	
	Mjahed	16.7	Not reported	SMR 0.87	low	low	
	Muench	0	Not reported	SMR 0	high	low	
	Paternina-Calcedo	4.27	original AUC 0.87: 0.80-0.93, adjusted AUC 0.87: 0.80-0.93	original SMR 0.3 (0.2-0.45), adjusted 0.85 (0.56-1.24), original Hosmer-Lemeshow 56.25(p<.001), adjusted Hosmer-Lemeshow 7.85 (.643)	high	low	
Simsek	20.63	Sensitivity 69.2% Specificity 70.0% for APACHE2 12.5	Not reported	high	low		
Tang	4.08	Not reported	SMR 0.22	high	low		
Thakur	4.35	Not reported	SMR 0.64	high	high		
Vasquez-1	11.18	Not reported	SMR obstetric cause 0.37	high	low		
Vasquez-2	3.59	AUC 0.886: 0.827-0.946	SMR 0.47	low	low		
Acute Physiology And Chronic Health Evaluation Score III	Hazलगrove	3.33	AUC 0.98 (0.96–1.0)	SMR 0.43 (0.15–0.93)	high	low	no
Crozier	1.67	Not reported	SMR 0.12	high	low		

Fig 4. Characteristics of predictive models originally developed from non-obstetric population. * Results were inconsistent across the full number of studies examining this model. ** Sensitivity analysis showed good discrimination and calibration, although results were inconsistent across the full number of studies examining this model. *** Due to concern of a high risk of bias due to a low SMR.

<https://doi.org/10.1371/journal.pone.0208563.g004>

(Fig 4)—that were initially developed from non-obstetric critically ill patient populations and incorporated *patient characteristics, comorbidities, physiological and laboratory-based data*. The AUC across these 7 models was near 0.80, demonstrating good discriminative ability. Most studies (n = 25) reported the standardized mortality ratio for calibration, which varied from 0 to 1.57, indicating that some models under-estimate and others over-estimate true mortality (Fig 4). Six studies of the Acute Physiology and Chronic Health Evaluation II [4,17,32,36,50,53], 6 studies of the Simplified Acute Physiology Score 2 [4,32,33,36,42,45], 4 studies of the Sequential Organ Failure Assessment [4,38,41,60] and 3 studies of the Mortality Probability Model [32,34,42] were pooled in separate meta-analyses [Fig 6]. SOFA had the highest discrimination followed by SAPS2 and APACHE II and MPM2. None of models originally developed primarily from non-obstetric patient populations had a designation of high “usability”.

Sensitivity analyses

A preplanned sensitivity analysis was performed, among 9 studies with low risk of bias and low concern of applicability [4,32–34,39,45,57,58,60]. Five studies evaluated APACHE II [4,32,39,57,58], 4 studies evaluated SAPS 2 [32,33,45,57], 2 studies evaluated SOFA [4,60] and

Prediction model	Variables and assigned points (if indicated)																																																																
Collaborative Integrated Pregnancy High-dependency Estimate of Risk	$\text{Logit}(p) = 1.057 + [-0.00109(\text{maternal age}^2)] + [-0.75348(\text{positive surgical status within 24hrs of admission})] + [-0.14360(\text{Glasgow Coma Scale})] + [-0.00839(\text{systolic blood pressure})] + [0.62835(\text{low white blood cell count})] + [-0.20439(\text{high white blood cell count})] + [0.00595(\text{serum bilirubin})] + [0.00414(\text{serum creatinine})] + [0.01732(\text{aPTT})]$ Maternal age (yr); surgical status (Y/N); Glasgow coma scale (ordinal units); systolic blood pressure (mmHg); low white blood cell count ($<5.6 \times 10^9/L$); high white blood cell count ($>17.0 \times 10^9/L$); serum bilirubin ($\mu\text{mol/L}$); serum creatinine ($\mu\text{mol/L}$)																																																																
WHO criteria	<table border="1"> <tr> <td>Cardiovascular dysfunction</td> <td>Shock Lactate > 5 mmol/L Cardiac arrest</td> <td>pH<7.1 Use of continuous vasoactive drugs Cardio-pulmonary resuscitation</td> </tr> <tr> <td>Respiratory dysfunction</td> <td>Acute cyanosis Respiratory rate >40 or <6 per minute SpO₂<90% for ≥ 60 minutes</td> <td>Gasping PaO₂/FiO₂<200mmHg Intubation & ventilation unrelated to anesthesia</td> </tr> <tr> <td>Renal dysfunction</td> <td>Oliguria non-responsive to fluids or diuretics Dialysis for acute renal failure</td> <td>Creatinine > 300 mmol/l or > 3.5 mg/dl</td> </tr> <tr> <td>Coagulation/hematological dysfunction</td> <td>Clotting failure Transfusion ≥ 5 units of blood/red cells</td> <td>Acute thrombocytopenia (<50,000 platelets/μL)</td> </tr> <tr> <td>Hepatic dysfunction</td> <td>Jaundice in the presence of pre-eclampsia</td> <td>Bilirubin>100mmol/l or >6.0mg/dl</td> </tr> <tr> <td>Neurological dysfunction</td> <td>Metabolic coma (loss of consciousness AND the presence of glucose and ketoacids in urine) Stroke</td> <td>Coma / loss of consciousness lasting 12 hours or more Status epilepticus / uncontrollable fits / total paralysis</td> </tr> <tr> <td>Uterine dysfunction</td> <td>Hysterectomy following infection or hemorrhage</td> <td></td> </tr> </table>	Cardiovascular dysfunction	Shock Lactate > 5 mmol/L Cardiac arrest	pH<7.1 Use of continuous vasoactive drugs Cardio-pulmonary resuscitation	Respiratory dysfunction	Acute cyanosis Respiratory rate >40 or <6 per minute SpO ₂ <90% for ≥ 60 minutes	Gasping PaO ₂ /FiO ₂ <200mmHg Intubation & ventilation unrelated to anesthesia	Renal dysfunction	Oliguria non-responsive to fluids or diuretics Dialysis for acute renal failure	Creatinine > 300 mmol/l or > 3.5 mg/dl	Coagulation/hematological dysfunction	Clotting failure Transfusion ≥ 5 units of blood/red cells	Acute thrombocytopenia (<50,000 platelets/ μL)	Hepatic dysfunction	Jaundice in the presence of pre-eclampsia	Bilirubin>100mmol/l or >6.0mg/dl	Neurological dysfunction	Metabolic coma (loss of consciousness AND the presence of glucose and ketoacids in urine) Stroke	Coma / loss of consciousness lasting 12 hours or more Status epilepticus / uncontrollable fits / total paralysis	Uterine dysfunction	Hysterectomy following infection or hemorrhage																																												
Cardiovascular dysfunction	Shock Lactate > 5 mmol/L Cardiac arrest	pH<7.1 Use of continuous vasoactive drugs Cardio-pulmonary resuscitation																																																															
Respiratory dysfunction	Acute cyanosis Respiratory rate >40 or <6 per minute SpO ₂ <90% for ≥ 60 minutes	Gasping PaO ₂ /FiO ₂ <200mmHg Intubation & ventilation unrelated to anesthesia																																																															
Renal dysfunction	Oliguria non-responsive to fluids or diuretics Dialysis for acute renal failure	Creatinine > 300 mmol/l or > 3.5 mg/dl																																																															
Coagulation/hematological dysfunction	Clotting failure Transfusion ≥ 5 units of blood/red cells	Acute thrombocytopenia (<50,000 platelets/ μL)																																																															
Hepatic dysfunction	Jaundice in the presence of pre-eclampsia	Bilirubin>100mmol/l or >6.0mg/dl																																																															
Neurological dysfunction	Metabolic coma (loss of consciousness AND the presence of glucose and ketoacids in urine) Stroke	Coma / loss of consciousness lasting 12 hours or more Status epilepticus / uncontrollable fits / total paralysis																																																															
Uterine dysfunction	Hysterectomy following infection or hemorrhage																																																																
Maternal Severity Index	$\text{Logit} = -7.540 + 0.309(\text{number of life-threatening conditions}^*) + 0.287(\text{Life-threatening conditions identified in the first 24 hours of hospital stay}) - 0.579(\text{Severe pre-eclampsia}) + 3.492(\text{cancer}) + 4.209(\text{any marker of cardiovascular failure}^{**}) + 1.513(\text{any marker of respiratory failure}^{**}) - 1.169(\text{Hysterectomy})$ Maternal Severity Index = $e^{\text{Logit}} / (1 + e^{\text{Logit}})$ If the condition exists, then 1. Otherwise, 0. *: life-threatening conditions are as same as ones listed in WHO criteria **: Presence of any of the following conditions: pH<7.1, Use of continuous vasoactive drugs, Cardiac arrest, Cardio-pulmonary resuscitation ***: Presence of any of the following conditions: Gasping, PaO ₂ /FiO ₂ <200 mmHg, Intubation & ventilation not related to anesthesia																																																																
Obstetric Early Warning Score	<table border="1"> <thead> <tr> <th>Variable\Points</th> <th>3</th> <th>2</th> <th>1</th> <th>0</th> <th>1</th> <th>2</th> <th>3</th> </tr> </thead> <tbody> <tr> <td>Systolic blood pressure (mmHg)</td> <td>< 80</td> <td>80-89</td> <td></td> <td>90-139</td> <td>140-149</td> <td>150-159</td> <td>≥ 160</td> </tr> <tr> <td>Diastolic blood pressure (mmHg)</td> <td></td> <td></td> <td></td> <td>< 90</td> <td>90-99</td> <td>100-109</td> <td>≥ 110</td> </tr> <tr> <td>Respiratory rate (per minute)</td> <td>< 10</td> <td></td> <td></td> <td>10-17</td> <td>18-24</td> <td>25-29</td> <td>≥ 30</td> </tr> <tr> <td>Heart rate (per minute)</td> <td>< 60</td> <td></td> <td></td> <td>60-110</td> <td></td> <td>111-149</td> <td>≥ 150</td> </tr> <tr> <td>% O₂ required to maintain SpO₂≥96%</td> <td></td> <td></td> <td></td> <td>Room air</td> <td>24-39%</td> <td></td> <td>≥ 40%</td> </tr> <tr> <td>Temperature (°C)</td> <td>< 34.0</td> <td></td> <td>34.0-35.0</td> <td>35.1-37.9</td> <td>38.0-38.9</td> <td></td> <td>≥ 39</td> </tr> <tr> <td>Conscious level</td> <td></td> <td></td> <td></td> <td>Alert</td> <td></td> <td></td> <td>Not alert</td> </tr> </tbody> </table>	Variable\Points	3	2	1	0	1	2	3	Systolic blood pressure (mmHg)	< 80	80-89		90-139	140-149	150-159	≥ 160	Diastolic blood pressure (mmHg)				< 90	90-99	100-109	≥ 110	Respiratory rate (per minute)	< 10			10-17	18-24	25-29	≥ 30	Heart rate (per minute)	< 60			60-110		111-149	≥ 150	% O ₂ required to maintain SpO ₂ ≥96%				Room air	24-39%		≥ 40%	Temperature (°C)	< 34.0		34.0-35.0	35.1-37.9	38.0-38.9		≥ 39	Conscious level				Alert			Not alert
Variable\Points	3	2	1	0	1	2	3																																																										
Systolic blood pressure (mmHg)	< 80	80-89		90-139	140-149	150-159	≥ 160																																																										
Diastolic blood pressure (mmHg)				< 90	90-99	100-109	≥ 110																																																										
Respiratory rate (per minute)	< 10			10-17	18-24	25-29	≥ 30																																																										
Heart rate (per minute)	< 60			60-110		111-149	≥ 150																																																										
% O ₂ required to maintain SpO ₂ ≥96%				Room air	24-39%		≥ 40%																																																										
Temperature (°C)	< 34.0		34.0-35.0	35.1-37.9	38.0-38.9		≥ 39																																																										
Conscious level				Alert			Not alert																																																										
Maternal Mortality Score	Parity greater than four: 3 points Fewer than three antenatal visits: 5 points Referral from other health facility: 8 points Malaria: 10 points Uterine rupture: 13 points Genital infection or sepsis: 13 points Preeclampsia or eclampsia: 14 points Antepartum or postpartum hemorrhage: 15 points Severe anemia: 19 points Total maternal mortality score: out of 100 points																																																																

Fig 5. Box summary of variables of predictive models originally developed from obstetric populations.

<https://doi.org/10.1371/journal.pone.0208563.g005>

2 studies evaluated MPM 2 [32,34]. APACHE II showed good discrimination, but over-estimated death. The MPM 2 showed good discrimination, but underestimated death. SOFA showed good discrimination, but calibration was not investigated. AUCs of SAPS 2 were

Study	N	Weight (%)	AUC (95%CI)
APACHE II			
el-Solh 1996	93	3.4%	0.90 (0.81, 0.98)
Hazelgrove 2001	210	1.5%	0.94 (0.86, 1.00)
Harrison 2005	1902	67.0%	0.84 (0.82, 0.86)
Lapinsky 2010	322	12.0%	0.82 (0.74, 0.88)
Paternina-Calcedo 2015	654	8.1%	0.87 (0.80, 0.93)
Vasquez 2015	362	7.9%	0.89 (0.83, 0.95)
APACHE II Overall	3543		0.85 (0.82, 0.88)
Heterogeneity: $I^2=15.7%$ (0.0%, 78.6%)			
SAPS 2			
el-Solh 1996	93	17.1%	0.93 (0.89, 0.98)
Tempe 2007	57	11.9%	0.92 (0.80, 0.97)
Lapinsky 2010	322	21.0%	0.78 (0.70, 0.85)
Hazelgrove 2001	210	13.1%	0.92 (0.85, 0.99)
Gombar 2014	151	19.4%	0.93 (0.89, 0.96)
Rojas-Suarez 2014	554	17.6%	0.87 (0.79, 0.94)
SAPS 2 Overall	1387		0.90 (0.83, 0.94)
Heterogeneity: $I^2=74.0%$ (40.7%, 88.6%)			
SOFA			
Oliveira-Neto 2012	673	23.9%	0.96 (0.91, 1.00)
Kallur 2014	69	6.5%	0.77 (0.46, 1.00)
Vasquez 2015	362	44.3%	0.89 (0.83, 0.95)
Jain 2015	90	25.3%	0.95 (0.87, 0.98)
SOFA Overall	1194		0.92 (0.81, 0.97)
Heterogeneity: $I^2=32.5%$ (0.0%, 76.0%)			
MPM 2			
el-Solh 1996	93	24.3%	0.91 (0.83, 1.00)
Gupta 2011	24	25.2%	0.74 (0.50, 0.89)
Rojas-Suarez 2014	723	50.5%	0.87 (0.79, 0.94)
MPM 2 Overall	840		0.85 (0.59, 0.96)
Heterogeneity: $I^2=34.5%$ (0.0%, 78.7%)			

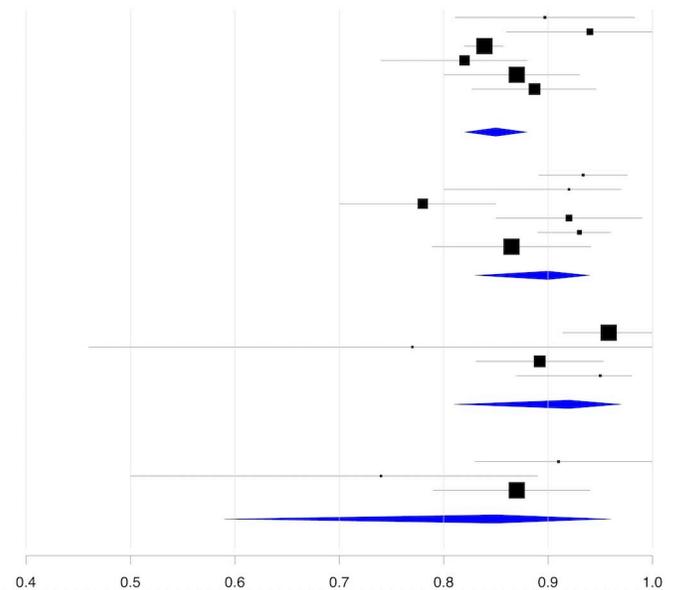


Fig 6. Pooled AUC of APACHE II, SAPS, SOFA and MPM2.

<https://doi.org/10.1371/journal.pone.0208563.g006>

pooled in this sensitivity analysis. The SAPS 2 summary AUC was 0.84 (95% CIs [0.92–0.94] and $I^2 = 0%$) with good discrimination, and calibration [S1 Fig].

A *post-hoc* sensitivity analysis was conducted according to mortality rates. AUCs were not pooled in this analysis because of the limited number of studies for each model. In studies from countries with low mortality—less than 1%—the Maternal Severity Index is the only model that is externally validated. In studies from countries with moderate mortality rates—less than 10%—all predictive models originally developed from non-obstetric population, except SAPS3, were investigated. However, the CIPHER model is the only well validated model for both discrimination and calibration. Only predictive models developed from non-obstetric population were investigated in studies with high and very high mortality. Predictive performance (i.e. discrimination and calibration) of models in studies with high and very high mortality were inconsistent across studies.

Subgroup analyses

Two *post-hoc* subgroup analyses were carried out. In predictive models developed from obstetric population, the CIPHER model and the Maternal Severity Index were investigated in both high- and low-income countries. Studies for predictive models developed from non-obstetric populations were well balanced between high- and low-income countries, and predictive performance seemed similar. The Maternal Severity Index was the only model that was investigated in general obstetric ward patients; therefore, it was difficult to estimate how other settings might affect predictive performance of the models.

Discussion

Main findings

This systematic review identified 38 studies that developed and/or validated 12 models for predicting mortality among hospitalized pregnant and postpartum women. The Collaborative

Integrated Pregnancy High-dependency Estimate of Risk (CIPHER) for hospitalized critically ill obstetric populations, and the Maternal Severity Index for hospitalized general obstetric populations have good discrimination, calibration, were developed from studies with a low risk of bias and internally and/or externally validated for critically ill pregnant and postpartum women. Prediction models developed from non-obstetric patients and from general ICU patient populations have very good discrimination but are at risk of over- or under-estimation of true mortality.

Interpretations

Maternal death is rare event. Hence, predictive models for maternal death seem to show high discriminative performance. In this context, calibration is important in assessing overall predictive performance. However, most of eligible studies in this review reported SMR, which is known to be a relatively crude measure of calibration [18] and ideally should be considered across the full range of outcome rates [64]. The Hosmer-Lemeshow goodness-of-fit test is very sensitive to sample size and also therefore an imperfect measure of calibration [65]. Therefore, calibration in models reporting SMR and Hosmer-Lemeshow goodness-of-fit test need to be interpreted in the context of the setting of the study, in addition to the setting of its application.

The wide range of mortality of the eligible 38 studies helps to explain high heterogeneity we found for various prediction models, across the studies. Also, the studies were assessed in different clinical settings and among patients of varying initial severity of illness, making comparability of predictive performance challenging. For example, a predictive model developed in a country where mortality is high might falsely under- or over-estimate mortality in a country where mortality is much lower, or vice versa. On the other hand, the average mortality among studies of predictive models originally developed from obstetric populations was higher than one from non-obstetric populations. This might explain why the calibration among studies of predictive models from non-obstetric populations appears to be worse than one from obstetric populations.

Implications

In observational studies, a risk adjustment tool is essential to help take into account characteristics (e.g. the severity of a patient's illness) that may influence or confound an attempt to estimate the magnitude and significance of new factor of interest on a *cohort of patients'* clinical outcome such as death. Both CIPHER and the Maternal Severity Index models might also be used to help risk adjust mortality differences between health facilities as part of quality assurance and improvement initiatives [29,54,63].

However, none of the indices studied have sufficient predictive ability to be used in determining the outcome of an *individual* obstetric patient. A risk prediction model revealing a 90% risk of death in a selected population of 100 pregnant or post-partum women cannot differentiate which *individual* 10 women will survive and which will die. That is, there is a risk of underestimating risk in low-risk patients, and of overestimating risk in high-risk patients. Therefore, estimates based on risk prediction models should not directly affect the decisions for withdrawing or withholding management of individual seriously ill pregnant or postpartum women. However, despite uncertainty of predictive performance in assessing individual risk, prediction models may help to identify pregnant and peri-partum women at high risk of critical illness or death and stimulate increased monitoring or preventive measures.

Risk prediction models developed from non-obstetric patient populations should generally not be applied to obstetric patient populations, *if a better alternative exists*. The CIPHER

model for hospitalized critically ill obstetric populations, and the Maternal Severity Index for hospitalized general obstetric populations are suggested for use, when sufficient data exists. In low/middle income countries, because of the large number of variables required for the models, feasibility is a concern, and therefore, the Maternal Mortality Score may be more appropriate. For future research, investigation of few different prediction models within the same population is the ideal study design, to determine the best model for predicting maternal mortality.

Strengths and limitations

This systematic review has a number of strengths. This is the first systematic review of risk prediction models for maternal mortality in obstetric populations. We followed the most recent guideline of systematic reviews for risk prediction models [18] and the PRISMA-P 2015 statement (S4 File) [19]. We employed a formal and broad search strategy, without language restriction and differentiated between risk prediction models for obstetric and non-obstetric populations. We used a robust tool to identify studies at risk of bias using the Prediction model study Risk Of Bias Assessment Tool [25], which allowed us to conduct sensitivity analyses on the most applicable studies for obstetric populations. Next, we have developed our systematic review according to the ROBIS guidelines for detecting bias in systematic reviews (S5 File) [66]. Lastly, our findings are applicable to pregnant and postpartum women who are admitted to either ICUs or acute care wards of general hospitals.

This systematic review also has certain limitations. First, individual patient data from each prediction model evaluation was not available, which precludes an opportunity re-calculate model performance characteristics, and precludes an opportunity to meta-analyze some metrics we evaluated. Second, the data sources used in the various evaluations were derived from diverse clinical settings across and within countries where patient characteristics and clinical practice vary. Yet, this clinical heterogeneity of included studies allows us to make inferences across diverse settings, for patients in an ICU or a hospital. Third, the eligible studies reported mortality at different measurement points (e.g. ICU mortality, hospital mortality), which was challenging to meta-analyze. Fourth, some risk prediction models we identified in the current study could perform differently in certain common diseases of maternal death (e.g. infection) [67], although our findings are likely applicable to most general obstetric populations. Fifth, physiology-based predictions models developed from non-obstetric populations may be challenging to apply to pregnant patients because of changes in physiology (heart rate, blood pressure, respiratory rate for example) that occur as a usual course of pregnancy. Importantly, there were a limited number of studies validating prediction models developed from obstetric population.

Conclusion

Mortality risk prediction models developed from obstetric patient populations, such as the Collaborative Integrated Pregnancy High-dependency Estimate of Risk (CIPHER) model and the Maternal Severity Index, have good discrimination and calibration, developed/validated from studies with a low risk of bias, and should be encouraged for use in prospectively designed studies, trials and quality improvement research among critically ill pregnant and postpartum women. While prediction models previously developed from general and non-obstetric patient populations such as the APACHE, MPM, SAPS, and SOFA scores are at some risk of over- or under-estimating mortality, they generally have good discrimination and may reasonably be used when pregnancy-specific models are unavailable.

Supporting information

S1 Table. Explanation for excluded studies while full text assessment.

(PDF)

S1 Fig. Sensitivity analysis, pooled AUC of SAPS 2.

(TIFF)

S1 File. Search strategy for MEDLINE, EMBASE, PubMed and web of science.

(DOCX)

S2 File. Agreed TRIPOD checklist for all eligible 38 studies.

(XLSX)

S3 File. Agreed risk of bias assessment for all eligible 38 studies (PROBAST).

(XLSX)

S4 File. PRISMA checklist.

(DOC)

S5 File. ROBIS checklist.

(PDF)

Acknowledgments

KA receives Fellowship funding from the Canadian Institutes of Health Research, to pursue the PhD stream of Clinical Epidemiology & Health Care Research at Institute of Health Policy, Management and Evaluation, University of Toronto.

Author Contributions

Conceptualization: Kazuyoshi Aoyama, Joel G. Ray, Damon C. Scales, Stephen E. Lapinsky, Prakesh S. Shah, Robert A. Fowler.

Data curation: Kazuyoshi Aoyama, Rohan D'Souza, Robert A. Fowler.

Formal analysis: Kazuyoshi Aoyama, Ruxandra Pinto, Prakesh S. Shah, Robert A. Fowler.

Funding acquisition: Kazuyoshi Aoyama, Ruxandra Pinto, Joel G. Ray, Damon C. Scales, Stephen E. Lapinsky, Gareth R. Seaward, Michelle Hladunewich, Prakesh S. Shah, Robert A. Fowler.

Investigation: Kazuyoshi Aoyama, Rohan D'Souza, Ruxandra Pinto, Joel G. Ray, Andrea Hill, Damon C. Scales, Stephen E. Lapinsky, Gareth R. Seaward, Michelle Hladunewich, Prakesh S. Shah, Robert A. Fowler.

Methodology: Kazuyoshi Aoyama, Rohan D'Souza, Ruxandra Pinto, Joel G. Ray, Andrea Hill, Damon C. Scales, Stephen E. Lapinsky, Gareth R. Seaward, Michelle Hladunewich, Prakesh S. Shah, Robert A. Fowler.

Project administration: Kazuyoshi Aoyama, Robert A. Fowler.

Resources: Kazuyoshi Aoyama, Rohan D'Souza, Joel G. Ray, Stephen E. Lapinsky, Robert A. Fowler.

Software: Ruxandra Pinto, Prakesh S. Shah, Robert A. Fowler.

Supervision: Joel G. Ray, Damon C. Scales, Stephen E. Lapinsky, Prakesh S. Shah, Robert A. Fowler.

- Validation:** Kazuyoshi Aoyama, Rohan D’Souza, Ruxandra Pinto, Joel G. Ray, Damon C. Scales, Stephen E. Lapinsky, Gareth R. Seaward, Michelle Hladunewich, Robert A. Fowler.
- Visualization:** Kazuyoshi Aoyama, Ruxandra Pinto, Joel G. Ray, Andrea Hill, Damon C. Scales, Stephen E. Lapinsky, Gareth R. Seaward, Robert A. Fowler.
- Writing – original draft:** Kazuyoshi Aoyama, Ruxandra Pinto, Joel G. Ray, Damon C. Scales, Stephen E. Lapinsky, Michelle Hladunewich, Prakesh S. Shah, Robert A. Fowler.
- Writing – review & editing:** Kazuyoshi Aoyama, Rohan D’Souza, Ruxandra Pinto, Joel G. Ray, Andrea Hill, Damon C. Scales, Stephen E. Lapinsky, Gareth R. Seaward, Michelle Hladunewich, Prakesh S. Shah, Robert A. Fowler.

References

1. Pollock W, Rose L, Dennis C. Pregnant and postpartum admissions to the intensive care unit: a systematic review. *Intensive Care Med.* 2010; 36: 1465–74. <https://doi.org/10.1007/s00134-010-1951-0> PMID: 20631987
2. Callaghan WM. Overview of maternal mortality in the United States. *Semin Perinatol.* 2012; 36: 2–6. <https://doi.org/10.1053/j.semperi.2011.09.002> PMID: 22280858
3. Wanderer JP, Leffert LR, Mhyre JM, Kuklina E V, Callaghan WM, Bateman BT. Epidemiology of obstetric-related ICU admissions in Maryland: 1999–2008*. *Crit Care Med.* 2013; 41: 1844–52. <https://doi.org/10.1097/CCM.0b013e31828a3e24> PMID: 23648568
4. Vasquez DN, Das Neves A V, Vidal L, Moseinco M, Lapadula J, Zakalik G, et al. Characteristics, Outcomes, and Predictability of Critically Ill Obstetric Patients: A Multicenter Prospective Cohort Study. *Crit Care Med.* 2015; 43: 1887–1897. <https://doi.org/10.1097/CCM.0000000000001139> PMID: 26121075
5. Einav S, Bromiker R, Sela HY. Maternal Critical Illness. *Curr Anesthesiol Rep. Current Anesthesiology Reports;* 2017; 7: 55–66. <https://doi.org/10.1007/s40140-017-0198-5>
6. Le Gall J-R, Lemeshow S, Saulnier F. Simplified Acute Physiology Score (SAPS II) Based on a European / North American multicenter study. *JAMA.* 1993; 270: 2957–2963. <https://doi.org/10.1001/jama.270.24.2957> PMID: 8254858
7. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med.* 1985; 13: 818–29. PMID: 3928249
8. Teres D, Lemeshow S, Avrunin JS, Pastides H. Validation of the mortality prediction model for ICU patients. *Crit Care Med.* 1987; 15: 208–13. PMID: 3816253
9. Lemeshow S, Teres D, Klar J, Avrunin JS, Gehlbach SH, Rapoport J. Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients. *JAMA.* 1993; 270: 2478–2486. <https://doi.org/10.1001/jama.270.20.2478> PMID: 8230626
10. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med.* 1996; 22: 707–10. PMID: 8844239
11. Ryan HM, Sharma S, Magee LA, Ansermino JM, MacDonell K, Payne BA, et al. The Usefulness of the APACHE II Score in Obstetric Critical Care: A Structured Review. *J Obstet Gynaecol Canada. Elsevier Inc;* 2016; 38: 909–918. <https://doi.org/10.1016/j.jogc.2016.06.013> PMID: 27720089
12. Von Dadelszen P, Payne B, Li J, Ansermino JM, Pipkin FB, Côté AM, et al. Prediction of adverse maternal outcomes in pre-eclampsia: Development and validation of the fullPIERS model. *Lancet. Elsevier Ltd;* 2011; 377: 219–227. [https://doi.org/10.1016/S0140-6736\(10\)61351-7](https://doi.org/10.1016/S0140-6736(10)61351-7) PMID: 21185591
13. Prick BW, Schuit E, Mignini L, Jansen AJG, van Rhenen DJ, Steegers EAP, et al. Prediction of escape red blood cell transfusion in expectantly managed women with acute anaemia after postpartum haemorrhage. *BJOG.* 2015; 122: 1789–97. <https://doi.org/10.1111/1471-0528.13224> PMID: 25600160
14. Bateman BT, Mhyre JM, Hernandez-Diaz S, Huybrechts KF, Fischer MA, Creanga AA, et al. Development of a comorbidity index for use in obstetric patients. *Obstet Gynecol.* 2013; 122: 957–65. <https://doi.org/10.1097/AOG.0b013e31828a603bb> PMID: 24104771
15. Metcalfe A, Lix LM, Johnson JA, Currie G, Lyon AW, Bernier F, et al. Validation of an obstetric comorbidity index in an external population. *BJOG An Int J Obstet Gynaecol.* 2015; 122: 1748–1755. <https://doi.org/10.1111/1471-0528.13254> PMID: 25559311

16. Aoyama K, D'Souza R, Inada E, Lapinsky SE, Fowler RA. Measurement properties of comorbidity indices in maternal health research: a systematic review. *BMC Pregnancy Childbirth*. 2017; 17: 372. <https://doi.org/10.1186/s12884-017-1558-3> PMID: 29132349
17. Lapinsky SE, Hallett D, Collop N, Drover J, Lavercombe P, Leeman M, et al. Evaluation of standard and modified severity of illness scores in the obstetric patient. *J Crit Care*. Elsevier B.V.; 2011; 26: 535.e1–7. <https://doi.org/10.1016/j.jcrc.2010.10.003> PMID: 21106337
18. Debray TPA, Damen JAAG, Snell KIE, Ensor J, Hooft L, Reitsma JB, et al. A guide to systematic review and meta-analysis of prediction model performance. *Bmj*. 2017; i6460. <https://doi.org/10.1136/bmj.i6460> PMID: 28057641
19. 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. <https://doi.org/10.1186/2046-4053-4-1> PMID: 25554246
20. Nassar AP, Malbouisson LMS, Moreno R. Evaluation of simplified acute physiology score 3 performance: a systematic review of external validation studies. *Crit Care*. 2014; 18: R117. <https://doi.org/10.1186/cc13911> PMID: 24906651
21. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015; 350: g7594. <https://doi.org/10.1136/bmj.g7594> PMID: 25569120
22. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *AnnInternMed*. 2015; 162: W1–73. <https://doi.org/10.7326/M14-0698> PMID: 25560730
23. Hayden JA, Côté P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. *Ann Intern Med*. 2006; 144: 427–37. PMID: 16549855
24. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. Elsevier Inc; 2010; 63: 737–45. <https://doi.org/10.1016/j.jclinepi.2010.02.006> PMID: 20494804
25. Wolff R. Prediction risk of bias assessment tool (PROBAST). Kleijnen Syst Rev Ltd, 2013 [accessed 22/07/2015].
26. IntHout J, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med Res Methodol*. BMC Medical Research Methodology; 2014; 14: 25. <https://doi.org/10.1186/1471-2288-14-25> PMID: 24548571
27. Sterne J, Egger M, Moher D. Addressing reporting biases. In: Higgins J, Green S, editors. *Cochrane Handbook for Systematic Reviews of Intervention Version 510*. The Cochrane Collaboration, 2011.; 2011.
28. Aldawood A. Clinical characteristics and outcomes of critically ill obstetric patients: A ten-year review. *Ann Saudi Med*. 2011; 31: 518–522. <https://doi.org/10.4103/0256-4947.84631> PMID: 21911991
29. Souza JP, Gülmezoglu AM, Vogel J, Carroli G, Lumbiganon P, Qureshi Z, et al. Moving beyond essential interventions for reduction of maternal mortality (the WHO Multicountry Survey on Maternal and Newborn Health): A cross-sectional study. *Lancet*. 2013; 381: 1747–1755. [https://doi.org/10.1016/S0140-6736\(13\)60686-8](https://doi.org/10.1016/S0140-6736(13)60686-8)
30. Carle C, Alexander P, Columb M, Johal J. Design and internal validation of an obstetric early warning score: Secondary analysis of the Intensive Care National Audit and Research Centre Case Mix Programme database. *Anaesthesia*. 2013; 68: 354–367. <https://doi.org/10.1111/anae.12180> PMID: 23488833
31. Cecatti JG, Souza JP, Oliveira Neto AF, Parpinelli MA, Sousa MH, Say L, et al. Pre-validation of the WHO organ dysfunction based criteria for identification of maternal near miss. *Reprod Health*. BioMed Central Ltd; 2011; 8: 22. <https://doi.org/10.1186/1742-4755-8-22> PMID: 21810265
32. el-Solh AA, Grant BJ. A comparison of severity of illness scoring systems for critically ill obstetric patients. *Chest*. 1996; 110: 1299–304. PMID: 8915238
33. Gombar Satinder, Ahuja Vanita and AJ. A retrospective analysis of obstetric patient's outcome in intensive care unit of a tertiary care center. *J Anesth Clin Pharmacol*. 2014; 30: 502–507. <https://doi.org/10.4103/0970-9185.142843> PMID: 25425775
34. Gupta S, Naithani U, Doshi V, Bhargava V, Vijay BS. Obstetric critical care: A prospective analysis of clinical characteristics, predictability, and fetomaternal outcome in a new dedicated obstetric intensive care unit. *Indian J Anaesth*. 2011; 55: 146–153. <https://doi.org/10.4103/0019-5049.79895> PMID: 21712871

35. Harde M, Dave S, Wagh S, Gujjar P, Bhadade R, Bapat A. Prospective evaluation of maternal morbidity and mortality in post-cesarean section patients admitted to postanesthesia intensive care unit. *J Anaesthesiol Clin Pharmacol*. 2014; 30: 508–513. <https://doi.org/10.4103/0970-9185.142844> PMID: [25425776](https://pubmed.ncbi.nlm.nih.gov/25425776/)
36. Hazelgrove JF, Price C, Pappachan VJ, Smith GB. Multicenter study of obstetric admissions to 14 intensive care units in southern England. *Crit Care Med*. 2001; 29: 770–5. PMID: [11373467](https://pubmed.ncbi.nlm.nih.gov/11373467/)
37. Huchon C, Dumont A, Traore M, Abrahamowicz M, Fauconnier A, Fraser W, et al. A prediction score for maternal mortality in Senegal and Mali. *Obstet Gynecol*. 2013; 121: 1049–1056. <https://doi.org/10.1097/AOG.0b013e31828b33a4> PMID: [23635742](https://pubmed.ncbi.nlm.nih.gov/23635742/)
38. Kallur SD, Patil Bada V, Reddy P, Pandya S, Nirmalan PK. Organ dysfunction and organ failure as predictors of outcomes of severe maternal morbidity in an obstetric intensive care unit. *J Clin Diagn Res*. 2014; 8: OC06–8. <https://doi.org/10.7860/JCDR/2014/8068.4213> PMID: [24959481](https://pubmed.ncbi.nlm.nih.gov/24959481/)
39. Karnad DR, Lapsia V, Krishnan A, Salvi VS. Prognostic factors in obstetric patients admitted to an Indian intensive care unit. *Crit Care Med*. 2004; 32: 1294–1299. <https://doi.org/10.1097/01.CCM.0000128549.72276.00> PMID: [15187509](https://pubmed.ncbi.nlm.nih.gov/15187509/)
40. Muench M V, Baschat AA, Malinow AM, Mighty HE. Analysis of disease in the obstetric intensive care unit at a university referral center: a 24-month review of prospective data. *J Reprod Med*. 2008; 53: 914–20. PMID: [19160649](https://pubmed.ncbi.nlm.nih.gov/19160649/)
41. Oliveira-Neto A, Parpinelli M a., Cecatti JG, Souza JP, Sousa MH. Sequential Organ Failure Assessment Score for Evaluating Organ Failure and Outcome of Severe Maternal Morbidity in Obstetric Intensive Care. *Sci World J*. 2012; 2012: 1–8. <https://doi.org/10.1100/2012/172145> PMID: [22454600](https://pubmed.ncbi.nlm.nih.gov/22454600/)
42. Rojas-Suarez J, Paternina-Caicedo AJ, Miranda J, Mendoza R, Dueñas-Castel C, Bourjeily G. Comparison of severity-of-illness scores in critically ill obstetric patients: a 6-year retrospective cohort. *Crit Care Med*. 2014; 42: 1047–54. <https://doi.org/10.1097/CCM.000000000000124> PMID: [24394629](https://pubmed.ncbi.nlm.nih.gov/24394629/)
43. Şimşek T, Eylgör C, Uyar M, Karaman S, Moral AR. Retrospective review of critically ill obstetrical patients: a decade ' s experience. *J Med*. 2011; 41: 1059–1064. <https://doi.org/10.3906/sag-1009-5>
44. Tang LCH, Kwok ACW, Wong AYK, Lee YY, Sun KO, So APC. Critical care in obstetrical patients: An eight-year review. *Chin Med J (Engl)*. 1997; 110: 936–941.
45. Tempe A, Wadhwa L, Gupta S, Bansal S, Satyanarayana L. Prediction of mortality and morbidity by simplified acute physiology score II in obstetric intensive care unit admissions. *Indian J Med Sci*. 2007; 61: 179–85. PMID: [17401254](https://pubmed.ncbi.nlm.nih.gov/17401254/)
46. Togal T, Yucel N, Gedik E, Gulhas N, Toprak HI, Ersoy MO. Obstetric admissions to the intensive care unit in a tertiary referral hospital. *J Crit Care*. Elsevier Inc.; 2010; 25: 628–33. <https://doi.org/10.1016/j.jcrc.2010.02.015> PMID: [20381297](https://pubmed.ncbi.nlm.nih.gov/20381297/)
47. Afessa B, Green B, Delke I, Koch K. Systemic inflammatory response syndrome, organ failure, and outcome in critically ill obstetric patients treated in an ICU. *Chest*. 2001; 120: 1271–7. PMID: [11591571](https://pubmed.ncbi.nlm.nih.gov/11591571/)
48. Gilbert TT, Smulian JC, Martin AA, Ananth C V., Scorza W, Scardella AT. Obstetric admissions to the intensive care unit: Outcomes and severity of illness. *Obstet Gynecol*. 2003; 102: 897–903. [https://doi.org/10.1016/S0029-7844\(03\)00767-1](https://doi.org/10.1016/S0029-7844(03)00767-1) PMID: [14672460](https://pubmed.ncbi.nlm.nih.gov/14672460/)
49. Haddad SM, Cecatti JG, Souza JP, Sousa MH, Parpinelli MA, Costa ML, et al. Applying the Maternal Near Miss Approach for the Evaluation of Quality of Obstetric Care: A Worked Example from a Multicenter Surveillance Study. *Biomed Res Int*. 2014; 2014: 1–10. <https://doi.org/10.1155/2014/989815> PMID: [25147830](https://pubmed.ncbi.nlm.nih.gov/25147830/)
50. Harrison D, Penny J, Yentis S, Fayek S, Brady A. Case mix, outcome and activity for obstetric admissions to adult, general critical care units: a secondary analysis of the ICNARC Case Mix Programme Database. *Crit Care*. 2005; 9: S25–S37. <https://doi.org/10.1186/cc3782> PMID: [16221316](https://pubmed.ncbi.nlm.nih.gov/16221316/)
51. Lenz A, Rabal M, Pateisky N, Gustorff B, Lehner R. PCharacteristics and morbidity of obstetric patients requiring intensive care. *Geburtsh Frauenhelik*. 2007; 67: 1120–1125. <https://doi.org/10.1055/s-2007-965965>
52. Lewinsohn G, Herman A, Leonov Y, Klinowski E. Critically ill obstetrical patients: outcome and predictability. *Crit Care Med*. 1994; 22: 1412–4. PMID: [8062563](https://pubmed.ncbi.nlm.nih.gov/8062563/)
53. Paternina-Caicedo AJ, Rojas-Suarez JA, Duenas-Castel C, Miranda-Quintero JE, Bourjeily G. Mortality Risk Prediction With an Updated Acute Physiology and Chronic Health Evaluation II Score in Critically Ill Obstetric Patients: A Cohort Study. *J Intensive Care Med*. 2015; 30: 97–102. <https://doi.org/10.1177/0885066613502450> PMID: [24004907](https://pubmed.ncbi.nlm.nih.gov/24004907/)
54. Souza JP, Cecatti JG, Haddad SM, Parpinelli MA, Costa ML, Katz L, et al. The WHO Maternal Near-Miss Approach and the Maternal Severity Index Model (MSI): Tools for Assessing the Management of Severe Maternal Morbidity. *PLoS One*. 2012; 7. <https://doi.org/10.1371/journal.pone.0044129> PMID: [22952897](https://pubmed.ncbi.nlm.nih.gov/22952897/)

55. Vasquez DN, Estenssoro E, Canales HS, Reina R, Saenz MG, Das Neves A V., et al. Clinical characteristics and outcomes of obstetric patients requiring ICU admission. *Chest*. 2007; 131: 718–724. <https://doi.org/10.1378/chest.06-2388> PMID: 17356085
56. Thakur M, Gonik B, Gill N, Awonuga AO, Rocha FG, Gonzalez JM. Intensive Care Admissions in Pregnancy: Analysis of a Level of Support Scoring System. *Matern Child Health J*. Springer US; 2016; 20: 106–113. <https://doi.org/10.1007/s10995-015-1808-9> PMID: 26318180
57. Mjahed K, Hamoudi D, Salmi S, Barrou L. Obstetric patients in a surgical intensive care unit: prognostic factors and outcome. *J Obs Gynaecol*. 2006; 26: 418–423. <https://doi.org/10.1080/01443610600720188> PMID: 16846867
58. Bhadade R, More A, de' Souza R, Harde M. Maternal outcomes in critically ill obstetrics patients: A unique challenge. *Indian J Crit Care Med*. 2012; 16: 8. <https://doi.org/10.4103/0972-5229.94416> PMID: 22557826
59. Mahutte NG, Murphy-Kaulbeck L, Le Q, Solomon J, Benjamin A, Boyd ME. Obstetric admissions to the intensive care unit. *Obstet Gynecol*. 1999; 94: 263–6. [https://doi.org/10.1016/S0029-7844\(99\)00274-4](https://doi.org/10.1016/S0029-7844(99)00274-4) PMID: 10432140
60. Jain S, Guleria K, Suneja A, Vaid NB, Ahuja S. Use of the Sequential Organ Failure Assessment score for evaluating outcome among obstetric patients admitted to the intensive care unit. *Int J Gynecol Obstet*. International Federation of Gynecology and Obstetrics; 2016; 132: 332–336. <https://doi.org/10.1016/j.ijgo.2015.08.005> PMID: 26792141
61. Paternina-Caicedo A, Miranda J, Bourjeily G, Levinson A, Dueñas C, Bello-Muñoz C, et al. Performance of the Obstetric Early Warning Score in critically ill patients for the prediction of maternal death. *Am J Obstet Gynecol*. Elsevier Inc.; 2017; 216: 58.e1–58.e8. <https://doi.org/10.1016/j.ajog.2016.09.103> PMID: 27751799
62. Crozier TM, Wallace EM. Obstetric admissions to an integrated general intensive care unit in a quaternary maternity facility. *Aust New Zeal J Obstet Gynaecol*. 2011; 51: 233–238. <https://doi.org/10.1111/j.1479-828X.2011.01303.x> PMID: 21631442
63. Payne B, Ryan H, Magee L, Von Dadelszen P. Precision medicine for pregnant and recently pregnant women receiving critical care: "the CIPHER (Collaborative Integrated Pregnancy High-dependency Estimate of Risk) model. *Crit Care*. 2018; (in press).
64. Kipnis P, Liu V, Escobar GJ. Accuracy of hospital standardized mortality rates: Effects of model calibration. *Med Care*. 2014; 52: 378–384. <https://doi.org/10.1097/MLR.000000000000111> PMID: 24848209
65. Paul P, Pennell ML, Lemeshow S. Standardizing the power of the Hosmer-Lemeshow goodness of fit test in large data sets. *Stat Med*. 2013; 32: 67–80. <https://doi.org/10.1002/sim.5525> PMID: 22833304
66. Whiting P, Savović J, Higgins JPT, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. 2016; 69: 225–234. <https://doi.org/10.1016/j.jclinepi.2015.06.005> PMID: 26092286
67. Aarvold ABR, Ryan HM, Magee LA, Von Dadelszen P, Fjell C, Walley KR. Multiple Organ Dysfunction Score Is Superior to the Obstetric-Specific Sepsis in Obstetrics Score in Predicting Mortality in Septic Obstetric Patients. *Crit Care Med*. 2017; 45: e49–e57. <https://doi.org/10.1097/CCM.0000000000002018> PMID: 27618276