healthcare

MDPI

Article

Leveraging Machine Learning to Predict and Assess Disparities in Severe Maternal Morbidity in Maryland

Qingfeng Li [1,*], Y. Natalia Alfonso [1], Carrie Wolfson [1], Khyzer B. Aziz [2] and Andreea A. Creanga [1,3]

[1] Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA; cwolfso2@jhu.edu (C.W.)
[2] Johns Hopkins Children's Center, Johns Hopkins School of Medicine, Baltimore, MD 21205, USA
[3] Department of Gynecology and Obstetrics, Johns Hopkins School of Medicine, Baltimore, MD 21205, USA
* Correspondence: qli28@jhu.edu; Tel.: +1-4432874738

Abstract: Background: Severe maternal morbidity (SMM) is increasing in the United States. The main objective of this study is to test the use of machine learning (ML) techniques to develop models for predicting SMM during delivery hospitalizations in Maryland. Secondarily, we examine disparities in SMM by key sociodemographic characteristics. Methods: We used the linked State Inpatient Database (SID) and the American Hospital Association (AHA) Annual Survey data from Maryland for 2016–2019 (N = 261,226 delivery hospitalizations). We first estimated relative risks for SMM across key sociodemographic factors (e.g., race, income, insurance, and primary language). Then, we fitted LASSO and, for comparison, Logit models with 75 and 18 features. The selection of SMM features was based on clinical expert opinion, a literature review, statistical significance, and computational resource constraints. Various model performance metrics, including the area under the receiver operating characteristic curve (AUC), accuracy, precision, and recall values were computed to compare predictive performance. Results: During 2016–2019, 76 per 10,000 deliveries (1976 of 261,226) were in patients who experienced an SMM event. The Logit model with a full list of 75 features achieved an AUC of 0.71 in the validation dataset, which marginally decreased to 0.69 in the reduced model with 18 features. The LASSO algorithm with the same 18 features demonstrated slightly superior predictive performance and an AUC of 0.80. We found significant disparities in SMM among patients living in low-income areas, with public insurance, and who were non-Hispanic Black or non-English speakers. Conclusion: Our results demonstrate the feasibility of utilizing ML and administrative hospital discharge data for SMM prediction. The low recall score is a limitation across all models we compared, signifying that the algorithms struggle with identifying all SMM cases. This study identified substantial disparities in SMM across various sociodemographic factors. Addressing these disparities requires multifaceted interventions that include improving access to quality care, enhancing cultural competence among healthcare providers, and implementing policies that help mitigate social determinants of health.

Keywords: severe maternal morbidity; relative risks; machine learning

check for updates

Academic Editor: Ines Aguinaga-Ontoso

Received: 26 December 2024
Revised: 23 January 2025
Accepted: 27 January 2025
Published: 31 January 2025

Citation: Li, Q.; Alfonso, Y.N.; Wolfson, C.; Aziz, K.B.; Creanga, A.A. Leveraging Machine Learning to Predict and Assess Disparities in Severe Maternal Morbidity in Maryland. *Healthcare* **2025**, *13*, 284. https://doi.org/10.3390/healthcare13030284

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Severe maternal morbidity (SMM), defined by the Centers for Disease Control and Prevention (CDC) as "unexpected outcomes of labor and delivery that result in significant short- or long-term consequences to a woman's health," has been increasing in the United States [1]. SMM carries a range of adverse consequences for women's health, encompassing elevated medical costs, prolonged hospitalization stays, and enduring health effects [2].

Additionally, SMM is often accompanied by adverse neonatal outcomes, including preterm birth, stillbirth, and mortality [3].

The CDC utilizes a list of 21 indicators and corresponding International Classification of Diseases (ICD-10-CM) codes from administrative hospital discharge data to identify delivery hospitalizations with SMM [1]. Following its validation in medical records data, the algorithm is widely used at both national and state levels for program benchmarking [4]. A recent study, drawing on 11.6 million delivery-related hospitalizations in the United States, found that the prevalence of SMM increased from 146.8 to 179.8 per 10,000 discharges during the period 2008–2021 [4]. This SMM burden translates to an annual incidence of over 50,000 SMM cases in the United States. Disparities across racial and ethnic groups are also consistently demonstrated in both national and state data [5].

The progression of pregnancy complications to severe illness and death is largely predictable and preventable [6,7]. Studies suggest that between one-third and two-fifths of maternal adverse events can be averted [8–10]. Machine learning (ML) has emerged as a powerful technique, showcasing promising prediction performance in various medical fields. Compared to traditional statistical approaches, ML algorithms exhibit superiority in handling large data with numerous features and achieving high predictive accuracy. ML algorithms have only recently found applications in maternal health [11,12].

The main objective of this study is to explore the feasibility and efficacy of ML techniques to accurately predict SMM. Our research question is whether the application of ML techniques to hospital discharge records can predict delivery hospitalizations with SMM. Several modeling approaches were explored, and the results from the best-performing model are presented in this report. The secondary objective of this study is to examine disparities in SMM by key sociodemographic characteristics based on the ML approach. Such an exploration holds significant implications. Successful early detection provides opportunities for timely and appropriate intervention, which may prevent the progression of SMM or mitigate its onset and consequences. This underscores the importance of identifying at-risk individuals at the earliest possible stage to improve maternal health and reduce the burden of SMM.

In this report, we first describe the dataset utilized in our study and outline our methodologies, then present and discuss our results, and conclude with an acknowledgment of the limitations of the current research; our conclusion highlights the implications of this study and offers suggestions for future research.

2. Materials and Methods

We used linked data from the Agency for Healthcare Research and Quality (AHRQ)'s State Inpatient Database (SID) and the American Hospital Association (AHA) Annual Survey data from Maryland for 2016–2019. The Maryland SID includes the universe of inpatient discharge records from community hospitals in Maryland. The SID data from Maryland are superior to corresponding databases in most other states in that they include information about intensive care unit admissions, obesity status, and history of substance use and allow longitudinal linking of patients over time via a unique patient identifier. The elements of the AHA survey that we linked with the SID data include hospital demographics, facility and service-line offerings, beds, utilization, finance and staffing characteristics, hospital organization, and leadership. The analytic sample comprised all delivery hospitalizations in Maryland hospitals during this period.

CDC's SMM algorithm was used to identify deliveries during which patients experienced SMM (excluding blood transfusion) [1]. Starting in October 2015, the International Classification of Diseases 10th Revision codes are used to identify 21 markers of organ-system dysfunction during delivery hospitalizations with SMM using administrative

hospital discharge data. The algorithm is applied to delivery hospitalizations in patients aged 12–55 years.

We organized the potential features of SMM in our linked SID-AHA data into four broad domains: delivery hospital characteristics (from AHA), patients' sociodemographic characteristics (from SID), general pregnancy risk factors (from SID), and comorbidities and pregnancy complications (from SID). Altogether, we identified 75 potential risk factors of SMM across these four domains, of which 64 are categorical and 11 are numerical variables.

We first examined the unadjusted relative risk of SMM across key sociodemographic factors (i.e., race–ethnicity, household income, insurance status, and primary language), aiming to provide insights into inequities in this outcome.

For formal modeling, there exists a wide array of options for modeling a binary target variable. Among these, the traditional and widely used approach is logistic regression, which provides a straightforward and interpretable framework for predicting binary targets. Building on this foundation, a natural extension is the LASSO (Least Absolute Shrinkage and Selection Operator) regression. LASSO not only estimates the model parameters but also incorporates feature selection by applying a penalty to the coefficients, effectively shrinking some of them to zero. This dual functionality of LASSO makes it particularly valuable in scenarios where the dataset contains a large number of features, as it helps to identify the most relevant features while simultaneously improving model interpretability and reducing overfitting. These methods, among others, offer robust tools for addressing binary classification problems in various research and practical applications.

In this study, we started with a traditional Logit regression model of SMM on the 75 features available in our dataset. Subsequently, we employed LASSO for SMM prediction. The Logit regression for a binary target variable $Y \in \{0,1\}$ can be expressed as

$$logit(P(Y = 1|\boldsymbol{X})) = \boldsymbol{\beta X} \quad (1)$$

where $\boldsymbol{X}$ denotes the features and $\boldsymbol{\beta}$ denotes the coefficients. Adding a penalty to the log-likelihood gives a LASSO regression as follows:

$$\sum_{i=1}^{n} (y_i \log(p_i) + (1 - y_i)\log(1 - p_i)) - \lambda \sum_{j=1}^{p} |\beta_j| \quad (2)$$

where y_i and p_i denote the observed value and expected probability for the i-th observation, respectively; n denotes the number of records; and λ denotes the hyperparameter, which controls the strength of the penalty on the coefficients. A small λ leads to minimal penalty, allowing more coefficients to be non-zero, i.e., less feature selection and larger variance. A larger λ results in more coefficients being shrunk to zero, which aids in feature selection.

The LASSO approach serves for both variable selection and regularization in predictive analysis. By introducing a penalty term based on the absolute values of model coefficients, LASSO effectively shrinks some coefficients to zero [13]. The best value of λ was tuned using cross-validation to balance the bias–variance trade-off. This ability proves particularly advantageous in scenarios with a substantial number of features, preventing overfitting and enhancing interpretability [14]. The LASSO model outperformed the other modeling approaches we explored for the dataset, including support vector machines (SVMs) and random forests. A 10-fold leave 5% out cross-validation approach was employed to assess predictive performance. In each of the ten iterations, a randomly chosen 5% sample was excluded from model training, and out-of-sample predictions for those samples were compared with the true values.

Many of the 75 features exhibited high correlations with each other, leading to singular matrices and non-convergence in model fitting with both the Logit and LASSO models. To address the convergence issue and explore result sensitivity, the Logit, and LASSO models

with a short list of 18 features were fitted and compared. Variable selection for the reduced 18-feature model was based on literature reviews [15,16], expert opinion, exploratory correlation analysis, statistical significance, and computational resource constraints. Due to the large number of records in the dataset, including too many features in the model, makes the model fitting highly computationally intensive, often exceeding the capacity of even a high-performance desktop computer. The 18 features judged to represent a minimum necessary list of features are maternal age, presence of comorbid conditions (including chronic medical conditions, obesity), substance use during pregnancy, delivery hospital's teaching status, and annual delivery volumes. Some women contributed more than one delivery during the research period, which was accounted for by the woman-level clustering effect.

We computed multiple performance metrics, including true positive (TP), true negative (TN), false positive (FP), false negative (FN), AUC, accuracy, precision, recall score, and F1 score.

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$F1 = 2 * \frac{precision * recall}{precision + recall} \tag{5}$$

The F1 score is used to evaluate the performance of a classification model in scenarios where class imbalance may exist. It is the harmonic mean of precision and recall, providing a single score that balances both metrics. The F1 score ranges from 0 to 1, where 1 indicates perfect precision and recall and 0 indicates poor performance.

We used the ROC (receiver operating characteristics curve) and AUC (area under the ROC curve) to assess and compare the predictive performance of the Logit and LASSO models. An ROC curve plots two parameters (sensitivity; 1—specificity) of a prediction model at all classification thresholds. AUC measures the entire two-dimensional area under the complete ROC curve. Although several other measures exist, many are inappropriate for highly imbalanced datasets where one target value is much more frequent than the other. Both the Logit and LASSO models were estimated in R (version 4.4.0). The model fitting was performed on a desktop computer (Windows 11; CPU i9-12900KF; 64 G RAM). The LASSO model with 18 features took about 10 min to converge, while the Logit model took less than 1 min.

Data are from public de-identified databases. For this reason, this study was exempt from review by the Institutional Review Board at our institution.

3. Results

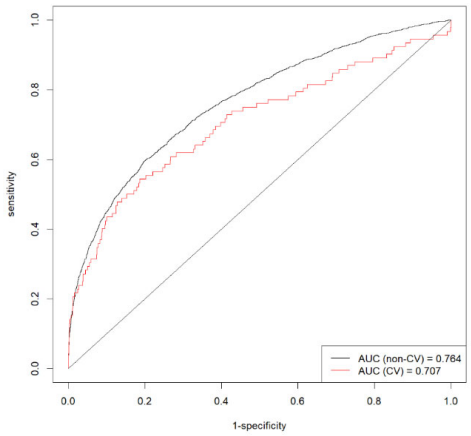
3.1. SMM Rates

Our analytic dataset included 261,226 delivery hospitalizations in Maryland during 2016–2019. The SMM rate was 0.76 per 10,000 deliveries, ranging between 0.55 and 1.12 per 10,000 deliveries among non-Hispanic White and non-Hispanic Black patients.

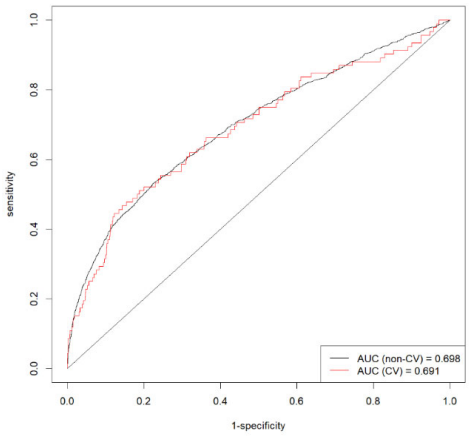
The unadjusted relative risk (RR) of SMM varied markedly across different racial and ethnic groups. Non-Hispanic Black women face a substantially higher risk of experiencing SMM compared to non-Hispanic White women (RR = 2.03). Hispanic and non-Hispanic other groups also experience elevated risks (RR = 1.18 and 1.27, respectively). We also found a linear relationship between SMM and median household income for the residence area, with women in the lowest income quartile having significantly higher risks of SMM than women in the highest quartile. Women with public insurance have the highest relative risk of SMM (RR = 1.36) compared to those with private insurance. Also, women whose primary language is not English face a marginally higher risk (RR = 1.06) compared to English speakers.

3.2. Logit Modeling

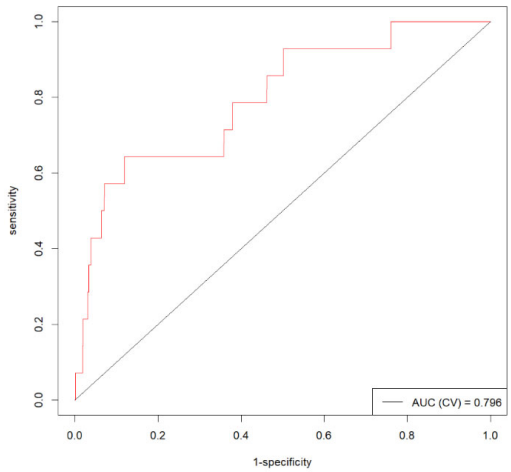
For the 75-feature Logit model, the AUC for non-cross-validation (non-CV) was 0.764, and the AUC for the cross-validation was 0.707 (Figure 1A). The Logit model with 18 features achieved comparable predictive performance, with a non-CV AUC of 0.698 and a CV AUC of 0.691 (Figure 1B). We used the default threshold of 0.5 for the binary classification task when calculating the metrics reported in Table 1. This threshold implies that predictions with a probability of 0.5 or higher were classified as positive, while those below 0.5 were classified as negative. The choice of this default threshold reflects standard practice in binary classification tasks, providing an initial evaluation of model performance.



(**A**) Logit model with 75 features



(**B**) Logit model with 18 features



(**C**) LASSO model with 18 features

Figure 1. Severe maternal morbidity predicting performance for the Logit models with 75 features and the Logit and LASSO models with 18 features. Notes: Reduction from 75 to 18 features based on literature reviews, exploratory correlation analysis, statistical significance, and computational resource constraints. The 18 features include year, maternal age, primary language, insurance status, homeless status, median household income quartile for the patient zip code, level of maternal care, teaching status of the hospital, hospital contracts with payors being tied to performance on quality/safety measures, hospital has patient/family advisory, hospital encounter in the past 30 days before delivery hospitalization, obesity, multiple gestations, supervision of high-risk pregnancy, hypertensive disease, comorbidities, annual delivery volume for the hospital, and % of deliveries to minority (non-White non-Hispanic) women in the hospital.

Table 1. Performance of the Logit and LASSO models in predicting severe maternal morbidity using hospital discharge data.

Algorithm	TP	FP	TN	FN	AUC	Accuracy	Precision	Recall
Logit-75	50	42	10559	2419	0.71	0.81	0.54	0.02
Logit-18	44	48	10979	1999	0.69	0.84	0.48	0.02
LASSO-18	45	25	11454	1546	0.80	0.88	0.64	0.03

Notes: TP = true positive; FP = false positive; TN = true negative; FN = false negative; AUC = area under the receiver operating characteristic curve. Accuracy refers to the proportion of correct predictions out of the total number of predictions. Precision refers to the proportion of correctly predicted positive cases out of all cases predicted as positive. Recall refers to the proportion of actual positive cases that were correctly predicted by the model.

3.3. Comparison Between Predicting Performance with LASSO vs. Logit Modeling

The LASSO model with the full list of 75 features did not converge, while the model with 18 features did. Non-CV essentially represents in-sample validation, where the whole dataset is used in both model training and validation. Including the prediction dataset in training usually increases the AUC, making non-CV a reference for model performance.

Compared with the Logit models, LASSO achieved a higher CV AUC of 0.796 (Figure 1C). Table 1 presents model performance metrics for the validation dataset of 13,070 deliveries, representing a 5% random sample of the total 261,226 deliveries in the full dataset. Overall, all models exhibited good predictive performance. LASSO's precision score of 0.64 signifies that 64% of instances predicted as positive are indeed true SMM. A precision score of this magnitude is indicative of an acceptable false positive rate. Across all performance metrics, LASSO consistently outperforms the Logit models.

Recall values remain low in all models, suggesting low sensitivity and a limited ability to identify all positive cases. This raises concerns and underscores the need for further research. The persistently low recall values are a common target value in machine learning classification tasks with highly skewed data [17,18]. The high precision and low recall values imply that the optimal application of the prediction algorithm is for initial screening rather than final confirmation.

3.4. Comparison with Previous Models

Our model demonstrates performance that is comparable to, or in some cases superior to, existing models documented in the literature (Table 2). However, model features included and metrics reported vary significantly across different works, making a direct and comprehensive comparison challenging.

Table 2. A Summary of the predictive performance of machine learning models for SMM.

Article	Algorithm	Precision	Recall	AUC or F1
Gao 2019 [19]	Regularized Logit	0.22 to 0.35	Sensitivity: 0.614 to 0.765	AUC: 0.790 to 0.937
Rodríguez 2016 [20]	Logit	NA	NA	AUC 0.66
Xu 2023 [21]	EBM; Logit	NA	0.59	AUC EBM 0.70 AUC Logit 0.69
Lengerich 2024 [22]	GAM	0.0152	0.369	AUC: 0.67
Rodríguez 2020 [23]	Logit; SVM	Logit: 0.518; SVM: 0.279	Logit: 0.977: SVM: 1	F1: Logit 0.677 vs SVM: 0.436
Clapp 2022a [24]	LASSO with BoW	NA	NA	AUC for SMM: 0.67–0.72; AUC for NT SMM: 0.72–0.76
Clapp 2022b [25]	NLP with BoW; LASSO with BoW	0.194 (SMM); 0.084 (NT SMM)	0.287 (SMM); 0.298 (NT SMM)	AUC for SMM: 0.76; AUC for NT SMM: 0.75
Clapp 2021 [26]	LASSO; Elastic Net, Ridge	0.075	0.11	AUC 0.611
Leonard 2022 [27]	Logit	NA	NA	AUC: 0.68–0.76

Notes: BoW, bag of words; EBM, explainable boosting machine; GAM, Generalized Additive Model; NLP, natural language processing; NPV, negative predictive value; NT SMM, non-transfusion SMM; SPR, screen positive rate.

4. Discussion

Our findings indicate that both the Logit and LASSO models can effectively predict SMM with acceptable accuracy, even with a relatively short list of features. The AUC values are comparable to similar applications in maternal health [28–30]. The LASSO model outperformed the Logit model in terms of predictive metrics, although the margin of outperformance is small. This may be attributed to the complex nature of SMM [31–33] and highly imbalanced data, with SMM prevalence <1%. ML models, by nature, learn much more about non-SMM cases than SMM cases. However, methodologically, this study employs an ML technique that is unconstrained by the preset assumptions of statistical distributions regarding variables and parameters. To our knowledge, this is one of few attempts to predict SMM using ML techniques [19] and the only study to employ LASSO. Hence, our results can serve as a comparison for future ML applications to predict SMM and other adverse maternal health outcomes.

Our scoping review of the literature shows how the prediction of maternal health outcomes varies and changes over time (Figure 2). Studies in the 1970s until the 2000s employed correlation and regression analyses to examine associations (or lack thereof) between specific clinical factors and adverse maternal outcomes [34–36]. The 2000s and 2010s brought an interest in early prediction of adverse outcomes in patients with or without documented pregnancy complications using clinical and laboratory tests [37–39]. The use of ML and NLP techniques is relatively new in maternal health, likely due to limitations in the data available and the fact that adverse outcomes are statistically rare events [24,40–43]. Currently, the focus is on identifying the best-performing ML techniques for predicting outcomes, such as SMM, and this is where our study makes a contribution [44,45].

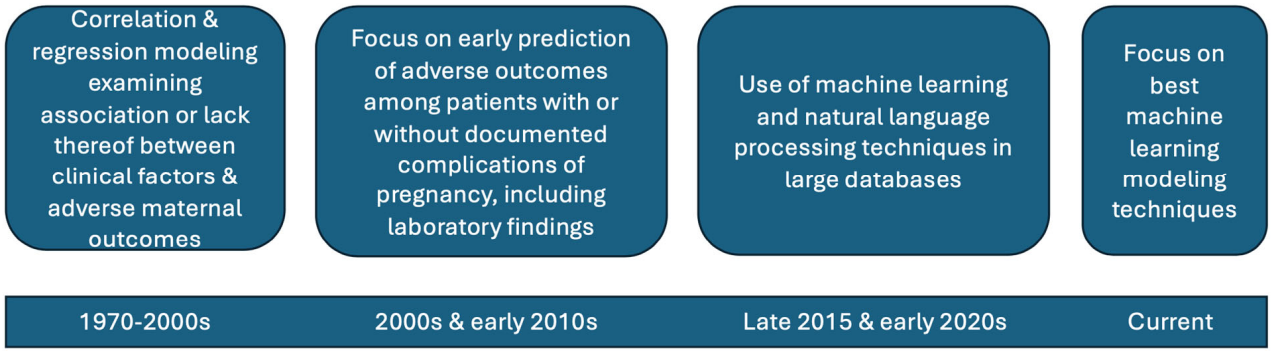


Figure 2. Efforts to predict maternal morbidity and severe outcomes over time.

Some studies aim to predict maternal outcomes to improve clinical care by evaluating new or improved screening guidelines or tools (e.g., the inclusion of proteinuria) [37–39,46–49]. Others assess the effect of potential features available through advances in technology or data access (e.g., placental imaging techniques, genetic data, novel biomarkers such as placental growth factor, platelet parameters, plasma cell-free RNA) [34–36,43,50–57]. And, only in the past decade or so have studies started to compare the predictive power of different statistical classification models (i.e., ML models, logistic regression, naive Bayes, adaptive models) [24,40,44,58–60].

Studies can also be distinguished by the timing of screening. Among those predicting preeclampsia risk in pregnant women using first-trimester data, some have used clinical and laboratory data to compare the predictive power of different classification models. These studies show that models can perform well, with AUCs as high as 0.89 (using the elastic net algorithm) or 0.86 (using random forest) [41,60]. Similar studies examining biomarkers, specifically systolic blood pressure polygenic risk scores, suggest this feature does not improve preeclampsia predictions [45]. Other studies combining multiple

biomarkers (e.g., placental growth factor, clinical, and biophysical features) demonstrate that this combination improves predictions compared to PLGF alone (AUC 0.937 with logistic regression) [39]. Additional studies using first-trimester data provide more insights into biomarker quality assessment (e.g., cumulative sum and target plots to assess blood pressure or placental growth factor) [48]. Studies assessing features obtained during the second trimester also looked at blood pressure to predict the risk of hypertension or preeclampsia. These studies include some of the earliest prediction models for pregnant women's risks and found evidence for predicting hypertension but not preeclampsia [34,35]. Other studies assess specific biomarkers, irrespective of the trimester of data collection, including platelet parameters or plasma cell-free RNA data [42,57]. More recent studies evaluating features obtained during the second or third trimesters focused on the risk of other pregnancy outcomes, postpartum hemorrhage, SMM, or fetal outcomes [24,46,49,55,61]. A recent study examining the risk of SMM assessed the predictive power of natural language processing (NLP) models using history and physical note texts (excluding additional patient information, such as demographics and diagnosis codes) and obtained an AUC of 0.76 [24].

The remaining studies assessed the risk of adverse pregnancy outcomes among women diagnosed with preeclampsia. The most recent study comparing different classification models found that among the nine models tested, elastic net, stochastic gradient boosting, extreme gradient boosting, and random forest produced the strongest discrimination between true and false outcomes, with AUCs between 0.860 and 0.973 [44]. Other studies evaluated predictions with specific biomarkers (e.g., liver function tests, creatinine concentrations from lab data, placental growth factors) [36,43,50]. Among these, the strongest performing model achieved an AUC of 0.88 using random forest [43]. Additional studies used logistic regression to evaluate the effect of routine biomarkers and clinical characteristics data on prediction performance, and these models achieved AUCs ranging between 0.71 (with external validation) and 0.88 or 0.91 with and without cross-validation, respectively [37,38,62].

Our study also reveals significant inequities in SMM across key sociodemographic factors. Women from low-income households, non-Hispanic Black women, those with public insurance, and non-English speakers are particularly vulnerable. Addressing these disparities requires multifaceted interventions that include improving access to quality care, enhancing cultural competence among healthcare providers, and implementing policies that mitigate the social determinants of health. By focusing on these areas, we can work towards a more equitable maternal healthcare system that ensures all women have the opportunity for safe and healthy pregnancies.

This study has several limitations. First, the ML algorithms are trained and validated using only data from Maryland. They may need further fine tuning using more diverse datasets to enhance generalization performance. Such external validation on an entirely separate dataset from a different source or population could provide further insights. However, such datasets were not available to us. Secondly, we explored several other ML models that we identified through our literature search, including decision trees, random forests, K-Nearest Neighbors (KNNs), and support vector machines (SVMs). Although these models have been extensively applied in classification tasks, they either did not converge or performed poorly in classifying SMM in our data. This is likely due to the highly imbalanced distribution of the SMM outcomes in our data. Thirdly, there are more advanced ML algorithms, such as deep neural networks, which offer greater modeling flexibility and predictive power. However, these sophisticated methods come with their own set of limitations, including increased complexity, higher computational costs, and challenges in implementation and interpretation. As a result, they fall outside the scope of

the present study, which focuses on more accessible and interpretable approaches. Those advanced methods can be explored in future research.

5. Conclusions

Our study highlights the potential of ML algorithms in predicting SMM in large databases. More research is warranted to validate and refine these algorithms and achieve more accurate SMM detection. Potential directions for future research include leveraging advanced methods, such as the Synthetic Minority Oversampling Technique, to address the highly imbalanced nature of the data, or employing sophisticated deep neural networks to enhance predictive performance. Our study also confirms substantial equity issues in maternal health across various sociodemographic aspects in Maryland. By focusing on these areas, we can work towards a more equitable maternal healthcare system that ensures all women have the opportunity for safe and healthy pregnancies.

Author Contributions: Q.L. had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: Q.L., A.A.C. Acquisition, analysis, or interpretation of data: All authors. Drafting of the manuscript: Q.L., Y.N.A. Critical revision of the manuscript for important intellectual content: All authors. Statistical analysis: Q.L. Obtained funding: A.A.C. Administrative, technical, or material support: A.A.C. Supervision: A.A.C. All authors have read and agreed to the published version of the manuscript.

Funding: Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Numbers R01HD112442 and R03HD095057. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH.

Institutional Review Board Statement: Data are from public de-identified databases, for this reason, the study was exempt from review by the Institutional Review Board at our institution.

Data Availability Statement: The original data presented in the study are openly available through AHRQ's HCUP project.

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

References

1. CDC. Severe Maternal Morbidity. Available online: https://www.cdc.gov/maternal-infant-health/php/severe-maternal-morbidity/index.html (accessed on 13 January 2024).
2. Callaghan, W.M.; Creanga, A.A.; Kuklina, E.V. Severe maternal morbidity among delivery and postpartum hospitaliza-tions in the United States. *Obstet. Gynecol.* **2012**, *120*, 1029–1036. [CrossRef]
3. Zeitlin, J.; Egorova, N.N.; Janevic, T.; Hebert, P.L.; Lebreton, E.; Balbierz, A.; Howell, E.A. The Impact of Severe Maternal Morbidity on Very Preterm Infant Outcomes. *J. Pediatr.* **2019**, *215*, 56–63.e1. [CrossRef]
4. Fink, D.A.; Kilday, D.; Cao, Z.; Larson, K.; Smith, A.; Lipkin, C.; Perigard, R.; Marshall, R.; Deirmenjian, T.; Finke, A.; 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*, e2317641. [CrossRef]
5. Wolfson, C.; Qian, J.; Chin, P.; Downey, C.; Mattingly, K.J.; Jones-Beatty, K.; Olaku, J.; Qureshi, S.; Rhule, J.; Silldorff, D.; et al. Findings From Severe Maternal Morbidity Surveillance and Review in Maryland. *JAMA Netw. Open* **2022**, *5*, e2244077. [CrossRef] [PubMed] [PubMed Central]
6. Kilpatrick, S.J.; Crabtree, K.E.; Kemp, A.; Geller, S. Preventability of maternal deaths: Comparison between Zambian and American referral hospitals. *Obstet. Gynecol.* **2002**, *100*, 321–326. [CrossRef]
7. Nannini, A.; Weiss, J.; Goldstein, R.; Fogerty, S. Pregnancy-associated mortality at the end of the twentieth century: Massachusetts, 1990–1999. *J. Am. Med. Womens Assoc.* **2002**, *57*, 140–143.
8. Berg, C.J.; Harper, M.A.; Atkinson, S.M.; Bell, E.A.; Brown, H.L.; Hage, M.L.; Mitra, A.G.; Moise, K.J., Jr.; Callaghan, W.M. Preventability of pregnancy-related deaths: Results of a state-wide review. *Obstet. Gynecol.* **2005**, *106*, 1228–1234. [CrossRef]
9. Geller, S.E.; Cox, S.M.; Kilpatrick, S.J. A descriptive model of preventability in maternal morbidity and mortality. *J. Perinatol.* **2006**, *26*, 79–84. [CrossRef] [PubMed]

10. Geller, S.E.; Koch, A.R.; Martin, N.J.; Rosenberg, D.; Bigger, H.R. Assessing preventability of maternal mortality in Illinois: 2002-2012. *Am. J. Obstet. Gynecol.* **2014**, *211*, 698.e1–698.e11. [CrossRef]
11. Ranjbar, A.; Montazeri, F.; Farashah, M.V.; Mehrnoush, V.; Darsareh, F.; Roozbeh, N. Machine learning-based approach for predicting low birth weight. *BMC Pregnancy Childbirth* **2023**, *23*, 803. [CrossRef] [PubMed]
12. Lodi, M.; Poterie, A.; Exarchakis, G.; Brien, C.; de Micheaux, P.L.; Deruelle, P.; Gallix, B. Prediction of cesarean delivery in class III obese nulliparous women: An externally validated model using machine learning. *J. Gynecol. Obstet. Hum. Reprod.* **2023**, *52*, 102624. [CrossRef] [PubMed]
13. Tibshirani, R. Regression Shrinkage and Selection Via the Lasso. *J. R. Stat. Soc. Ser. B (Methodol.)* **2018**, *58*, 267–288. [CrossRef]
14. Ranstam, J.; Cook, J.A. LASSO regression. *Br. J. Surg.* **2018**, *105*, 1348. [CrossRef]
15. Geller, S.E.; Rosenberg, D.; Cox, S.M.; Brown, M.L.; Simonson, L.; Driscoll, C.A.; Kilpatrick, S.J. The continuum of maternal morbidity and mortality: Factors associated with severity. *Am. J. Obstet. Gynecol.* **2004**, *191*, 939–944. [CrossRef]
16. Clark, S.L.; Belfort, M.A.; Dildy, G.A.; Herbst, M.A.; Meyers, J.A.; Hankins, G.D. Maternal death in the 21st century: Causes, prevention, and relationship to cesarean delivery. *Am. J. Obstet. Gynecol.* **2008**, *199*, 36.e1-5; discussion 91-2. e7-11. [CrossRef] [PubMed]
17. Schaudt, D.; von Schwerin, R.; Hafner, A.; Riedel, P.; Reichert, M.; von Schwerin, M.; Beer, M.; Kloth, C. Augmentation strategies for an imbalanced learning problem on a novel COVID-19 severity dataset. *Sci. Rep.* **2023**, *13*, 18299. [CrossRef]
18. Ahmed, J.; Ii, R.C.G. Predicting severely imbalanced data disk drive failures with machine learning models. *Mach. Learn. Appl.* **2022**, *9*, 100361. [CrossRef]
19. Gao, C.; Osmundson, S.; Yan, X.; Edwards, D.V.; Malin, B.A.; Chen, Y. Learning to identify severe maternal morbidity from electronic health records. *Stud. Health Technol. Inform.* **2019**, *264*, 143–147. [PubMed]
20. Rodríguez, E.A.; Estrada, F.E.; Torres, W.C.; Santos, J.C.M. Early Prediction of Severe Maternal Morbidity Using Machine Learning Techniques. In *Ibero-American Conference on Artificial Intelligence*; Springer: Cham, Switzerland, 2016; pp. 259–270. [CrossRef]
21. Xu, Z.; Bosschieter, T.M.; Lan, H.; Lengerich, B.; Nori, H.; Sitcov, K.; Painter, I.; Souter, V.; Caruana, R. Predicting severe maternal morbidity at admission for delivery using intelligible machine learning. *Am. J. Obstet. Gynecol.* **2023**, *228*, S404–S405. [CrossRef]
22. Lengerich, B.J.; Caruana, R.; Painter, I.; Weeks, W.B.; Sitcov, K.; Souter, V. Interpretable machine learning predicts postpartum hemorrhage with severe maternal morbidity in a lower-risk laboring obstetric population. *Am. J. Obstet. Gynecol. MFM* **2024**, *6*, 101391. [CrossRef] [PubMed]
23. Arrieta Rodríguez, E.; López-Martínez, F.; Santos, J.C.M. *A Machine Learning Approach for Severe Maternal Morbidity Prediction at Rafael Calvo Clinic in Cartagena-Colombia*; Springer International Publishing: Cham, Switzerland, 2020.
24. Clapp, M.A.; Kim, E.; James, K.E.; Perlis, R.H.; Kaimal, A.J.; McCoy, T.H. Natural language processing of admission notes to predict severe maternal morbidity during the delivery encounter. *Am. J. Obstet. Gynecol.* **2022**, *227*, 511.e1–511.e8. [CrossRef] [PubMed]
25. Clapp, M.A.; Kim, E.; James, K.E.; Perlis, R.H.; Kaimal, A.J.; McCoy, T.H.; Easter, S.R. Comparison of Natural Language Processing of Clinical Notes With a Validated Risk-Stratification Tool to Predict Severe Maternal Morbidity. *JAMA Netw. Open* **2022**, *5*, e2234924. [CrossRef] [PubMed]
26. Clapp, M.A.; McCoy, T.H., Jr.; James, K.E.; Kaimal, A.J.; Perlis, R.H. Derivation and external validation of risk stratification models for severe maternal morbidity using pre-natal encounter diagnosis codes. *J Perinatol.* **2021**, *41*, 2590–2596. [CrossRef]
27. Leonard, S.A.; Main, E.K.; Lyell, D.J.; Carmichael, S.L.; Kennedy, C.J.; Johnson, C.; Mujahid, M.S. Obstetric comorbidity scores and disparities in severe maternal morbidity across marginalized groups. *Am. J. Obstet. Gynecol. MFM* **2022**, *4*, 100530. [CrossRef] [PubMed]
28. Betts, K.S.; Kisely, S.; Alati, R. Predicting common maternal postpartum complications: Leveraging health administrative data and machine learning. *BJOG Int. J. Obstet. Gynaecol.* **2019**, *126*, 702–709. [CrossRef]
29. Wang, S.; Pathak, J.; Zhang, Y. Using Electronic Health Records and Machine Learning to Predict Postpartum Depression. *Stud. Health Technol. Inform.* **2019**, *264*, 888–892. [PubMed]
30. Sun, Q.; Zou, X.; Yan, Y.; Zhang, H.; Wang, S.; Gao, Y.; Liu, H.; Liu, S.; Lu, J.; Yang, Y.; et al. Machine Learning-Based Prediction Model of Preterm Birth Using Electronic Health Record. *J. Health Eng.* **2022**, *2022*, 9635526. [CrossRef] [PubMed]
31. Gray, K.E.; Wallace, E.R.; Nelson, K.R.; Reed, S.D.; Schiff, M.A. Population-Based Study of Risk Factors for Severe Maternal Morbidity. *Paediatr. Périnat. Epidemiol.* **2012**, *26*, 506–514. [CrossRef]
32. Madeiro, A.P.; Rufino, A.C.; Lacerda, É.Z.G.; Brasil, L.G. Incidence and determinants of severe maternal morbidity: A transversal study in a referral hospital in Teresina, Piaui, Brazil. *BMC Pregnancy Childbirth* **2015**, *15*, 210. [CrossRef] [PubMed]
33. Himes, K.P.; Bodnar, L.M. Validation of criteria to identify severe maternal morbidity. *Paediatr. Périnat. Epidemiol.* **2020**, *34*, 408–415. [CrossRef]
34. Friedman, E.A.; Neff, R.K. Hypertension-Hypotension in Pregnancy: Correlation with Fetal Outcome. *JAMA J. Am. Med. Assoc.* **1978**, *239*, 2249–2251. [CrossRef]

35. Chesley, L.C.; Sibai, B.M. Clinical significance of elevated mean arterial pressure in the second trimester. *Am. J. Obstet. Gynecol.* **1988**, *159*, 275–279. [CrossRef]
36. Thangaratinam, S.; Koopmans, C.M.; Iyengar, S.; Zamora, J.; Ismail, K.M.; Mol, B.W.; Khan, K.S. Accuracy of liver function tests for predicting adverse maternal and fetal outcomes in women with preeclampsia: A systematic review. *Acta Obstet. Gynecol. Scand.* **2011**, *90*, 574–585. [CrossRef] [PubMed]
37. Von Dadelszen, P.; Payne, B.; Li, J.; Ansermino, J.M.; Pipkin, F.B.; Côté, A.-M.; Douglas, M.J.; Gruslin, A.; A Hutcheon, J.; Joseph, K.; et al. Prediction of adverse maternal outcomes in preeclampsia: Development and validation of the fullPIERS model. *Lancet* **2011**, *377*, 219–227. [CrossRef] [PubMed]
38. Payne, B.A.; Hutcheon, J.A.; Ansermino, J.M.; Hall, D.R.; Bhutta, Z.A.; Bhutta, S.Z.; Biryabarema, C.; Grobman, W.A.; Groen, H.; Haniff, F.; et al. A Risk Prediction Model for the Assessment and Triage of Women with Hypertensive Disorders of Pregnancy in Low-Resourced Settings: The miniPIERS (Preeclampsia Integrated Estimate of RiSk) Multi-country Prospective Cohort Study. *PLoS Med.* **2014**, *11*, e1001589. [CrossRef]
39. Agarwal, R.; Chaudhary, S.; Kar, R.; Radhakrishnan, G.; Tandon, A. Prediction of preeclampsia in primigravida in late first trimester using serum placental growth factor alone and by combination model. *J. Obstet. Gynaecol.* **2017**, *37*, 877–882. [CrossRef] [PubMed]
40. Marić, I.; Tsur, A.; Aghaeepour, N.; Montanari, A.; Stevenson, D.K.; Shaw, G.M.; Winn, V.D. Early Prediction of Preeclampsia via Machine Learning. *Am. J. Obstet. Gynecol. MFM* **2020**, *2*, 100100. [CrossRef]
41. Liu, M.; Yang, X.; Chen, G.; Ding, Y.; Shi, M.; Sun, L.; Huang, Z.; Liu, J.; Liu, T.; Yan, R.; et al. Development of a prediction model on preeclampsia using machine learning-based method: A retrospective cohort study in China. *Front. Physiol.* **2022**, *13*, 896969. [CrossRef]
42. Rasmussen, M.; Reddy, M.; Nolan, R.; Camunas-Soler, J.; Khodursky, A.; Scheller, N.M.; Cantonwine, D.E.; Engelbrechtsen, L.; Mi, J.D.; Dutta, A.; et al. RNA profiles reveal signatures of future health and disease in pregnancy. *Nature* **2022**, *601*, 422–427. [CrossRef]
43. Schmidt, L.J.; Rieger, O.; Neznansky, M.; Hackelöer, M.; Dröge, L.A.; Henrich, W.; Higgins, D.; Verlohren, S. A machine-learning–based algorithm improves prediction of preeclampsia-associated adverse outcomes. *Am. J. Obstet. Gynecol.* **2022**, *227*, 77.e1–77.e30. [CrossRef] [PubMed]
44. Ranjbar, A.; Montazeri, F.; Ghamsari, S.R.; Mehrnoush, V.; Roozbeh, N.; Darsareh, F. Machine learning models for predicting preeclampsia: A systematic review. *BMC Pregnancy Childbirth* **2024**, *24*, 6. [CrossRef] [PubMed]
45. Kovacheva, V.P.; Eberhard, B.W.; Cohen, R.Y.; Maher, M.; Saxena, R.; Gray, K.J. Preeclampsia Prediction Using Machine Learning and Polygenic Risk Scores from Clinical and Genetic Risk Factors in Early and Late Pregnancies. *Hypertension* **2024**, *81*, 264–272. [CrossRef]
46. Shinar, S.; Melamed, N.; Abdulaziz, K.E.; Ray, J.G.; Riddell, C.; Barrett, J.; Murray-Davis, B.; Mawjee, K.; Mcdonald, S.D.; Geary, M.; et al. Changes in rate of preterm birth and adverse pregnancy outcomes attributed to preeclampsia after introduction of a refined definition of preeclampsia: A population-based study. *Acta Obstet. Gynecol. Scand.* **2021**, *100*, 1627–1635. [CrossRef]
47. Hu, M.; Shi, J.; Lu, W. Association between proteinuria and adverse pregnancy outcomes: A retrospective cohort study. *J. Obstet. Gynaecol.* **2022**, *43*, 2126299. [CrossRef] [PubMed]
48. Chaemsaithong, P.; Sahota, D.S.; Poon, L.C. First trimester preeclampsia screening and prediction. *Am. J. Obstet. Gynecol.* **2022**, *226*, S1071–S1097.e2. [CrossRef] [PubMed]
49. Zhang, J.; Klebanoff, M.A.; Roberts, J.M. Prediction of adverse outcomes by common definitions of hypertension in pregnancy. *Obstet. Gynecol.* **2001**, *97*, 261–267.
50. Conti-Ramsden, F.I.; Nathan, H.L.; De Greeff, A.; Hall, D.R.; Seed, P.T.; Chappell, L.C.; Shennan, A.H.; Bramham, K. Pregnancy-related acute kidney injury in preeclampsia: Risk factors and renal outcomes. *Hypertension* **2019**, *74*, 1144–1151. [CrossRef]
51. Irene, K.; Amubuomombe, P.P.; Mogeni, R.; Andrew, C.; Mwangi, A.; Omenge, O.E. Maternal and perinatal outcomes in women with eclampsia by mode of delivery at Riley mother baby hospital: A longitudinal case-series study. *BMC Pregnancy Childbirth* **2021**, *21*, 439. [CrossRef]
52. Kelly, B.S.; Judge, C.; Bollard, S.M.; Clifford, S.M.; Healy, G.M.; Aziz, A.; Mathur, P.; Islam, S.; Yeom, K.W.; Lawlor, A.; et al. Radiology artificial intelligence: A systematic review and evaluation of methods (RAISE). *Eur. Radiol.* **2022**, *32*, 7998–8007. [CrossRef] [PubMed]
53. Giorgione, V.; Quintero Mendez, O.; Pinas, A.; Ansley, W.; Thilaganathan, B. Routine first-trimester preeclampsia screening and risk of preterm birth. *Ultrasound Obstet. Gynecol.* **2022**, *60*, 185–191. [CrossRef]
54. Hackelöer, M.; Schmidt, L.; Verlohren, S. New advances in prediction and surveillance of preeclampsia: Role of machine learning approaches and remote monitoring. *Arch. Gynecol. Obstet.* **2023**, *308*, 1663–1677. [CrossRef]
55. Sun, Z.; Wu, W.; Zhao, P.; Wang, Q.; Woodard, P.K.; Nelson, D.M.; Odibo, A.; Cahill, A.; Wang, Y. Association of intraplacental oxygenation patterns on dual-contrast MRI with placental abnormality and fetal brain oxygenation. *Ultrasound Obstet. Gynecol.* **2023**, *61*, 215–223. [CrossRef] [PubMed]

56. Kenny, L.C.; English, F.; McCarthy, F.P. Risk factors and effective management of preeclampsia. *Integr. Blood Press. Control.* **2015**, *8*, 7–12. [CrossRef]
57. Bawore, S.G.; Adissu, W.; Niguse, B.; Larebo, Y.M.; Ermolo, N.A.; Gedefaw, L. A pattern of platelet indices as a potential marker for prediction of pre-eclampsia among pregnant women attending a Tertiary Hospital, Ethiopia: A case-control study. *PLoS ONE* **2021**, *16*, e0259543. [CrossRef] [PubMed]
58. Aljameel, S.S.; Alzahrani, M.; Almusharraf, R.; Altukhais, M.; Alshaia, S.; Sahlouli, H.; Aslam, N.; Khan, I.U.; Alabbad, D.A.; Alsumayt, A. Prediction of Preeclampsia Using Machine Learning and Deep Learning Models: A Review. *Big Data Cogn. Comput.* **2023**, *7*, 32. [CrossRef]
59. Hennessy, A.; Tran, T.H.; Sasikumar, S.N.; Al-Falahi, Z. Machine learning, advanced data analysis, and a role in pregnancy care? How can we help improve preeclampsia outcomes? *Pregnancy Hypertens.* **2024**, *37*, 101137. [CrossRef] [PubMed]
60. Vasilache, I.A.; Scripcariu, I.S.; Doroftei, B.; Bernad, R.L.; Cărăuleanu, A.; Socolov, D.; Melinte-Popescu, A.-S.; Vicoveanu, P.; Harabor, V.; Mihalceanu, E.; et al. Prediction of Intrauterine Growth Restriction and Preeclampsia Using Machine Learning-Based Algorithms: A Prospective Study. *Diagnostics* **2024**, *14*, 453. [CrossRef] [PubMed]
61. Ende, H.B.; Domenico, H.J.; Polic, A.; Wesoloski, A.; Zuckerwise, L.C.; Mccoy, A.B.; Woytash, A.R.D.; Moore, R.P.; Byrne, D.W. Development and Validation of an Automated, Real-Time Predictive Model for Postpartum Hemorrhage. *Obstet. Gynecol.* **2024**, *144*, 109–117. [CrossRef] [PubMed]
62. Edvinsson, C.; Björnsson, O.; Erlandsson, L.; Hansson, S.R. Predicting intensive care need in women with preeclampsia using machine learning—A pilot study. *Hypertens. Pregnancy* **2024**, *43*, 2312165. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.