BMC Pregnancy and Childbirth

RESEARCH

Open Access

Mortality-weighted severe maternal morbidity: a novel approach to assessing maternal health outcomes

Deborah Kilday[1], Drew Tatum[2], Michael Korvink[2*], Kathy W. Belk[2], Joseph IV Beals[2], Ashley Finke[2], Raymond Perigard[2], Stephanie Martin[3], Richelle Marshall[1], Yue Jiang[4], Joseph G. Ibrahim[5], John Yeh[6] and Dorothy A. Fink[1]

Abstract

Objective The aim of this study is to develop a robust mortality-weighted severe maternal morbidity (SMM) composite measure that provides increased clinical granularity by measuring the individual and combined risk of inpatient mortality across multiple SMM indicators. The enhanced metric better reflects the complexity gradient of individual patients through a weighted SMM (w-SMM) composite score expressed on a continuous scale.

Methods This retrospective cohort study analyzed 7,174,412 inpatient deliveries from 2016 to 2023 extracted from the Premier Healthcare Database, an all-payer, nationally representative administrative database. Harm weights for each SMM indicator were calculated using the conditional probability of in-hospital maternal mortality given each SMM indicator. The resulting mortality-based harm weights were employed to combine individual SMM indicators into a composite index. The index was scaled to create the in-hospital w-SMM composite score, ranging from 0 to 1. Out-of-sample performance of the proposed w-SMM and established binary SMM (b-SMM) indices was evaluated using an 80:20 derivation-validation split framework.

Results An SMM indicator was identified in 96% of 433 inpatient maternal mortality cases in the study population. The incidence and crude mortality percentages varied by SMM indicator, with cardiac arrest/ventricular fibrillation and conversion of cardiac rhythm having the highest rates of mortality percentages (33% and 29%, respectively), and therefore, the highest severity weights. Approximately 21% of patients identified with an SMM indicator had multiple SMM indicators. The w-SMM composite score, scaled between 0 and 1, accounts for the specific additive risk of multiple indicators, with higher scores indicating higher mortality risk. The w-SMM demonstrated stronger association with in-hospital mortality than the b-SMM measure, with a precision-recall area under the curve (PR-AUC) of 0.269, compared to the current binary composite PR-AUC of 0.007.

Conclusion The mortality-weighted w-SMM index improves upon the binary SMM composite measure by accounting for the varying severity of individual and combined SMM indicators and can further serve as a more granular measure of in-hospital mortality risk. The w-SMM index can be used retrospectively to stratify high-risk

*Correspondence:
Michael Korvink
michael_korvink@premierinc.com

Full list of author information is available at the end of the article

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

patients, enhance risk adjustment models, and support comparative outcomes research that informs public health policy development, future research, policy decisions, and clinical practice.

Keywords Severe maternal morbidity, Index, Maternal mortality risk, Harm-weighting, Maternal health

Introduction

The US Centers for Disease Control (CDC) describes severe maternal morbidity (SMM) as "*unexpected outcomes of labor and delivery that can result in significant short- or long-term health consequences*" [1]. From 2008 to 2021, SMM has increased by 22% in the United States (U.S.) [2]. In the U.S. Surgeon General's *Call to Action* to improve maternal health, the U.S. Secretary of Health and Human Services designated maternal mortality and morbidity as a public health crisis [3].

While there is no global standard for defining SMM, various composite measures have been developed using different conceptual rubrics focused on maternal mortality, resource utilization, or long-term health outcomes [4–7]. As a result, the conditions included in SMM can vary, leading to differences in reported SMM rates across studies and settings. Historically, in the United States, the most widely used measure of SMM is the CDC's binary composite measure, which is based on the presence of one or more of 21 conditions and procedures (including blood transfusion) defined using International Classification of Diseases, Tenth Revision, Clinical Modification diagnosis codes (ICD-10-CM) and ICD-10 Procedure Coding System (ICD-10-PCS) procedure codes. In recent years, the CDC has excluded blood transfusion from the SMM definition because transfusion alone often does not reflect a severe maternal complication. Therefore, the CDC-defined binary composite currently comprises 20 conditions [5, 8].

While defining SMM as a binary composite measure offers simplicity and convenience, it fails to account for the varying degrees of risk associated with individual SMM indicators, as well as the increased risk from the co-occurrence of multiple SMM indicators. While researchers frequently address the challenge of analyzing rare events in healthcare research by grouping such events into binary composite measures, including those employed in obstetric and maternal studies, such measures often suffer from methodological shortcomings [9–11]. As a binary measure, the CDC's approach to SMM obscures underlying drivers of variation in maternal mortality and lacks sufficient granularity to stratify across a spectrum of maternal acuity [12, 13]. This lack of specificity limits the interpretation of differences in outcomes, such as mortality or resource utilization, among patients defined by the presence of SMM.

This study enhances the existing CDC binary SMM composite measure, referred to here as b-SMM, by developing a continuous, patient-level, and weighted SMM index (w-SMM). While b-SMM captures only the presence or absence of any of the 20 SMM indicators, the w-SMM assigns each indicator a mortality-based harm weight, resulting in a continuous SMM score. The w-SMM score ranges from 0 to 1 and is constructed by aggregating the mortality-based weights. This captures a continuous spectrum of risk, unlike the b-SMM, which assigns only two risk levels, and allows for more granular risk stratification. This approach aligns with the CDC's maternal condition framework while offering a more nuanced and holistic understanding of SMM risk, particularly in the presence of multiple indicators.

The w-SMM is designed to improve upon the existing binary composite indicator by severity-adjusting the underlying SMM components, thus resulting in a finer-grain morbidity composite. Just as the binary SMM composite can be risk-standardized to produce a homogeneous indicator of quality that accounts for case-mix differences across facilities, providers, and service lines, the w-SMM index can also be risk-standardized to address the unique needs of the research question. The unadjusted w-SMM index, as presented here, can provide not only a better understanding of individual mortality risk, but also insight into changes in the risk associated with specific SMM indicators as well as their co-occurrence. As a continuous, mortality-weighted score, the w-SMM supports population-level risk stratification and may further inform maternal care intervention and maternal health public policy.

Methods

Data source

All analyses in this study were based on administrative data extracted from the Premier Healthcare Database (PHD). Data in the PHD are statistically de-identified in compliance with the Health Insurance Portability and Accountability Act (HIPAA). This study is exempt from institutional review board approval under U.S. Title 45 Code of Federal Regulations, Part 46.

The PHD is a comprehensive all-payer administrative database that accounts for approximately 25% of annual inpatient admissions in the United States and includes more than 1,400 geographically diverse hospitals and health systems. In addition to patient demographics and visit characteristics, the PHD includes complete ICD-10 diagnosis and procedure codes [14]. This study adhered to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for observational studies.

Study population

This retrospective observational study evaluated 7,174,412 hospital inpatient delivery discharges from 2016 to 2023 for patients aged 10 to 64 years. Delivery visits were identified using the CDC's SMM birth volume denominator definition, which includes Medicare Severity-Diagnosis Related Groups (MS-DRG) and ICD-10 diagnosis or procedure codes indicative of a delivery visit. Patients with abortive outcomes (e.g., ectopic pregnancy, spontaneous abortion) were excluded using the CDC's SMM ICD-10-based denominator exclusion criteria [15]. Additionally, the study population was limited to live and stillbirth deliveries completing at least 20 weeks of gestation, excluding those with unknown weeks of gestation. Each discharge record represented a delivery hospitalization at a single hospital. The final study population includes delivery hospitalizations from 49 U.S. states and the District of Columbia, representing a wide geographic and demographic cross-section of birthing individuals. Definitions of delivery and abortive outcomes, used for inclusion and exclusion, are provided in Supplemental Table 1.

Study variables

The primary endpoint used to develop the w-SMM composite was in-hospital maternal mortality, identified through an "expired" discharge status on the Uniform Billing (UB) 04 form at the hospital where the delivery occurred. Each of the 20 SMM indicators was extracted based on its corresponding CDC ICD-10 definition [8]. Consistent with the CDC definition, a binary SMM composite (b-SMM), based on the presence of at least one individual SMM indicator, was calculated as a comparator to the w-SMM. The complete list of SMM indicators, along with their respective ICD-10 diagnosis and procedure codes, is provided in Supplemental Table 2.

Framework for w-SMM composite

The w-SMM composite was developed using the conditional probability of in-hospital mortality given each SMM indicator *j*. This approach accounts for the unique incidence and severity of each SMM indicator.

$$c_j = Pr\left(mortality|\, SMM_j\right)$$

Consistent with approaches used in other harm-weighted outcome indices, each indicator-specific weight c_j (i.e., conditional probability) was further normalized by dividing each weight by the sum of all 20 SMM indicator weights c_k. This ensured the composite ranged between 0 and 1, with the 20 normalized SMM weights, w_j, summing to one [16].

$$w_j = \frac{c_j}{\sum_{k=1}^{20} c_k}$$

The final hospitalization-level harm-weighted SMM composite, or w-SMM, was calculated for each patient *i*, as a weighted sum of the normalized weight w_j when the corresponding SMM condition is present. An example of the calculation for a mock patient is shown in Supplemental Document 1.

$$wSMM_i = \sum_{j=1}^{20} w_j SMM_{i,j}$$

Treating each SMM indicator as a separate contribution to the final composite assumes independent additive risk among the indicators. To assess the potential limitation of this additive assumption, we employed a partial least squares (PLS) regression model to compare composite weights. PLS regression captures latent variables by maximizing the variance among the SMM indicators [17]. These latent variables enable the model to account for potential interactions and address multicollinearity by constructing orthogonal vectors based on these latent factors, albeit at the cost of a more complex framework. The top three components, selected for their strong covariance with mortality and ability to capture the majority of its total variability, were used to construct the PLS-based composite. For each SMM indicator, the corresponding regression coefficients from these components were averaged and then summed to create a single composite score, which was further normalized.

The approach was evaluated using an 80:20 derivation-validation split framework. Since inpatient mortality is a highly imbalanced outcome due to its rarity, random stratified sampling by mortality status was implemented when splitting the dataset. Unique SMM indicator weights were calculated using the 80% derivation dataset. The derived weights were then applied to patient data in both the derivation and validation datasets to produce w-SMM composite scores.

To evaluate the utility of the w-SMM composite relative to the existing binary composite, b-SMM, separate logistic regression models were fit, regressing the inpatient mortality indicator on the respective SMM composite indices.

Traditional evaluation methods, such as receiver operating characteristic (ROC) curves or simple prediction accuracy, can overestimate performance when evaluating rare events like mortality, due to the overrepresentation of true negatives among the predicted results [18]. Therefore, precision-recall (PR) curves were used to assess the area under the curve (AUC). Unlike ROC curves, PR curves are less prone to bias in these scenarios, as

they focus on the balance between precision and recall, explicitly emphasizing the minority class (i.e., mortality). Although the PR-AUC offers a useful summary measure of model performance, direct comparison between the w-SMM and b-SMM should be interpreted with caution. The b-SMM has only two distinct predicted probabilities of mortality, conditional on the binary SMM, whereas the w-SMM incorporates a continuous risk gradient comprising a broad range of predicted probabilities. This wider distribution of risk highlights the advantage of using a continuous composite measure for risk detection. A baseline AUC was calculated based on the crude mortality rate in the test dataset. As such, a PR-AUC greater than the baseline indicates improved utility in distinguishing the outcome beyond random chance.

A final weighted composite was calculated for the entire study population using weights derived from the entire dataset. Patients with no observed SMM indicators received a score of 0, and those with all 20 SMM indicators received a score of 1. Although a score of 1 is theoretically possible, it is very rarely, if ever, observed. For improved interpretability of the risk gradient, the w-SMM values were binned into severity levels. Bins were defined manually using easily interpretable breakpoints (e.g., 0.025, 0.05, 0.1, etc.), guided by the frequency of each bin, with case counts and mortality percentages calculated across severity levels.

To assess the robustness of the w-SMM composite for mortality risk detection and risk adjustment, the existing patient-level w-SMM scores were stratified by race, age group, and payer. A reduced set of w-SMM bins was used to assess the crude mortality rates within each stratum.

Sensitivity analysis

In addition to the recent exclusion of blood transfusion, other modifications have been proposed to refine the severity of the SMM composite measure. In 2024 the Agency for Healthcare Research and Quality (AHRQ) released a refined SMM measure (MHI-03) in their Quality Indicator software, modifying the acute renal failure (ARF) definition so that an ARF diagnosis is documented only when paired with at least one dialysis procedure and omitting three diagnosis codes from the disseminated intravascular coagulation (DIC) definition. These modified definitions resulted in fewer overall SMM cases; however, rates of ICU admission, Cesarean delivery, in-hospital mortality, and extended length of stay were higher using the narrower AHRQ-modified definition [19].

A sensitivity analysis was conducted to assess the impact of refinements to the AHRQ SMM indicators. A modified final weighted composite was calculated using the alternative MHI-03 definitions for ARF and DIC, and the relative importance of the indicators was compared between the AHRQ and CDC definitions. Statistical analyses were performed using R software, version 4.4.1.

Results

This study population included 7,174,412 unique inpatient delivery discharges from 2016 to 2023. Supplemental Table 3 provides summary statistics for the demographic characteristics of the study population. Half of all deliveries occurred among women aged 25 to 33. Women identifying as White accounted for 63% of the cohort, while those identifying as Black and Asian accounted for 15% and 5%, respectively. Commercial insurance and Medicaid covered 52% and 42% of deliveries, respectively. Vaginal births accounted for 68% of deliveries, while 32% were by Cesarean section. As shown in Table 1 and 96% of the in-hospital maternal deaths were identified as having at least one SMM indicator beforehand (416 of 433 maternal mortality cases). Among all deliveries in this study population, 79 per 10,000 had experienced an SMM indicator event, and 6 per 100,000 resulted in hospital inpatient maternal mortality.

Incidence, co-occurrence, and mortality among SMM indicators

The incidence of each SMM indicator differed, as well as the maternal mortality percentage when the SMM indicator was present. DIC, ARF, and hysterectomy procedures were the three most frequently occurring SMM indicators. Cardiac arrest/ventricular fibrillation, conversion of cardiac rhythm, and heart failure arrest during the procedure or surgery exhibited the highest inpatient mortality percentages (Table 2).

Among patients experiencing SMM in the study population, 21% experienced co-occurring SMM indicators, and 2.4% had five or more SMM indicators. Further, individuals with multiple SMM indicators experienced 10 to 354 times higher unadjusted odds of in hospital mortality compared to those with one SMM indicator (Table 3).

Final weighted SMM composite using CDC-defined conditions

Cardiac arrest/ventricular fibrillation and conversion of cardiac rhythm had the highest w-SMM weights, while eclampsia and hysterectomy procedures had the lowest

Table 1 In-hospital morality by SMM status

SMM	Number of Discharges	In-Hospital Mortality Count	In-Hospital Mortality Percentage
No	7,117,800	17	0.0002%
Yes	56,612	416	0.7348%

Table 2 SMM indicator incidence and in-hospital mortality percentage

SMM Indicator	Cumulative Incidence (per 10k Discharges)	In-Hospital Mortality Count	Mortality Percentage When SMM Present
Cardiac Arrest/Ventricular Fibrillation	755 (1.1)	249	32.98%
Conversion of Cardiac Rhythm	701 (1.0)	203	28.96%
Heart Failure/Arrest During Surgery or Procedure	71 (0.1)	16	22.54%
Amniotic Fluid Embolism	401 (0.6)	60	14.96%
Temporary Tracheostomy	198 (0.3)	22	11.11%
Ventilation	4,347 (6.1)	283	6.51%
Acute Myocardial Infarction	370 (0.5)	15	4.05%
Shock	5,144 (7.2)	170	3.3%
Acute Respiratory Distress Syndrome	9,214 (12.8)	272	2.95%
Aneurysm	233 (0.3)	5	2.15%
Air and Thrombotic Embolism	2,251 (3.1)	47	2.09%
Severe Anesthesia Complications	432 (0.6)	9	2.08%
Puerperal Cerebrovascular Disorders	2,446 (3.4)	42	1.72%
Sepsis	7,063 (9.8)	118	1.67%
Acute Renal Failure	11,486 (16.0)	150	1.31%
Disseminated Intravascular Coagulation	14,537 (20.3)	164	1.13%
Pulmonary Edema/Acute Heart Failure	5,049 (7.0)	56	1.11%
Sickle Cell Disease with Crisis	688 (1.0)	5	0.73%
Hysterectomy	9,252 (12.9)	61	0.66%
Eclampsia	5,076 (7.1)	26	0.51%

Table 3 Inpatient mortality rate by number of SMM indicators

Number of SMM Indicators	Number of Discharges	Number of Mortalities	Mortality Percentage
0	7,117,800	17	< 0.001%
1	44,448	20	0.045%
2	6,590	31	0.470%
3	2,790	54	1.935%
4	1,425	94	6.596%
5+	1,359	217	15.968%

weights. Table 4 shows the complete set of normalized weights for each SMM indicator. Table 5 details the relationship between the w-SMM severity level and the percentage of in-hospital maternal mortality.

In-hospital mortality rates exhibited a monotonic increase across w-SMM severity levels. Similarly, as shown in Supplemental Table 4, the transfer-out rate followed a similar pattern across most severity levels, with both mortality and transfer-out rates increasing as the w-SMM value approached the highest severity bin. Additionally, Supplemental Table 5 shows that each SMM indicator had a distinct transfer rate. The majority of indicators had a relative transfer rank within three positions of their mortality rank. The largest discrepancies between transfer and mortality ranks were observed for severe anesthesia complications (difference of 6), puerperal cerebrovascular disorders (difference of 5), and temporary tracheostomy (difference of 4). Notably, more than 50% of temporary tracheostomy events resulted in a transfer.

Validity and utility of weighted SMM composite

Inpatient mortality occurred in 6 per 100,000 deliveries in the dataset. Within the randomly stratified derivation and validation datasets, there were 5,739,530 and 1,434,882 deliveries, respectively.

The PR-AUC for the w-SMM composite was 0.269 (Supplemental Fig. 1). The w-SMM had a higher PR-AUC than the b-SMM measure of 0.007 and baseline PR-AUC of 0.00006 (i.e., the prevalence of mortality in the validation dataset). Using the full study cohort, patient-level w-SMM values were stratified by race, age group, and payer. Across all stratification levels, mortality percentages increased with higher w-SMM severity bins, consistent with the pattern observed in the overall unstratified complete study cohort (Supplemental Tables 6–8).

Normalized weight comparison: conditional probability vs. PLS regression

A separate set of normalized SMM indicator weights was derived from the PLS regression model (Supplemental

Table 4 Final w-SMM composite weights

SMM Indicator	Normalized Weight
Cardiac Arrest/Ventricular Fibrillation	0.2314
Conversion of Cardiac Rhythm	0.2032
Heart Failure/Arrest During Surgery or Procedure	0.1581
Amniotic Fluid Embolism	0.1050
Temporary Tracheostomy	0.0780
Ventilation	0.0457
Acute Myocardial Infarction	0.0284
Shock	0.0232
Acute Respiratory Distress Syndrome	0.0207
Aneurysm	0.0151
Air and Thrombotic Embolism	0.0147
Severe Anesthesia Complications	0.0146
Puerperal Cerebrovascular Disorders	0.0120
Sepsis	0.0117
Acute Renal Failure	0.0092
Disseminated Intravascular Coagulation	0.0079
Pulmonary Edema/Acute Heart Failure	0.0078
Sickle Cell Disease with Crisis	0.0051
Hysterectomy	0.0046
Eclampsia	0.0036

Fig. 2). The three latent factors used to derive the PLS index captured 39% of the variability of the SMM indicators and 90% of the variability of mortality explained by all 20 indicators. Both approaches identified cardiac arrest/ventricular fibrillation and conversion of cardiac rhythm as the two primary contributors to the w-SMM composite, accounting for more than one-third of the total composite weight. The largest absolute difference in normalized weights between the two approaches was observed for heart failure/arrest during surgery or procedure (difference of 0.12) and ventilation (difference of 0.04). The PR-AUC for the PLS index was 0.279, consistent with the PR-AUC of the w-SMM composite.

AHRQ MHI-03 modified DIC and ARF definitions

In a sensitivity analysis using the modified AHRQ SMM indicator definitions, the prevalence of DIC and ARF in this study population (the two most frequently occurring CDC-defined SMM indicators) decreased by factors of 3 and 30 (Supplemental Table 9). With their reduced diagnostic frequency, maternal mortality rates for the modified definition DIC and ARF cases increased by 2- and 7-fold, respectively. Using the modified AHRQ definition, the relative rank of the DIC w-SMM weight shifted from 16th to 11th, and the ARF rank changed from 15th to 6th, indicating an increase in w-SMM weights across both indicators with the modified criteria (Supplemental Table 10).

Discussion

This study proposes and evaluates a novel composite measure of SMM, w-SMM, designed to improve upon existing binary SMM approaches. The method reported here assigns mortality-informed weights to individual SMM indicators, which serve as inputs to a singular continuous measure of severe maternal morbidity. This approach contrasts with existing SMM composite measures that collapse multiple individual SMM indicators into an overall binary classification. Furthermore, the enhanced approach considers the additive effect of co-occurring SMM indicators, thus recognizing that the presence of multiple conditions can compound and considerably elevate overall mortality risk. As such, the proposed approach yields a finer-grained SMM composite with improved maternal mortality predictive accuracy compared to the existing CDC-defined binary SMM composite. Further, the w-SMM index enables more specific stratification across the maternal mortality risk gradient. Additionally, this study highlights the potential underestimation of mortality among severe cases of SMM, due to unobserved mortality among patients transferred to higher acuity care settings (Supplemental Table 4).

Implementation and applicability

The continuous mortality-informed SMM index presented in this study provides the ability to differentiate mortality risk more precisely and is a necessary step in understanding how preexisting risk factors are associated with in-hospital maternal mortality and the incidence of SMM. The w-SMM demonstrated consistent performance across key demographic strata, including race, age group, and payer, supporting its applicability in diverse

Table 5 Inpatient mortality by w-SMM severity level

SMM Severity Level	Number of Discharges	In-Hospital Mortality Count	Mortality Percentage
[0]	7,117,800	17	< 0.001%
(0-0.025)	46,632	9	0.019%
[0.025–0.05)	4,689	14	0.299%
[0.05–0.1)	3,316	47	1.417%
[0.1–0.2)	872	49	5.619%
[0.2–0.3)	416	54	12.981%
[0.3–0.5)	393	128	32.570%
[0.5-1]	294	115	39.116%

patient populations. SMM has long-term health consequences, such as an increased risk of mortality beyond delivery into the postpartum period, as well as a higher likelihood of developing chronic conditions like cardiovascular disease [1, 20, 21]. The w-SMM index, paired with preexisting comorbidities, can further serve as a foundational element in assessing post-partum outcome risk. From a policy perspective, this measure of SMM could inform national or institutional guidelines for maternal care.

While SMM is strongly associated with inpatient mortality, research has shown that, after adjusting for patient and hospital characteristics, SMM rates have increased amid declining mortality rates [2]. This pattern may suggest that the severity distribution of SMM may be decreasing. The w-SMM index, as a continuous measure, provides greater delineation of the changes in overall SMM-related severity and may help in understanding the origins of such trends.

Ongoing development of SMM measures

The refinement and improvement of SMM measures remains a key focus in maternal health research. Several research groups have also proposed alternative indices to enhance the CDC's SMM definition for hospital performance assessment. These variations have included adding serious in-hospital events such as ICU utilization, risk adjustment using pre-existing conditions, and incorporating factors such as patient demographics, social risk factors, and other clinical measures [22–24]. However, to the authors' knowledge, the approach employed in this study is unique in creating a continuous SMM composite score that incorporates the CDC's current list of indicators, while addressing indicator-specific severity.

Additionally, organizations such as AHRQ have recently refined some SMM indicator definitions, resulting in lower incidence but higher adverse outcome rates for indicators like DIC and ARF compared to the standard CDC definitions. Sensitivity analysis here demonstrated that the modified and more stringent definitions of DIC and ARF elevated their relative severity rankings among SMM indicators compared to the current CDC definitions.

Limitations

Despite its strengths, this study has several limitations. All data were derived from administrative records, which are subject to coding errors and inconsistencies. Furthermore, as demonstrated by the sensitivity analysis, w-SMM weights are sensitive to differences in the set of included SMM indicators and their respective definitions. Similarly, the severity weights are derived from conditional probabilities of mortality observed during inpatient delivery hospitalizations with specific inclusion and exclusion criteria. While this study leverages a large, diverse dataset of more than 7 million delivery encounters across 49 U.S. states and the District of Columbia, differences in population characteristics, care settings, or inclusion criteria in other studies may influence the relative frequency of SMM and associated mortality events. As such, caution is warranted when applying the current weighting scheme to substantially different populations or healthcare contexts. Despite these limitations, this study highlights the differences in risks associated with different SMM conditions in relation to inpatient mortality.

A key limitation of this analysis is the restriction to delivery hospitalizations in the inpatient setting, which excludes outcomes that could have occurred after transfer to another care facility. This may result in the potential for unobserved mortality among high-risk patients transferred to higher-acuity care settings, as elevated transfer rates to other facilities were observed across more severe w-SMM levels. Transfers may vary widely in clinical acuity and purpose, ranging from emergency transfers for specialized maternal-fetal care to transfers for closer proximity to the patient's established residence. Further studies should characterize these nuances in transfer types and integrate linked data across facilities to more accurately capture the full spectrum of delayed maternal outcomes post discharge. Adopting a competing risks framework, jointly modeling the risks of transfers and mortality, could offer a more comprehensive understanding of these maternal outcomes and better quantify the true mortality burden associated with SMM. Additionally, further research is needed to explore the impact of the w-SMM measure on postpartum mortality, as studies have shown that the distribution of SMM indicators may differ between the delivery hospital admission episode and the postpartum period [25].

Additionally, a common concern with composite indices is the potential to overestimate risk when combining highly correlated indicators [26]. To mitigate this concern, alternative statistical methods, such as partial least squares, was considered, yielding an overall performance comparable to the current w-SMM composite based on their PR-AUC values. In both approaches, the majority of the normalized weights were distributed similarly across the SMM indicators, with a high correlation (r >0.8) between weighting schemes. Although the statistical performance of both indices was comparable, non-trivial differences in the rank and magnitude of the indicator weights were observed between the two methods, suggesting some degree of duplicative information content across the indicators—potentially due to unmeasured causal pathways between indicators. While the PLS-based approach offers a more robust method for accounting for shared information across indicators, it relies on

methods that are more opaque and less interpretable to clinical stakeholders. The parsimonious approach used here was preferred, given its similar performance, greater interpretability, and ease of implementation, at the expense of some statistical precision at the indicator level. Although some differences in weights may signal interactions between specific SMM indicators, the additive assumption remains a justifiable and effective method. Future research examining interactions among SMM indicators could provide additional clinical insight.

Conclusion

This study highlights the utility of an enhanced SMM composite measure (w-SMM) that offers a weighted view of SMM severity, which was developed in relation to inpatient maternal mortality. Beyond retrospective validation, the w-SMM composite could benefit from prospective benchmarking against additional endpoints such as ICU utilization, length of stay, and quality of life assessments. As a finer-grained metric, the w-SMM supports the stratification of high-risk patient populations and informs future research, policy decisions, and clinical practice, enabling improved outcome comparison and quality improvement efforts.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12884-025-08526-8.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Supplementary Material 4.

Supplementary Material 5.

Supplementary Material 6.

Supplementary Material 7.

Supplementary Material 8.

Supplementary Material 9.

Supplementary Material 10.

Supplementary Material 11.

Supplementary Material 12.

Supplementary Material 13.

Acknowledgements

Tara McFarlin (Premier) assisted with project management, manuscript revision, analysis, and technology support. Cate Polacek (Premier) contributed to manuscript revision. Chelsea Wagner, Candice Michael, and Rachel Sheehan (Premier) provided clinical guidance.

Authors' contributions

Conceptualization, D.K., D.T., M.K., K.B., J.B., and D.F.; methodology, D.K., D.T., M.K., S.M., K.B., J.B., Y.J., J.I., J.Y.; formal analysis, D.T., M.K., K.B., J.B., A.F., R.P.; data curation, D.T., M.K., A.F., R.P., and K.B.; writing—original draft preparation, D.K., D.T., M.K., K.B., and J.B.; writing—review and editing, D.K., D.T., M.K., K.B., J.B., A.F., S.M., R.P., R.M., J.Y., and D.F.; supervision, D.K., M.K., S.M., R.M., Y.J., J.I., J.Y., and D.F. All authors have read and agreed to the published version of the manuscript.

Funding

The US Department of Health and Human Services Office on Women's Health supported this project under contract 75P00120C00066.

Data availability

The de-identified dataset used in this study is subject to ethical, legal, and data disclosure restrictions. While the dataset has been anonymized, it contains potentially sensitive patient information that could be indirectly identifiable. These restrictions are in place to comply with privacy regulations (e.g., HIPAA). Additionally, the data is owned by Premier, Inc. and is governed by their Institutional Review Board (IRB) or Research Ethics Committee (REC). The dataset, as disclosed in the manuscript, includes aggregated data from at least five different healthcare facilities or health systems to ensure privacy and prevent identification of individuals. However, these confidentiality safeguards cannot be maintained if patient-level data is shared, as such disclosure could lead to potential re-identification, especially when the number of facilities represented is limited. Being a commercially available database, access to the PHD is associated with a financial charge. Requests for access to the de-identified dataset may be submitted to the following data access committee or institutional body: Michael Korvink. Email: michael_korvink@premierinc.com. Phone:1 (704) 816-5312.

Declarations

Ethics approval and consent to participate

The PHD is considered exempt from Institutional Review Board (IRB) oversight under Title 45 Code of Federal Regulations, Part 46 of the United States, specifically 45 CFR 46.101(b)(4). In accordance with the HIPAA Privacy Rule, disclosed data from the PHD are considered de-identified per 45 CFR 164.506(d)(2)(ii)(B) through the "Expert Determination" method.
Because this study involved only de-identified secondary data, informed consent from participants was not required and was waived in accordance with U.S. federal regulations (45 CFR 46.104(d)(4)). No contact with human subjects occurred, and no identifiable private information was accessed. This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki and the Belmont Report, as applicable to research using de-identified secondary data.

Consent for publication

Not applicable.

Competing interests

The authors of this manuscript have the following competing interests: This study was supported through a contractual agreement between the U.S. Department of Health and Human Services (HHS) and Premier, Inc. (contract 75P00120C00066), which provided financial compensation for the research. HHS participated in the study design, data collection, analysis, data interpretation, manuscript review, and the decision to publish. The authors declare no other competing interests.

Author details

[1]U.S. Department of Health & Human Services, Office of the Assistant Secretary for Health, Office on Women's Health, Rockville, MD 20852, USA
[2]Advisory Services, Premier Inc., 13520 Ballantyne Corporate Place, Charlotte, NC 28277, USA
[3]Clinical Concepts in Obstetrics, Scottsdale, AZ 85250, USA
[4]Department of Statistical Science, Duke University, Durham, NC 27708, USA
[5]Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
[6]Department of Obstetrics, Gynecology, and Reproductive Biology, Harvard Medical School, Boston, MA 02115, USA

Received: 12 August 2025 / Accepted: 21 November 2025
Published online: 25 November 2025

References

1. Centers for Disease Control and Prevention. Severe Maternal Morbidity. Maternal Infant Health. 2024. https://www.cdc.gov/maternal-infant-health/php/severe-maternal-morbidity/index.html. Accessed 4 Dec 2024.
2. Fink DA, Kilday D, Cao Z, et al. Trends in maternal mortality and severe maternal morbidity during Delivery-Related hospitalizations in the united States, 2008 to 2021. JAMA Netw Open. 2023;6(6):e2317641. https://doi.org/10.1001/jamanetworkopen.2023.17641.
3. Mehta LS, Sharma G, Creanga AA, et al. Call to action: maternal health and saving mothers: A policy statement from the American heart association. Circulation. 2021;144(15):e251–69. https://doi.org/10.1161/CIR.0000000000001000.
4. Dzakpasu S, Deb-Rinker P, Arbour L, et al. Severe maternal morbidity surveillance: monitoring pregnant women at high risk for prolonged hospitalisation and death. Paediatr Perinat Epidemiol. 2020;34(4):427–39. https://doi.org/10.1111/ppe.12574.
5. Snowden JM, Lyndon A, Kan P, El Ayadi A, Main E, Carmichael SL. Severe maternal morbidity: A comparison of definitions and data sources. Am J Epidemiol. 2021;190(9):1890–7. https://doi.org/10.1093/aje/kwab077.
6. Kern-Goldberger AR, Howell EA, Srinivas SK, Levine LD. What we talk about when we talk about severe maternal morbidity: a call to action to critically review severe maternal morbidity as an outcome measure. Am J Obstet Gynecol MFM. 2023;5(5):100882. https://doi.org/10.1016/j.ajogmf.2023.100882.
7. Geller SE, Koch AR, Garland CE, MacDonald EJ, Storey F, Lawton B. A global view of severe maternal morbidity: moving beyond maternal mortality. Reprod Health. 2018;15(1):98. https://doi.org/10.1186/s12978-018-0527-2.
8. Centers for Disease Control and Prevention. Identifying Severe Maternal Morbidity (SMM). Maternal Infant Health. 2024. https://www.cdc.gov/maternal-infant-health/php/severe-maternal-morbidity/icd.html. Accessed 4 Dec 2024.
9. Cordoba G, Schwartz L, Woloshin S, Bae H, Gøtzsche PC. Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. BMJ. 2010;341:c3920. https://doi.org/10.1136/bmj.c3920.
10. Gleason JL, Grewal J, Chen Z, Cernich AN, Grantz KL. Risk of adverse maternal outcomes in pregnant women with disabilities. JAMA Netw Open. 2021;4(12):e2138414. https://doi.org/10.1001/jamanetworkopen.2021.38414.
11. Herman D, Lor KY, Qadree A, Horn D, D'Souza R. Composite adverse outcomes in obstetric studies: a systematic review. BMC Pregnancy Childbirth. 2021;21(1):107. https://doi.org/10.1186/s12884-021-03588-w.
12. McKenna SP, Heaney A. Composite outcome measurement in clinical research: the triumph of illusion over reality? J Med Econ. 2020;23(10):1196–204. https://doi.org/10.1080/13696998.2020.1797755.
13. Chou D, Tunçalp Ö, Firoz T, et al. Constructing maternal morbidity - towards a standard tool to measure and monitor maternal health beyond mortality. BMC Pregnancy Childbirth. 2016;16:45. https://doi.org/10.1186/s12884-015-0789-4.
14. Premier Applied Sciences. Best in Class: Data that Informs and Performs. https://products.premierinc.com/downloads/PremierHealthcareDatabaseWhitepaper.pdf. Accessed 5 Dec 2024.
15. Alliance for Innovation on Maternal Health. AIM Guide to Implementing the Severe Maternal Morbidity Algorithm. AIM. 2022. https://saferbirth.org/wp-content/uploads/AIMData_SMMGuide_5CR_042922.pdf. Accessed 5 Dec 2024.
16. Agency for Healthcare Research and Quality. Quality Indicator User Guide: Inpatient Quality Indicators (IQI) Composite Measures. 2024. https://qualityindicators.ahrq.gov/Downloads/Modules/IQI/v2024/IQI_Composite_Measures.pdf. Accessed 5 Dec 2024.
17. Tobias RD. An Introduction to Partial Least Squares Regression. https://stats.oarc.ucla.edu/wp-content/uploads/2016/02/pls.pdf. Accessed 4 Dec 2024.
18. Saito T, Rehmsmeier M. The Precision-Recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS ONE. 2015;10(3):e0118432. https://doi.org/10.1371/journal.pone.0118432.
19. Agency for Healthcare Research and Quality. Maternal Health Indicators. https://qualityindicators.ahrq.gov/measures/mhi_resources. Accessed 6 Dec 2024.
20. Ukah UV, Dayan N, Potter BJ, Ayoub A, Auger N. Severe maternal morbidity and risk of mortality beyond the postpartum period. Obstet Gynecol. 2021;137(2):277. https://doi.org/10.1097/AOG.0000000000004223.
21. Ukah UV, Dayan N, Potter BJ, Paradis G, Ayoub A, Auger N. Severe maternal morbidity and Long-Term risk of cardiovascular hospitalization. Circ Cardiovasc Qual Outcomes. 2022;15(2):e008393. https://doi.org/10.1161/CIRCOUTCOMES.121.008393.
22. Lazariu V, Nguyen T, McNutt LA, Jeffrey J, Kacica M. Severe maternal morbidity: A population-based study of an expanded measure and associated factors. PLoS ONE. 2017;12(8):e0182343. https://doi.org/10.1371/journal.pone.0182343.
23. Fridman M, Korst LM, Reynen DJ, et al. Using potentially preventable severe maternal morbidity to monitor hospital performance. Jt Comm J Qual Patient Saf. 2023;49(3):129–37. https://doi.org/10.1016/j.jcjq.2022.11.007.
24. Centers for Medicare & Medicaid Services. Severe Obstetric Complications Electronic Clinical Quality Measure (eCQM) Methodology Report. CMS. 2021. https://www.cms.gov/files/document/measure-methodology-report.pdf. Accessed 4 Dec 2024.
25. Chen J, Cox S, Kuklina EV, Ferre C, Barfield W, Li R. Assessment of incidence and factors associated with severe maternal morbidity after delivery discharge among women in the US. JAMA Netw Open. 2021;4(2):e2036148. https://doi.org/10.1001/jamanetworkopen.2020.36148.
26. Organisation for Economic Co-operation and Development. Handbook on Constructing Composite Indicators: Methodology and User Guide. Organisation for Economic Co-operation and Development. 2008. https://www.oecd-ilibrary.org/economics/handbook-on-constructing-composite-indicators-methodology-and-user-guide_9789264043466-en. Accessed 4 Dec 2024.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.