Original Research

Machine Learning and Statistical Models to Predict Postpartum Hemorrhage

Kartik K. Venkatesh, MD, PhD, Robert A. Strauss, MD, Chad A. Grotegut, MD, MBA, R. Philip Heine, MD, Nancy C. Chescheir, MD, Jeffrey S. A. Stringer, MD, David M. Stamilio, MD, MSCE, Katherine M. Menard, MD, MPH, and J. Eric Jelovsek, MD, MMEd

OBJECTIVE: To predict a woman's risk of postpartum hemorrhage at labor admission using machine learning and statistical models.

METHODS: Predictive models were constructed and compared using data from 10 of 12 sites in the U.S. Consortium for Safe Labor Study (2002–2008) that consistently reported estimated blood loss at delivery. The outcome was postpartum hemorrhage, defined as an estimated blood loss at least 1,000 mL. Fifty-five candidate risk factors routinely available on labor admission were considered. We used logistic regression with and without lasso regularization (lasso regression) as the two statistical models, and random forest and extreme gradient boosting as the two machine learning models to predict postpartum hemorrhage. Model performance was measured by C statistics (ie, concordance index), calibration, and decision curves. Models were constructed from the first phase (2002–2006) and externally validated (ie, temporally) in the second phase (2007–2008). Further validation was performed combining both temporal and site-specific validation.

RESULTS: Of the 152,279 assessed births, 7,279 (4.8%, 95% CI 4.7–4.9) had postpartum hemorrhage. All models had good-to-excellent discrimination. The extreme gradient boosting model had the best discriminative ability to predict postpartum hemorrhage (C statistic: 0.93; 95% CI 0.92–0.93), followed by random forest (C statistic: 0.92; 95% CI 0.91–0.92). The lasso regression model (C statistic: 0.87; 95% CI 0.86–0.88) and logistic regression (C statistic: 0.87; 95% CI 0.86–0.87) had lower-but-good discriminative ability. The above results held with validation across both time and sites. Decision curve analysis demonstrated that, although all models provided superior net benefit when clinical decision thresholds were between 0% and 80% predicted risk, the extreme gradient boosting model provided the greatest net benefit.

CONCLUSION: Postpartum hemorrhage on labor admission can be predicted with excellent discriminative ability using machine learning and statistical models. Further clinical application is needed, which may assist health care providers to be prepared and triage at-risk women.

(Obstet Gynecol 2020;135:935–44)

DOI: 10.1097/AOG.0000000000003759

Postpartum hemorrhage is the primary source of maternal morbidity and mortality worldwide, accounting for nearly a third of deaths of pregnant and postpartum women.[1] In the United States, the rate of postpartum hemorrhage has increased by at

From the Departments of Obstetrics and Gynecology, University of North Carolina at Chapel Hill, Chapel Hill, Duke University, Durham, Wake Forest University, Winston-Salem, North Carolina.

Supported by the Intramural Research Program of the National Institutes of Health, Eunice Kennedy Shriver *National Institute of Child Health and Human Development (NICHD). The Consortium on Safe Labor was funded by the Intramural Research Program of the NICHD, through Contract No. HHSN267200603425C. The study was funded by the Intramural Research Program of the* Eunice Kennedy Shriver *National Institute of Child Health and Human Development. Intramural investigators designed the study, and data were collected by clinical site investigators. The corresponding author has full access to the data and final responsibility for preparation and submission of the paper for publication.*

Presented at the Society for Maternal-Fetal Medicine's 40th Annual Pregnancy Meeting, February 3–8, 2020, Grapevine, Texas.

For a list of institutions involved in the Consortium on Safe Labor, see Appendix 1 online at http://links.lww.com/AOG/B788.

Dr. Chescheir, Editor-in-Chief of Obstetrics & Gynecology, *was not involved in the review or decision to publish this article.*

Each author has confirmed compliance with the journal's requirements for authorship.

Corresponding author: Kartik K. Venkatesh, MD, PhD, Department of Obstetrics and Gynecology, University of North Carolina at Chapel Hill, Chapel Hill, NC; email: kartik.k.venkatesh@gmail.com.

Financial Disclosure

David M. Stamilio reports receiving funds for being a symposium lecturer at the VI Congreso Javeriano de Ginecologia y Obstetricia Sep 12–14, 2019, and he was compensated for travel expenses by Pontifica Universidad Javeriana, Bogota, Colombia. He was also a co-investigator on a Bill & Melinda Gates Foundation grant (OPP1191684; PI Jeff Stringer) from July 12, 2018–April 30, 2019. The other authors did not report any potential conflicts of interest.

ISSN: 0029-7844/20

least 26% in the past decade.[2,3] Although maternal death is a rare outcome in the United States, blood transfusion after hemorrhage, which is 50 times more common than death, is the main diagnosis associated with severe maternal morbidity in the US.[4,5] Although the methods to estimate blood loss at delivery continue to evolve,[6] the American College of Obstetricians and Gynecologists (ACOG) currently defines postpartum hemorrhage as a cumulative blood loss of at least 1,000 mL or signs or symptoms of hypovolemia within 24 hours after birth.[7]

Predicting a woman's risk of postpartum hemorrhage on labor admission requires the obstetrician to incorporate known risk factors and then to approximate the probability of hemorrhage by using a risk strata scheme.[7,8] With an increasing focus on standardized guidelines to prevent and manage postpartum hemorrhage,[9,10] limited tools exist to accurately predict which women are at the highest risk for hemorrhage.[7] Current risk-based stratification guidelines adopted by ACOG and the California Maternal Quality Care Collaborative (CMQCC) include decision tree algorithms based on clinical consensus,[11] expert opinion, and prior observational data.[2,12,13] An accurate and validated clinical prediction model that could be deployed on the labor and delivery unit for postpartum hemorrhage is lacking.[14–17]

Current methods for predicting postpartum hemorrhage are based on risk stratification methods. Improved predictive ability could be achieved by applying traditional statistical and machine learning methods.[18] Recent advances in machine learning, which employs advanced computer-driven algorithms aimed at detecting patterns in data, have increasingly attracted attention because of their superior predictive ability primarily in determining intensive care unit admission and hospital readmission, compared with statistical models.[19,20] However, these innovative approaches have yet to be widely tested in obstetrics.[21] Advantages of machine learning include the ability to process nonadditive relationships and incorporate complex interactions between factors that do not need to be prespecified.[20] For these reasons, it is possible that machine learning approaches could accurately identify women at highest risk of postpartum hemorrhage and improve obstetric decision making,[21,22] and possibly improved clinical outcomes.

Our objective was to develop and validate prediction models for postpartum hemorrhage on labor admission. We quantified how closely the diagnosis of postpartum hemorrhage classified by machine learning and statistical models matched the known diagnosis of an estimated blood loss of at least 1,000 mL.

METHODS

Data from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development Consortium on Safe Labor were used for the development and validation of the prediction models. Briefly, the Consortium on Safe Labor was a retrospective cohort study of women delivering at 23 weeks of gestation or longer between 2002 and 2008 at 12 clinical sites with 19 hospitals across nine ACOG districts in the United States.[23] This cohort included data abstracted from electronic medical records, including demographics, prenatal complications, labor and delivery information, and maternal and neonatal outcomes. The Consortium on Safe Labor included a total of 228,438 deliveries at 23 weeks of gestation or longer, with 9.5% of women (n=5,053) contributing more than one birth over the study period. The present analysis was performed using a de-identified data set under a waiver of informed consent and was approved by the Institutional Review Board at the University of North Carolina at Chapel Hill. Methods and reporting guidelines were followed as proposed in the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) statement.[24]

The candidate predictors of postpartum hemorrhage for all models were chosen from routinely available data on labor admission and were compiled from expert opinion, consensus statements, prior observational cohorts, and collected from the Consortium on Safe Labor database.[25–27] We first included risk factors for postpartum hemorrhage a priori identified by the CMQCC and ACOG (Table 1 in Practice Bulletin No. 183).[7,28] Additional predictors included available patient socio-demographics (eg, age, race), obstetric diagnoses (eg, placenta previa, fetal macrosomia, preeclampsia), comorbid conditions (eg, chronic hypertension, diabetes), and vital signs on labor admission (Appendix 2, available online at http://links.lww.com/AOG/B788). Risk factors in the Consortium on Safe Labor were captured by searching for specific International Classification of Diseases, 9th Revision, Clinical Modification codes as well as from documentation in the labor and delivery clinical record. Some previously identified risk factors as outlined by CMQCC and ACOG that would be available to obstetric providers on labor admission were not available in this data set, including a history of prior postpartum hemorrhage and thrombocytopenia on admission.

The primary outcome of postpartum hemorrhage was defined as a cumulative estimated blood loss of at

least 1,000 mL, regardless of mode of delivery. This outcome was selected because it was consistent with the most recent definition of the ACOG reVITALize program,[28] as well as the consensus case definition of the Brighton Collaboration Primary Postpartum Hemorrhage Working Group,[25] the Royal College of Obstetricians and Gynaecologists,[26] and the World Health Organization.[27]

We developed two statistical models that used logistic regression and logistic regression with lasso regularization (lasso regression),[29] and two machine learning models that used random forest[30] and extreme gradient boosting algorithms.[31] Lasso regression also is referred to as penalized regression because a penalty is imposed on variables with high variance to eliminate the number of variables and improve model predictions. Recent review articles provide a framework for interpreting clinical studies that use machine learning methods for clinical readers.[29,32,33]

Variable reduction was performed during logistic regression using backwards stepwise elimination. The variable selection process started with the full model, and a bootstrap bias-corrected concordance index was used as the stopping criteria. Variables with individual $P>.05$ were left in the model if they offered information to improve the overall model. The removal of each variable was evaluated by determining which variable had the smallest effect on the adjusted R^2 and was stopped when the bootstrap concordance index had a change of 0.001. Variable reduction was performed during lasso regression by 10-fold cross-validation of the lambda value, and the final model incorporated the variables that were most predictive within one standard error of the best value.

Model performance was assessed using three recommended measures[24]: 1) the C statistic, or area under the receiver operating characteristic curve (AUC); 2) calibration curves; and 3) decision curves. The C statistic measures the model's overall ability to discriminate between high- and low-risk patients, but it does not allow one to understand how the model performs across the entire range of possible predictions, which is measured by the model's calibration curve. Calibration curves were plotted to show the relationship between the model's predicted outcomes against the cohort's observed outcome, where a perfectly calibrated model follows a 45° line.[34] In this study, decision curve analysis was used to quantify the net benefit of using each model and to visually compare the models.[35] Decision curve analysis evaluates the benefits of a diagnostic test, or a prediction model in this case, across a range of patient preferences for accepting risk of undertreatment and overtreatment to facilitate decisions about test selection and use.[35] Decision curve analysis assesses the value of information provided by the model by considering the likely range of a patient's risk and benefit preferences, without the need for actually measuring these preferences for a particular patient.[35] The net benefit is determined by calculating the difference between the expected benefit and the expected harm associated with each proposed testing and possible treatment strategies. The expected benefit is represented by the number of patients who have the outcome and who will receive treatment (true positives) using the proposed strategy.[36] The net benefit is calculated as: net benefit=true-positive rate−(false-positive rate×weighting factor), in which the weighting factor=threshold probability/1−threshold probability, and the threshold probability is a level of certainty above which the patient or physician would choose to intervene. Variable importance, a scaled measure with a maximum value of 100,[31] was plotted to understand the contribution of each predictor in the machine learning models. All models were internally validated using bootstrapping or cross-validation to measure optimism-corrected reliability. Missing predictor values were imputed using multiple imputed chained equations.[37] The frequency of missing data for each variable generally did not vary over time between the two time periods used in temporal validation (2002–2006 and 2007–2008) (Appendix 3, available online at http://links.lww.com/AOG/B788).

Measures, such as sensitivity, specificity, and false-positive and negative probabilities are not recommended when reporting performance of clinical prediction models because they are performance measures after introducing one or more artificial probability thresholds or catagories.[24] Although these are useful for estimating accuracy or classification measures often reported in a single diagnostic test or prognostic factor studies, such dichotomization and related classification measures lead to loss of information when providing a prediction for the future, and introducing such a threshold implies that it is relevant to clinical practice, which often is not the case.[24]

The data set was split by time (ie, temporal validation); we used the first phase (2002–2006) for model derivation and the second phase (2007–2008) for model validation. When the sample size is very large, this approach has been shown to be methodologically more rigorous than a simple random split of the data set.[38,39] Next, we combined both site-specific and temporal validation by using each site once as a validation sample, with the remaining sites used for model derivation during the first phase.

Site-specific estimates of discrimination (C statistic) and calibration were pooled and tested in the second phase. All analyses were performed with R statistical software 3.4.1.

RESULTS

Of 228,438 births in the Consortium on Safe Labor cohort across 12 sites, data were available from 10 of the sites for estimated blood loss (152,279, 66.6%) (Fig. 1). The mean estimated blood loss was 445 mL (SD 2,327) and was two times greater for cesarean deliveries than vaginal deliveries (769.9 mL vs 315.0 mL) (Appendix 4, available online at http://links.lww.com/AOG/B788). Of the 152,279 births assessed, 7,279 (4.7%, 95% CI 4.6–4.9) had postpartum hemorrhage as defined by an estimated blood loss of at least 1,000 mL, which was higher for cesarean deliveries (15.2%, 95% CI 14.9–15.6) than vaginal deliveries (0.6%, 95% CI 0.5–0.6). Of the 145,943 of 152,279 (95.8%) births with available transfusion data, the 11.3% (n=789) with postpartum hemorrhage had a transfusion compared with 1.9% (n=2,724) without postpartum hemorrhage.

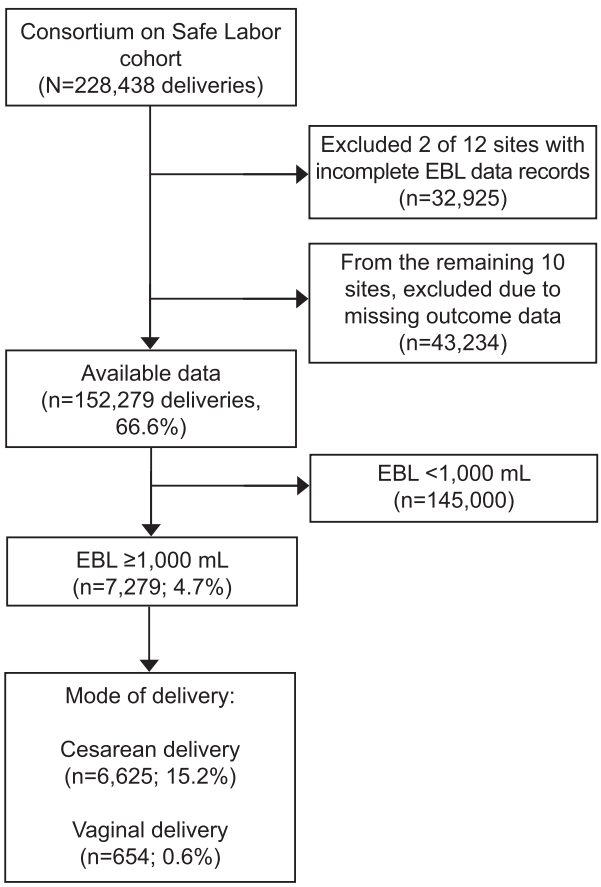


Fig. 1. Flow chart of women with postpartum hemorrhage (estimated blood loss [EBL] 1,000 mL or greater).

Venkatesh. Prediction of Postpartum Hemorrhage. Obstet Gynecol 2020.

A total of 55 candidate predictors of postpartum hemorrhage available on labor admission were assessed for possible model inclusion, including sociodemographic, obstetric, clinical, and physiologic variables (Appendix 2, http://links.lww.com/AOG/B788). Variables associated with postpartum hemorrhage included maternal age, prepregnancy body mass index, being of black race, residing in the South, delivery at a community hospital, gestational age at birth, maternal comorbid conditions (eg, pregestational and gestational diabetes, chronic and gestational hypertension), prior preterm birth, antepartum admission, threatened preterm birth, antenatal steroids, breech presentation, prelabor rupture of membranes, whether trial of labor was attempted, and whether labor was initiated spontaneously. The variables included in each model are presented in Appendix 5, available online at http://links.lww.com/AOG/B788.

After temporal and site validation, the best-performing model was the extreme gradient boosting model with the highest discriminative ability (C statistic 0.93, 95% CI 0.92–0.93) (Table 1). The random forest model also had a high discriminative ability (C statistic 0.92; 95% CI 0.91–0.92). The lasso regression (C statistic: 0.87; 95% CI 0.86–0.88), and the logistic regression model had lower discriminative ability (C statistic: 0.87; 95% CI 0.86–0.87). Appendix 6, available online at http://links.lww.com/AOG/B788, displays the receiver operating characteristic curves for the two machine learning and two statistical models.

Models were further validated both over time (ie, temporally) and by site for predicting postpartum hemorrhage. As above, we similarly noted good-to-excellent discrimination for prediction of postpartum hemorrhage with extreme gradient boosting (C statistic: 0.93; 95% CI 0.92–0.94), followed by random forest (C statistic: 0.92; 95% CI 0.91–0.92). Logistic regression (C statistic 0.87; 95% CI 0.86–0.87) and lasso regression (C statistic: 0.87; 95% CI 0.86–0.88) had lower discriminative ability.

Figure 2 demonstrates the overall calibration curves with 95% CIs of the four models. The calibration curve shows the variation in performance of each model in comparison with perfect agreement between the predicted probability of the model and the actual probability. We also present the best performing model, extreme gradient boosting, by each of the 10 assessed sites in Appendix 7, available online at http://links.lww.com/AOG/B788. This model assigned an accurate probability of postpartum hemorrhage when prediction ranged from 0% to 70–80%.

Table 1. Discrimination Between Women With and Without Postpartum Hemorrhage Using Two Machine Learning and Two Statistical Models

Model	Temporal Validation*		Temporal and Site Validation†	
	Before 2007	2007 or After	Before 2007	2007 or After
Extreme gradient boosting	0.95 (0.95–0.95)	0.93 (0.92–0.93)	0.93 (0.92–0.94)	0.93 (0.92–0.94)
Random forest	0.99 (0.99–1)	0.92 (0.91–0.92)	0.92 (0.91–0.92)	0.92 (0.91–0.92)
Logistic regression with lasso regularization	0.88 (0.87–0.88)	0.87 (0.86–0.88)	0.87 (0.86–0.88)	0.87 (0.86–0.88)
Logistic regression model	0.87 (0.87–0.88)	0.87 (0.86–0.87)	0.87 (0.86–0.87)	0.87 (0.86–0.87)

Data are concordance index (95% CI).

* In temporal validation, models were constructed from the first phase (2002–2006) and externally validated (ie, temporally) in the second phase (2007–2008).

† In temporal and site validation, both clinical site–specific (total of 10 sites) and temporal validation were combined by using each site once as a validation sample, with the remaining sites used for model derivation during the first phase.

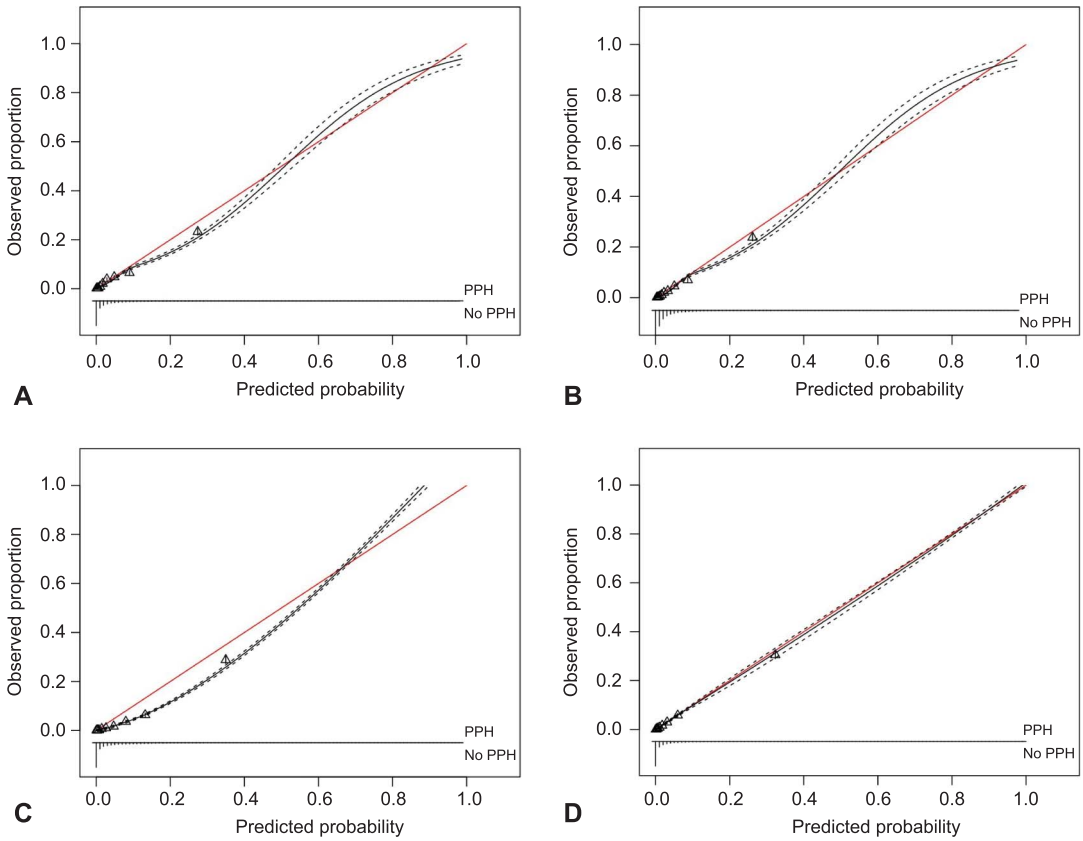


Fig. 2. Calibration curves demonstrating the performance of predicting postpartum hemorrhage (PPH) for all four models: logistic regression (**A**), logistic regression with lasso regularization (**B**), random forest (**C**), and extreme gradient boosting (**D**). The figure demonstrates the variation in each model's performance. The *red line* indicates perfect agreement between the predicted probability of the model and the actual probability. The *black line bounded by two dotted lines* indicates the overall calibration, with 95% CIs of each model. Each *triangle* represents a group of individualized risk. There are 10 *triangles*, and each *triangle* represents a decile of risk. **A**. Calibration (intercept: −0.17 [–0.22 to −0.11]; slope: 0.96 [0.92 to 1.00]); discrimination (C-statistic: 0.87 [0.86 to 0.87]). **B**. Calibration (intercept: −0.20 [–0.26 to −0.15]; slope: 1.08 [1.04 to 1.12]); discrimination (C-statistic: 0.87 [0.86 to 0.88]). **C**. Calibration (intercept: −0.60 [–0.66 to −0.54]; slope: 1.28 [1.23 to 1.33]); discrimination (C-statistic: 0.92 [0.91 to 0.92]). **D**. Calibration (intercept: −0.13 [–0.19 to −0.06]; slope: 1.02 [0.98 to 1.06]); discrimination (C-statistic: 0.93 [0.92 to 0.94]).

Venkatesh. Prediction of Postpartum Hemorrhage. Obstet Gynecol 2020.

In the decision curve analysis (Fig. 3), all models provided superior net benefit when clinical decision thresholds were between 0% and 80%. The net benefit for the extreme gradient boosting model was greatest across the range of threshold probabilities compared with the other models (ie, lasso, random forest, and logistic regression). The threshold probability is a level of certainty above which the patient or physician would choose to intervene. The probability threshold captures the relative value the patient or physician places on receiving an intervention for the outcome, if present, to the value of avoiding an intervention if the outcome is not present.

Figure 4 displays the variable importance in the extreme gradient boosting model for postpartum hemorrhage, which was the best performing model.

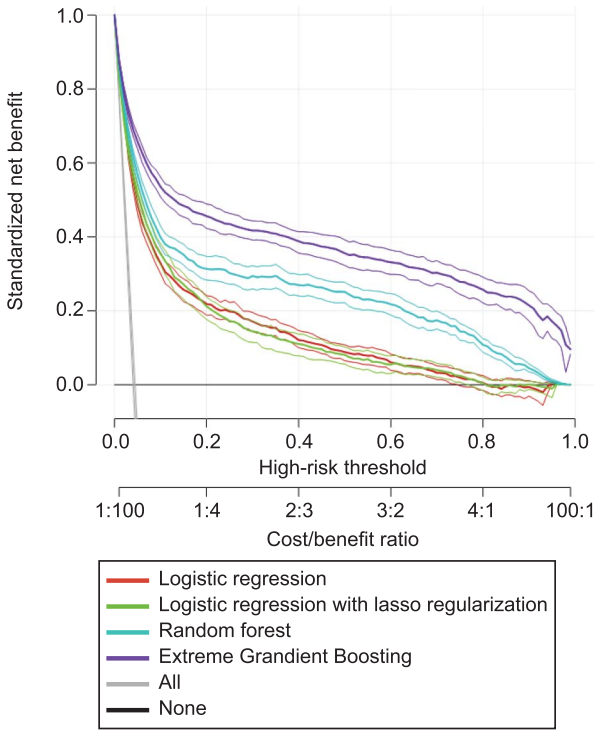


Fig. 3. Decision curve analysis of predicting postpartum hemorrhage by all models. The x-axis indicates the threshold probability for postpartum hemorrhage outcome. The y-axis indicates the net benefit. The net benefit is calculated as true-positive rate−(false-positive rate×weighting factor). Weighting factor is calculated as threshold probability/1−threshold probability. For example, when threshold probability is 0.1, weighting factor is 0.1/1−0.1=0.1/0.9. The decision curves indicate the net benefit of each model as well as two clinical alternatives (classifying no women as having the outcome vs classifying all women as having the outcome) over a specified range of threshold probabilities of outcome.

Venkatesh. Prediction of Postpartum Hemorrhage. Obstet Gynecol 2020.

The top 10 variables, ranked from most to least important, were prepregnancy maternal weight, admission maternal weight, prenatal diagnosis of fetal macrosomia, temperature at admission, attempted trial of labor on admission, prepregnancy maternal body mass index, admission systolic blood pressure, multiple gestation, anemia diagnosis during pregnancy, and spontaneous labor on admission.

We were unable to compare our models to current risk stratification strategies as outlined by CMQCC (ie, low, medium, and high risk for postpartum hemorrhage) because we only had two thirds of these variables available in the data set, and many of these variables (such a uterine fibroids, known bleeding disorder, or coagulopathy) were not routinely assessed on labor admission and, therefore, were not considered as candidate variables for our models.

DISCUSSION

We found that machine leaning and statistical models can accurately predict postpartum hemorrhage using

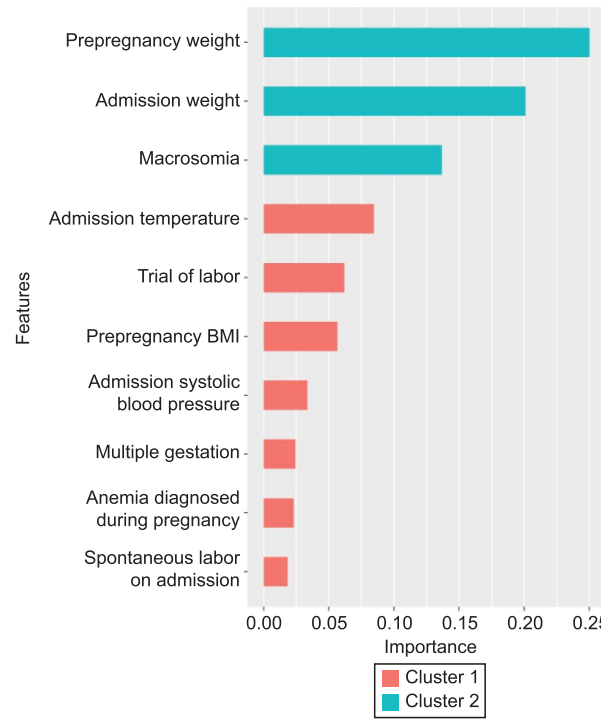


Fig. 4. Importance of each predictor in the extreme gradient boosting model to predict risk of postpartum hemorrhage. The variable importance is a measure scaled to have a maximum value of 1. Cluster indicates features that are similar to one another in importance value. BMI, body mass index.

Venkatesh. Prediction of Postpartum Hemorrhage. Obstet Gynecol 2020.

data available at the time of admission for labor. Machine leaning models performed the best, but at the cost of possibly increased complexity and minimal clinical significance. Extreme gradient boosting and random forest models provided excellent discriminative ability to predict postpartum hemorrhage. Importantly, these models achieved high predictive performance using clinical and physiologic data readily available to the obstetric provider at the time of labor admission.

Clinical application could allow obstetric providers to be prepared and, in some instances (ie, when not in spontaneous labor), to triage women at high risk of postpartum hemorrhage to the appropriate level of maternity care.[40] It would be reasonable to integrate the models into an online calculator or automated input in the electronic medical record for immediate use on labor admission. At the current time, statistical models are easier to integrate in this manner compared with many machine learning algorithms. Given that the statistical models performed well, this should also be considered. Regardless, any model will need to be prospectively assessed in local contemporary cohorts of pregnant women across regional settings. Our model focused on prediction at the time of labor admission and, hence, employed variables available at that time; it may also be reasonable to build models that include intrapartum variables that affect hemorrhage risk, such as length of labor and mode of delivery.

To date, accurate prediction models that use statistical models for postpartum hemorrhage have been lacking. Dilla et al[41] found that CMQCC risk strata of low, medium, and high risk generally predicted postpartum hemorrhage at a tertiary care center, but this analysis did not involve development of a prediction model that could be prospectively tested. Betts et al[42] aimed to identify women at risk for common postpartum complications using Australian administrative data and found good discrimination for postpartum hypertension and surgical site infection (both with AUC greater than 0.80), but not postpartum hemorrhage. Similarly, an earlier model did not accurately predict postpartum hemorrhage using data from the HYPITAT trial of women with gestational hypertension or preeclampsia in the Netherlands (AUC 0.59).[43] Albright et al[44] built a logistic regression model that adequately predicted transfusion after cesarean delivery using data from the Maternal-Fetal Medicine Units Cesarean Registry (AUC 0.82). Recently, models using traditional statistical methods have been developed with fair-to-good predictive ability for intensive care unit admission (AUC 0.81) and failed induction among obese women (AUC 0.79) using population-based administrative data.[45,46] In comparison with these prior models, our data suggest that both machine learning and statistical models can provide superior discriminative ability in the case of postpartum hemorrhage and that a final model should be chosen based on a combination of discrimination, calibration (especially accurate calibration in the range of predictions where clinical decisions will be affected), ease of use in the clinical setting, and acceptability by clinicians and patients.

Prediction of postpartum hemorrhage on labor admission could allow for optimizing labor and delivery unit health care resources, risk mitigation, and timely care.[47] Currently, health care providers rely largely on their clinical judgment and recognition of nonspecific risk factors. To further systematize this process, the Alliance for Innovation on Maternal Health has established an obstetric hemorrhage patient safety bundle, including policies, guidelines, and algorithms, to aid in the prompt recognition and management of postpartum hemorrhage.[8,28] The CMQCC categorizes pregnant women into low, medium, or high risk based on clinical or laboratory risk factors,[48] and statewide implementation of their hemorrhage protocol has been associated with a reduction in severe maternal morbidity.[49] Despite increasing impetus for wider adoption across the United States,[11] these risk-based stratification guidelines or decision tree algorithms are based on expert opinion and clinical consensus, and these methods do not provide an individualized risk prediction. In the current study, we did not formally compare our models to the CMQCC risk strata for both methodologic (ie, availability of only about two thirds of these variables) as well as pragmatic (ie, many of these variables are not routinely assessed on labor admission) considerations. Our definition of hemorrhage did not incorporate transfusion of blood products and further models to accurately predict transfusion beyond current guidelines are needed as recently highlighted.[50] We did note a 2% rate of transfusion among those women without postpartum hemorrhage, which may reflect outcome misclassification or cases of transfusion that occurred after delayed or late postpartum hemorrhage.

There are several study limitations. Missing data is an important limitation of this analysis, including restricting this analysis to the subset of the original cohort with available blood loss data, and the substantial proportion of covariates with incomplete data, although we used up-to-date imputation techniques. The proportion of missing data is a limitation to the

general application of these models, affecting generalizability of these results. However, it is likely that missing data or incomplete ascertainment will continue to be a limitation when applying these models in real time with electronic health record data.

Similar to clinical applications of precision medicine in other medical specialties, whether improved prediction will affect clinical outcomes and patient disposition compared with conventional clinical practice remains to be studied.[51] Alternatively stated, accurate prediction of postpartum hemorrhage may not necessarily change care or produce better patient outcomes. It is possible that inclusion of more predictors in the model, such as serial measurement of vital signs, physical examination findings, and laboratory data, may improve prediction, but it may not be feasible and could potentially delay intervention. The effect of a model on clinical decision making depends on multiple provider and care environment characteristics, including capacity to formulate a timely clinical response, weighing of risks and benefits of intervention, ability to execute that action, and patient (or health care provider) adherence to the recommended intervention.[19] Additional environmental constraints may include personnel, space, and equipment, which are not integrated into current prediction models. Furthermore, estimated blood loss is known to be imprecise, inaccurate, and often underestimated,[52] and there can be substantial variability in the relationship between blood loss and clinical signs and symptoms.[53]

Our definition of postpartum hemorrhage is consistent with current clinical guidelines; however, we did not assess other relevant clinical measures of acute blood loss, including a higher blood loss threshold, hemorrhage resulting in transfusion, drop of hemoglobin predelivery to postdelivery, or change in vital signs to suggest hemodynamic instability. Predictive models for these outcomes will need to be tested while using newer, more precise methods of assessing blood loss at delivery, including weight-based or photographic, colormetric quantitative blood loss approaches. Finally, machine learning approaches are data driven and depend on accurate data. Some important clinical variables were not measured in the Consortium on Safe Labor data set (eg, postpartum hemorrhage in prior pregnancy; thrombocytopenia; thromboprophylactic drug treatment; placental characteristics; and uterine fibroids). Some assessed variables, such as a prenatal diagnosis of macrosomia, were likely not universally captured, and it was unclear how they were defined. Additionally, the data set is now a decade old. These models will need to be replicated in more contemporary prospective data sets, ideally in real time as part of an integrated electronic medical record.

Strengths of this study include creation of a generalizable model drawn from a large data set from multiple hospitals across the United States over nearly a decade, as well as use of an analytical approach that has yet to widely studied and implemented in obstetrics. Our findings reinforce that machine learning approaches can be used to improve clinical prediction in obstetrics, and this "proof of concept" will need to be prospectively tested.

In conclusion, these findings are an opportunity to apply novel prediction approaches to support decision making on labor admission in an era of rising U.S. maternal morbidity and mortality. As predictive tools become more widely used in obstetric care,[45] they can be included in guidelines and care pathways for clinical use and further testing. Identification of women at high risk of postpartum hemorrhage on labor admission using both machine learning and statistical models could allow for more prompt diagnosis and possibly intervention, which may result in more accurate clinical care, improved patient outcomes, and better resource allocation.

REFERENCES

1. Say L, Chou D, Gemmill A, Tunçalp Ö, Moller AB, Daniels J, et al. Global causes of maternal death: a WHO systematic analysis. Lancet Glob Health 2014;2:e323–33.
2. Bateman B, Berman MF, Riley LE, Leffert LR. The epidemiology of postpartum hemorrhage in a large, nationwide sample of deliveries. Anaesth Analgesia 2010;110:1368–73.
3. Callaghan W, Kuklina EV, Berg CJ. Trends in postpartum hemorrhage: United States, 1994-2006. Am J Obstet Gynecol 2010;202:353.e1–6.
4. Creanga A, Berg CJ, Ko JY, Farr SL, Tong VT, Bruce FC, et al. Maternal mortality and morbidity in the United States: where are we now? J Womens Health 2014;23:3–9.
5. Callaghan W, Mackay AP, Berg CJ. Identification of severe maternal morbidity during delivery hospitalizations, United States, 1991-2003. Am J Obstet Gynecol 2008;199:133.e1–8.
6. Hamm R, Wang E, Romanos A, O'Rourke K, Srinivas SK. Implementation of quantification of blood loss does not improve prediction of hemoglobin drop in deliveries with average blood loss. Am J Perinatol 2018;35:134–9.
7. Postpartum hemorrhage. Practice Bulletin No. 183. American College of Obstetricians and Gynecologists. Obstet Gynecol 2017;130:e168–86.
8. Council on Patient Safety in Women's Health Care. Obstetric hemorrhage (+AIM). Available at: https://safehealthcareforeverywoman.org/patient-safety-bundles/obstetric-hemorrhage/. Retrieved March 25, 2019.
9. Dahlke J, Mendez-Figueroa H, Maggio L, Hauspurg AK, Sperling JD, Chauhan SP, et al. Prevention and management of postpartum hemorrhage: a comparison of 4 national guidelines. Am J Obstet Gynecol 2015;213:76.e1–10.
10. Shields L, Wiesner S, Fulton J, Pelletreau B. Comprehensive maternal hemorrhage protocols reduce the use of blood products and improve patient safety. Am J Obstet Gynecol 2015;212:272–80.
11. Main E, Goffman D, Scavone BM, Low LK, Bingham D, Fontaine PL, et al. National partnership for maternal safety;

council on patient safety in women's health care. National partnership for maternal safety: consensus bundle on obstetric hemorrhage. Obstetrics Gynecol 2015;126:155–62.

12. Wetta L, Szychowski JM, Seals S, Mancuso MS, Biggio JR, Tita AT. Risk factors for uterine atony/postpartum hemorrhage requiring treatment after vaginal delivery. Am J Obstetrics Gynecol 2013;209:51.e1–6.
13. Kramer M, Berg C, Abenhaim H, Dahhou M, Rouleau J, Mehrabadi A, et al. Incidence, risk factors, and temporal trends in severe postpartum hemorrhage. Am J Obstet Gynecol 2013;209:449.e1–7.
14. Prata N, Hamza S, Bell S, Karasek D, Vahidnia F, Holston M. Inability to predict postpartum hemorrhage: insights from Egyptian intervention data. BMC Pregnancy Childbirth 2011:11:97.
15. Mousa H, Cording V, Alfirevic Z. Risk factors and interventions associated with major primary postpartum hemorrhage unresponsive to first-line conventional therapy. Acta Obstet Gynecol Scand 2008;87:652–61.
16. Biguzzi E, Franchi F, Ambrogi F, Ibrahim B, Bucciarelli P, Acaia B, et al. Risk factors for postpartum hemorrhage in a cohort of 6011 Italian women. Thromb Res 2012;129:e1–7.
17. Helman S, Drukker L, Fruchtman H, Ioscovich A, Farkash R, Avitan T, et al. Revisit of risk factors for major obstetric hemorrhage: insights from a large medical center. Arch Obstet Gynecol 2015;292:819–28.
18. Cuocolo R, Perillo T, De Rosa E, Ugga L, Petretta M. Current applications of big data and machine learning in cardiology. J Geriatr Cardiol 2019;16:601–7.
19. Shah N, Milstein A, Bagley SC. Making machine learning models clinically useful. J Am Med Assoc 2019;322:1351.
20. Goto T, Camargo CA, Faridi MK, Freishtat RJ, Hasegawa K. Machine learning-based prediction of clinical outcomes for children during emergency department triage. JAMA Open Netw 2019;2:e186937.
21. Escobar G, Gupta NR, Walsh EM, Soltesz L, Terry SM, Kipnis P. Automated early detection of obstetric complications: theoretic and methodologic considerations. Am J Obstet Gynecol 2019;220:297–307.
22. Fohner A, Greene JD, Lawson BL, Chen JH, Kipnis P, Escobar GJ, et al. Assessing clinical heterogeneity in sepsis through treatment patterns and machine learning. J Am Med Inform Assoc 2019;26:1466–77.
23. Zhang J, Troendle J, Reddy UM, Laughon SK, Branch DW, Burkman R, et al. Contemporary cesarean delivery practice in the United States. Am J Obstet Gynecol 2010;203:326.e1–10.
24. Moons K, 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. Ann Intern Med 2015;162:W1–73.
25. Kerr R, Eckert LO, Winikoff B, Durocher J, Meher S, Fawcus S, et al. Postpartum haemorrhage: case definition and guidelines for data collection, analysis, and presentation of immunization safety data. Vaccine 2016;34:6102–9.
26. Royal College of Obstetricians and Gynaecologists. Postpartum haemorrhage, prevention and management (Green-top guideline No. 52). Available at: https://www.rcog.org.uk/en/guidelines-research-services/guidelines/gtg52/. Retrieved March 16, 2019.
27. World Health Organization. The prevention and management of postpartum haemorrhage: report of a technical working group.Geneva, Switzerland: World Health Organization; 1989.
28. Menard M, Main EK, Currigan SM. Executive summary of the reVITALize initiative: standardizing obstetric data definitions. Obstetrics Gynecol 2014;124:150–3.
29. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc 1996;58:267–88.
30. Breiman L. Random forests. Machine Learn 2001;45:5–32.
31. xgboost: extreme gradient boosting. Available at: https://cran.r-project.org/web/packages/caret/index.html. Retrieved October 12, 2019.
32. Bi Q, Goodman KE, Kaminsky J, Lessler J. What is machine learning? A primer for the epidemiologist. Am J Epidemiol 2019 [Epub ahead of print].
33. Liu Y, Chen PHC, Krause J, Peng L. How to read articles that use machine learning users' guides to the medical literature. J Am Med Assoc 2019;322:1806–16.
34. Jelovsek J, Hill AJ, Chagin KM, Kattan MW, Barber MD. Predicting risk of urinary incontinence and adverse events after midurethral sling surgery in women. Obstetrics Gynecol 2016;127:330–40.
35. Vickers A, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26:565–74.
36. Fitzgerald M, Saville BR, Lewis RJ. Decision curve analysis. J Am Med Assoc 2015;313:409–10.
37. White I, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med 2011;30:377–99.
38. Roberts D, Bahn V, Ciuti S, Boyce M, Elith J, Guillera-Arroita G, et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 2017;40:913–29.
39. Steyerberg E. Validation of prediction models. In: Clinical prediction models: a practical approach to development, validation, and updating (statistics for biology and health). 2nd ed. Cham, Switzerland: Springer; 2019.
40. Levels of maternal care. Obstetric Care Consensus No. 9 [published erratum appears in Obstet Gynecol 2019;134:883]. American College of Obstetricians and Gynecologists. Obstet Gynecol 2019;134:e41–55.
41. Dilla A, Waters JH, Yazer MH. Clinical validation of risk stratification criteria for peripartum hemorrhage. Obstetrics Gynecol 2013;122:120–6.
42. Betts K, Kisely S, Alati R. Predicting common maternal postpartum complications: leveraging health administrative data and machine learning. BJOG 2019;126:702–9.
43. Koopmans C, van der Tuuk K, Groen H, Doornbos JP, de Graaf IM, van der Salm PC, et al. Prediction of postpartum hemorrhage in women with gestational hypertension or mild preeclampsia at term. Acta Obstet Gynecol Scand 2014;93:399–407.
44. Albright C, Spillane TE, Hughes BL, Rouse DJ. A regression model for prediction of cesarean-associated blood transfusion. Am J Perinatol 2019;36:879–85.
45. Rossi R, Hall E, Dufendach K, DeFranco EA. Predictive model of factors associated with maternal intensive care unit admission. Obstetrics Gynecol 2019;134:216–24.
46. Rossi R, Requarth EW, Warshak CR, Dufendach K, Hall ES, DeFranco EA. Predictive model for failed induction of labor among obese women. Obstetrics Gynecol 2019;134:485–93.
47. Merriam A, Wright JD, Siddiq Z, D'Alton ME, Friedman AM, Ananth CV, et al. Risk for postpartum hemorrhage, transfusion, and hemorrhage-related morbidity at low, moderate, and high volume hospitals. J Matern Fetal Neonatal Med 2018;31:1025–34.
48. California Maternal Quality Care Collaborative. Improving health care response to obstetric hemorrhage version 2.0. A California quality improvement toolkit. Available at: https://www.cmqcc.org/resources-tool-kits/toolkits/ob-hemorrhage-toolkit. Retrieved October 10, 2019.

49. Main E, Cape V, Abreo A, Vasher J, Woods A, Carpenter A, et al. Reduction of severe maternal morbidity from hemorrhage using a state perinatal quality collaborative. Am J Obstet Gynecol 2017;216:298.e1–11.

50. Kawakita T, Mokhtari N, Huang JC, Landy HJ. Evaluation of risk-assessment tools for severe postpartum hemorrhage in women undergoing cesarean delivery. Obstetrics Gynecol 2019;134:1308–16.

51. Emanuel E, Wachter RM. Artificial intelligence in health care: will the value match the hype? J Am Med Assoc 2019;321:2281–2.

52. Dildy GA III, Paine AR, George NC, Velasco C. Estimating blood loss: can teaching significantly improve visual estimation? Obstetrics Gynecol 2004;104:601–6.

53. Pacagnella R, Souza JP, Durocher J, Perel P, Blum J, Winikoff B, et al. A systematic review of the relationship between blood loss and clinical signs. PLoS One 2013;8:e57594.

PEER REVIEW HISTORY

Received November 20, 2019. Received in revised form January 8, 2020. Accepted January 16, 2020. Peer reviews and author correspondence are available at http://links.lww.com/AOG/B789.

Appendix 1.

Institutions involved in the Consortium include, in alphabetical order: Baystate Medical Center, Springfield, MA; Cedars-Sinai Medical Center Burnes Allen Research Center, Los Angeles, CA; Christiana Care Health System, Newark, DE; Georgetown University Hospital , MedStar Health, Washington, DC; Indiana University Clarian Health, Indianapolis, IN; Intermountain Healthcare and the University of Utah, Salt Lake City, Utah; Maimonides Medical Center, Brooklyn, NY; MetroHealth Medical Center, Cleveland, OH.; Summa Health System, Akron City Hospital, Akron, OH; The EMMES Corporation, Rockville MD (Data Coordinating Center); University of Illinois at Chicago, Chicago, IL; University of Miami, Miami, FL; and University of Texas Health Science Center at Houston, Houston, Texas.

Venkatesh KK, Strauss RA, Grotegut C, Heine RP, Chescheir NC, Stringer JSA, et al. Machine learning and statistical models to predict postpartum hemorrhage. Obstet Gynecol 2020;135.
The authors provided this information as a supplement to their article.

Appendix 2. Characteristics Overall and by EBL			
Characteristic	**Overall N=228,438**	**EBL N=152,279**	
		EBL ≥1,000 mL	**EBL <1,000 mL**
		N=7,279	**N=145,000**
Mode of delivery, overall			
Vaginal	162,448 (71.1)	654 (8.9)	108,218 (74.6)
Cesarean	65,990 (28.8)	6,625 (91.0)	36,782 (25.3)
Mode of delivery, by subtype			
Vaginal	146,073 (63.9)	558 (7.6)	97,381 (67.1)
Operative	10,698 (4.6)	71 (0.9)	6,600 (4.5)
Vaginal birth after cesarean	5,677 (2.4)	25 (0.3)	4,237 (2.9)
Primary cesarean	39,577 (17.3)	4,303 (59.1)	22,011 (15.1)
Repeat cesarean	26,413 (11.5)	2,322 (31.9)	14,771 (10.1)
Socio-demographic variables			
Age, mean (SD), years	27.6 (6.19)	29.2 (6.60)	27.5 (6.29)
Age, median (IQR), years	29 (24, 30)	30 (25, 35)	27 (23, 32)
Missing n (%)	323 (0.1%)	3 (0.1%)	272 (0.1%)
Race			
White	113,224 (51.7)	3,078 (44.4)	69,410 (49.9)
Black	51,392 (23.5)	2,347 (33.8)	37,131 (26.7)
Latina	39,716 (18.1)	1,048 (15.1)	22,522 (16.2)
Asian	9,345 (4.3)	316 (4.5)	6,876 (4.9)
Other	5,400 (2.5)	142 (2.0)	2,953 (2.1)
Missing n (%)	9,361 (4.0)	119 (4.4)	6337 (4.2)
Region			
West	68,712 (30.0)	803 (11.0)	37,109 (25.5)
Midwest	34,683 (15.1)	1,824 (25.0)	30,793 (21.2)
South	69,268 (30.3)	2,822 (38.7)	44,080 (30.4)
Northeast	55,775 (24.4)	1,830 (25.1)	33,018 (22.7)
Missing n (%)	0 (--)	0 (--)	0 (--)
Academic hospital	130,205 (57.0)	3,591 (49.3)	72,987 (50.3)
Missing n (%)	0 (--)	0 (--)	0 (--)
Delivery year			
2002	7,357 (3.2)	455 (6.2)	6,385 (4.4)
2003	9,649 (4.2)	590 (8.1)	8,613 (5.9)
2004	15,576 (6.8)	755 (10.3)	13,331 (9.2)
2005	63,680 (32.0)	1,814 (24.9)	37,056 (25.5)
2006	73,154 (32.0)	2,071 (28.4)	43,625 (30.0)
2007	58,671 (25.6)	1,575 (21.6)	35,670 (24.6)
2008	292 (0.1)	3 (0.1)	278 (0.1)

Missing n (%)	59 (0)	0 (--)	44 (0.0)
Pre-pregnancy body mass index (BMI)			
Mean (SD), kg/m^2	25.4 (6.25)	29.0 (8.55)	25.4 (6.28)
Median (IQR), kg/m^2	23.9 (21.0, 28.2)	26.8 (22.5, 33.6)	23.8 (21.0, 28.2)
Missing n (%)	76,818 (33.6)	1,257 (46.5)	59,110 (39.5)
Pre-pregnancy weight			
Mean (SD), kg	68.0 (17.6)	77.4 (24.4)	67.8 (17.6)
Median (IQR), kg	63.5 (56.2, 75.8)	70.7 (58.9, 90.2)	63.5 (55.8, 75.7)
Missing n (%)	67,294 (29.4)	1,239 (45.8)	57,721 (38.5)
Admission weight			
Mean (SD), kg	82.5 (17.9)	92.2 (23.7)	82.3 (17.8)
Median (IQR), kg	79.3 (70.0, 91.1)	87.9 (75.0, 104.3)	78.9 (69.9, 91.1)
Missing n (%)	35,984 (15.7)	624 (23.0)	30,212 (20.1)
Tobacco use	15,247 (6.6)	499 (6.8)	9,889 (6.8)
Missing n (%)	0 (--)	0 (--)	0 (--)
Illicit drug use	4,725 (2.0)	220 (3.1)	3,337 (2.4)
Missing n (%)	22,544 (9.8%)	84 (3.1)	6,526 (4.3)
Maternal comorbidities			
Pre-gestational diabetes	5,305 (2.3)	495 (6.8)	3,747 (2.5)
Missing n (%)	0 (--)	0 (--)	0 (--)
Gestational diabetes	11,999 (5.2)	700 (9.6)	7,223 (4.9)
Missing n (%)	0 (--)	0 (--)	0 (--)
Chronic hypertension	7,690 (3.3)	467 (6.4)	4,073 (2.8)
Missing n (%)	0 (--)	0 (--)	0 (--)
Gestational hypertension	6,286 (2.7)	257 (3.5)	3,589 (2.4)
Missing n (%)	0 (--)	0 (--)	0 (--)
Heart disease	3,481 (1.5)	176 (2.4)	2,170 (1.5)
Missing n (%)	0 (--)	0 (--)	0 (--)
Thyroid disease	5,583 (2.5)	232 (3.2)	3,301 (2.3)
Missing n (%)	7,878 (3.4)	84 (3.1)	6,526 (4.3)
Renal disease	1,856 (0.8)	98 (1.3)	1,574 (1.0)
Missing n (%)	0 (--)	0 (--)	0 (--)
Asthma	17,490 (7.6)	775 (10.6)	11,996 (8.2)
Missing n (%)	0 (--)	0 (--)	0 (--)
Anemia diagnosis	23,057 (10.4)	1,900 (26.9)	16,157 (11.6)
Missing n (%)	7,877 (3.4)	84 (3.1)	6,526 (4.3)
Depression	9.850 (4.3)	494 (6.7)	6,956 (4.8)
Missing n (%)	0 (--)	0 (--)	0 (--)
Gastrointestinal disease	2,592 (1.1)	95 (1.3)	1,342 (0.9)
Missing n (%)	0 (--)	0 (--)	0 (--)
Active seizure disorder	210 (0.1)	21 (0.3)	160 (0.1)
Missing n (%)	24,679 (10.8)	122 (4.5)	6,214 (4.1)
History of seizure disorder	1,528 (0.7)	99 (1.4)	1,077 (0.7)
Missing n (%)	22,544 (9.8)	84 (3.1)	6,526 (4.3)
Obstetrical characteristics			
Parity			
Mean (SD)	1.1 (1.36)	1.1 (1.38)	1.1 (1.37)
Median (IQR)	1 (0, 2)	1 (0, 2)	1 (0, 2)
Missing n (%)	0 (--)	0 (--)	0 (--)
Assisted reproductive technology	1,101 (0.9)	123 (2.8)	607 (0.8)

Missing n (%)	107,479 (47.0)	1,216 (44.9)	72,745 (48.6)
Multiple gestation			
Twins	4,840 (2.1)	768 (10.5)	2,670 (1.8)
Higher order	213 (0.0)	69 (0.9)	87 (0.0)
Missing n (%)	0 (--)	0 (--)	0 (--)
Breech presentation	11,647 (5.1)	1,170 (16.0)	5,923 (4.0)
Missing n (%)	0 (--)	0 (--)	0 (--)
History of prior preterm birth	17,017 (7.6)	788 (10.8)	11,811 (8.1)
Missing n (%)	5,436 (2.3)	1 (--)	131 (0.1)
Polyhydramnios	885 (0.4)	82 (1.2)	509 (0.3)
Missing n (%)	14,169 (6.2)	356 (13.1)	13,022 (8.7)
Macrosomia	1,858 (1.5)	222 (3.2)	1,191 (1.5)
Missing n (%)	102,247 (44.7)	153 (5.6)	70,424 (47.0)
Pregnancy complications			
Prior cesarean delivery	31,321 (14.5)	2,270 (31.8)	18,403 (13.2)
Missing n (%)	13,219 (5.7)	66 (2.4)	5,638 (3.7)
Mild preeclampsia	7,457 (3.2)	483 (6.6)	5,127 (3.5)
Missing n (%)	0 (--)	0 (--)	0 (--)
Severe preeclampsia	3,772 (1.6)	278 (3.8)	2,363 (1.6)
Missing n (%)	0 (--)	0 (--)	0 (--)
Superimposed preeclampsia	1,971 (0.8)	151 (2.0)	1,064 (0.7)
Missing n (%)	0 (--)	0 (--)	0 (--)
Eclampsia	326 (0.1)	21 (0.3)	219 (0.1)
Missing n (%)	15,614 (6.8)	316 (11.6)	13,328 (8.9)
Fetal death	1,145 (0.5)	69 (0.9)	436 (0.3)
Missing n (%)	0 (--)	0 (--)	0 (--)
Group B streptococcus colonization	40,207 (17.6)	1,236 (16.9)	27,736 (19.1)
Missing n (%)	0 (--)	0 (--)	0 (--)
Intrapartum magnesium sulfate	8,763 (4.3)	708 (10.0)	6,648 (4.8)
Missing n (%)	26,136 (11.4)	84 (3.1)	6,526 (4.3)
Placental abruption	3,794 (1.6)	431 (5.9)	2,316 (1.6)
Missing n (%)	0 (--)	0 (--)	0 (--)
Fetal growth restriction	2,678 (1.2)	105 (1.6)	1,719 (1.3)
Missing n (%)	15,626 (6.8)	318 (11.7)	13,334 (8.9)
Large for gestational age, antepartum diagnosis	2,031 (1.4)	203 (3.5)	1,127 (1.1)
Missing n (%)	84,338 (36.9)	627 (23.2)	50,937 (34.0)
Antenatal steroids	5,945 (4.0)	388 (9.5)	3,447 (3.7)
Missing n (%)	82,507 (36.1)	1,104 (40.8)	56,141 (37.5)
Antepartum bleeding in third trimester	3,088 (1.7)	213 (3.9)	2,192 (2.1)
Missing n (%)	49,410 (21.6)	767 (28.3)	45,372 (30.3)
Placenta previa	1,647 (0.7)	357 (4.9)	738 (0.5)
Missing n (%)	0 (--)	0 (--)	0 (--)
Placenta accreta	138 (0.0)	69 (1.2)	51 (0.0)
Missing n (%)	82,335 (36.0)	767 (28.3)	45,372 (30.3)
Threatened preterm birth antepartum	7,880 (3.4)	366 (5.0)	4,367 (3.0)
Missing n (%)	0 (--)	0 (--)	0 (--)
Antepartum hospital admission	17,500 (7.6)	429 (13.2)	8,620 (11.3)
Missing n (%)	84,603 (37.0)	1,378 (50.9)	71,627 (47.8)
Gestational age at delivery, weeks			

Mean (SD) Median (IQR)	38.2 (2.47) 39 (38, 40)	37.2 (3.50) 38 (36, 39)	38.2 (2.43) 39 (38, 40)
Missing n (%)	0 (--)	0 (--)	0 (--)
Characteristics on L&D admission			
Temperature (F) Mean (SD)	98.0 (0.81)	98.1 (0.89)	98.0 (0.80)
Missing n (%)	47,977 (21.0)	682 (25.2)	30,097 (20.1)
Systolic blood pressure Mean (SD)	124.1 (14.89)	125.6 (16.81)	123.9 (14.73)
Missing n (%)	52,766 (23.0)	618 (22.8)	34,530 (23.0)
Diastolic blood pressure Mean (SD)	74.3 (11.6)	74.0 (13.07)	74.1 (11.76)
Missing n (%)		617 (22.8)	34,549 (23.0)
Trial of labor	192,074 (84.0)	3,684 (50.6)	124,529 (85.9)
Missing n (%)	0 (--)	0 (--)	0 (--)
Spontaneous labor	122,673 (53.7)	2,412 (33.1)	80,749 (55.6)
Missing n (%)	0 (--)	0 (--)	0 (--)
Premature rupture of membranes	16,219 (7.1)	620 (8.5)	11,768 (8.1)
Missing n (%)	0 (--)	0 (--)	0 (--)
Chorioamnionitis on admission	1,527 (1.0)	157 (2.6)	886 (0.9)
Missing n (%)	83,009 (36.3)	515 (19.0)	50,343 (33.6)

Appendix 3. Characteristics Overall and by Time Period			
Characteristic	**Overall** N=228,438	**Comparison between two time periods (2002-2006 vs. 2007-2008) for temporal validation N=152,232***	
		2002-2006	**2007-2008**
		N=114,695	**N=37,537**
Estimated blood loss (EBL)			
EBL ≥1,000 mL	--	5,685 (5.0)	1,589 (4.2)
EBL <1,000 mL		109,010 (95.0)	35,948 (95.8)
Mode of delivery, overall			
Vaginal	162,448 (71.1)	81,892 (71.5)	26,851 (71.5)
Cesarean	65,990 (28.8)	32,713 (28.5)	10,686 (28.5)
Mode of delivery, by subtype			
Vaginal	146,073 (63.9)	73,848 (64.4)	24,055 (64.1)
Operative	10,698 (4.6)	4,944 (4.3)	1,726 (4.6)
Vaginal birth after cesarean	5,677 (2.4)	3,190 (2.8)	1,070 (2.9)
Primary cesarean	39,577 (17.3)	20,028 (17.5)	6,282 (16.7)
Repeat cesarean	26,413 (11.5)	12,685 (11.1)	4,404 (11.7)
Socio-demographic variables			
Age, mean (SD), years	27.6 (6.19)	27.6 (6.33)	27.5 (6.28)
Age, median (IQR), years	29 (24, 30)	27 (23, 32)	27 (23, 32)
Missing n (%)	323 (0.1)	233 (0.2)	41 (0.1)
Race			
White	113,224 (51.7)	53,670 (48.9)	18,806 (52.1)
Black	51,392 (23.5)	29,914 (27.3)	9,541 (26.4)
Latina	39,716 (18.1)	18,382 (16.8)	5,187 (14.4)
Asian	9,345 (4.3)	5,325 (4.9)	1,865 (5.2)
Other	5,400 (2.5)	2,414 (2.2)	680 (1.9)
Missing n (%)	9,361 (4.0)	4,990 (4.4)	1,458 (3.9)
Region			
West	68,712 (30.0)	26,590 (23.2)	11,322 (30.2)
Midwest	34,683 (15.1)	23,044 (20.1)	9,571 (25.5)
South	69,268 (30.3)	36,870 (32.2)	9,987 (26.6)
Northeast	55,775 (24.4)	28,191 (24.6)	6,657 (17.7)
Missing n (%)	0 (--)	0 (--)	0 (--)
Academic hospital	130,205 (57.0)	57,584 (50.2)	18,959 (50.5)
Missing n (%)	0 (--)	0 (--)	0 (--)
Pre-pregnancy body mass index (BMI)			
Mean (SD), kg/m^2	25.4 (6.25)	25.5 (6.42)	25.6 (6.45)
Median (IQR), kg/m^2	23.9 (21.0, 28.2)	23.9 (21.0, 28.3)	23.9 (21.1, 28.5)
Missing n (%)	76,818 (33.6)	46,979 (40.9)	13,357 (35.6)
Pre-pregnancy weight			
Mean (SD), kg	68.0 (17.6)	68.2 (18.0)	68.3 (18.1)
Median (IQR), kg	63.5 (56.2, 75.8)	63.5 (56.2, 76.6)	63.5 (56.2, 77.0)
Missing n (%)	67,294 (29.4)	45,881 (40.0)	13,049 (34.7)

Admission weight			
Mean (SD), kg	82.5 (17.9)	82.8 (18.3)	82.7 (18.2)
Median (IQR), kg	79.3 (70.0, 91.1)	79.4 (70.0, 91.6)	79.4 (69.9, 91.6)
Missing n (%)	35,984 (15.7)	22,256 (19.4)	8,566 (22.8)
Tobacco use	15,247 (6.6)	7,751 (6.8)	2,637 (7.0)
Missing n (%)	0 (--)	0 (--)	0 (--)
Illicit drug use	4,725 (2.0)	2,698 (2.4)	858 (2.3)
Missing n (%)	22,544 (9.8)	4,864 (4.2)	1,746 (4.7)
Maternal comorbidities			
Pre-gestational diabetes	5,305 (2.3)	3,410 (3.0)	832 (2.2)
Missing n (%)	0 (--)	0 (--)	0 (--)
Gestational diabetes	11,999 (5.2)	6,044 (5.3)	1,877 (5.0)
Missing n (%)	0 (--)	0 (--)	0 (--)
Chronic hypertension	7,690 (3.3)	3,366 (2.9)	1,173 (3.1)
Missing n (%)	0 (--)	0 (--)	0 (--)
Gestational hypertension	6,286 (2.7)	2,964 (2.6)	881 (2.4)
Missing n (%)	0 (--)	0 (--)	0 (--)
Heart disease	3,481 (1.5)	1,785 (1.6)	561 (1.5)
Missing n (%)	0 (--)	0 (--)	0 (--)
Thyroid disease	5,583 (2.5)	2,637 (2.4)	896 (2.5)
Missing n (%)	7,878 (3.4)	4,864 (4.2)	1,746 (4.7)
Renal disease	1,856 (0.8)	1,259 (1.1)	413 (1.1)
Missing n (%)	0 (--)	0 (--)	0 (--)
Asthma	17,490 (7.6)	9,729 (8.5)	3,041 (8.1)
Missing n (%)	0 (--)	0 (--)	0 (--)
Anemia diagnosis	23,057 (10.4)	12,938 (11.3)	5,114 (13.6)
Missing n (%)	7,877 (3.4)	4,864 (4.2)	1,746 (4.6)
Depression	9.850 (4.3)	5,225 (4.6)	2,224 (5.9)
Missing n (%)	0 (--)	0 (--)	0 (--)
Gastrointestinal disease	2,592 (1.1)	1,031 (0.9)	406 (1.1)
Missing n (%)	0 (--)	0 (--)	0 (--)
Active seizure disorder	210 (0.1)	144 (0.1)	37 (0.1)
Missing n (%)	24,679 (10.8)	2,677 (2.3)	3,659 (9.8)
History of seizure disorder	1,528 (0.7)	889 (0.8)	287 (0.8)
Missing n (%)	22,544 (9.8)	4,864 (4.2)	1,746 (4.7)
Obstetrical characteristics			
Parity			
Mean (SD)	1.1 (1.36)	1.1 (1.37)	1.2 (1.37)
Median (IQR)	1 (0, 2)	1 (0, 2)	1 (0, 2)
Missing n (%)	0 (--)	0 (--)	0 (--)
Assisted reproductive technology	1,101 (0.9)	547 (0.5)	183 (0.5)
Missing n (%)	107,479 (47.0)	55,449 (48.3)	18,512 (49.3)
Multiple gestation			
Twins	4,840 (2.1)	2,650 (2.3)	782 (2.1)
Higher order	213 (0.0)	131 (0.1)	25 (0.1)
Missing n (%)	0 (--)	0 (--)	0 (--)
Breech presentation	11,647 (5.1)	6,123 (5.3)	1,863 (5.0)
Missing n (%)	0 (--)	0 (--)	0 (--)
History of prior preterm birth	17,017 (7.6)	9,416 (8.2)	3,181 (8.5)
Missing n (%)	5,436 (2.3)	83 (0.1)	49 (0.1)

Polyhydramnios	885 (0.4)	468 (0.4)	123 (0.3)
Missing n (%)	14,169 (6.2)	7,560 (6.6)	5,818 (15.5)
Macrosomia	1,858 (1.5)	1,069 (0.9)	344 (0.9)
Missing n (%)	102,247 (44.7)	53,835 (46.9)	16,713 (44.5)
Pregnancy complications			
Prior cesarean delivery	31,321 (14.5)	15,441 (13.5)	5,226 (13.9)
Missing n (%)	13,219 (5.7)	3,613 (3.2)	2,083 (5.6)
Mild preeclampsia	7,457 (3.2)	4,237 (3.7)	1,371 (3.7)
Missing n (%)	0 (--)	0 (--)	0 (--)
Severe preeclampsia	3,772 (1.6)	1,909 (1.7)	731 (2.0)
Missing n (%)	0 (--)	0 (--)	0 (--)
Superimposed preeclampsia	1,971 (0.8)	912 (0.8)	302 (0.8)
Missing n (%)	0 (--)	0 (--)	0 (--)
Eclampsia	326 (0.1)	183 (0.2)	57 (0.2)
Missing n (%)	15,614 (6.8)	9,741 (8.5)	3,903 (10.4)
Fetal death	1,145 (0.5)	559 (0.5)	168 (0.5)
Missing n (%)	0 (--)	0 (--)	0 (--)
Group B streptococcus colonization	40,207 (17.6)	21,533 (18.8)	7,432 (19.8)
Missing n (%)	0 (--)	0 (--)	0 (--)
Intrapartum magnesium sulfate	8,763 (4.3)	5,489 (4.8)	1,864 (5.0)
Missing n (%)	26,136 (11.4)	4,864 (4.2)	1,746 (4.7)
Placental abruption	3,794 (1.6)	2,095 (1.8)	652 (1.7)
Missing n (%)	0 (--)	0 (--)	0 (--)
Fetal growth restriction	2,678 (1.2)	1,451 (1.3)	373 (1.0)
Missing n (%)	15,626 (6.8)	9,747 (8.5)	3,905 (10.4)
Large for gestational age, antepartum diagnosis	2,031 (1.4)	1,073 (0.9)	257 (0.7)
Missing n (%)	84,338 (36.9)	36,337 (31.7)	15,227 (40.5)
Antenatal steroids	5,945 (4.0)	2,762 (2.4)	1,073 (2.9)
Missing n (%)	82,507 (36.1)	47,303 (41.2)	9,942 (26.5)
Antepartum bleeding in third trimester	3,088 (1.7)	1,836 (1.6)	569 (1.5)
Missing n (%)	49,410 (21.6)	36,293 (31.6)	9,846 (26.2)
Placenta previa	1,647 (0.7)	859 (0.8)	235 (0.6)
Missing n (%)	0 (--)	0 (--)	0 (--)
Placenta accreta	138 (0.0)	97 (0.1)	23 (0.1)
Missing n (%)	82,335 (36.0)	36,293 (31.6)	9,846 (26.2)
Threatened preterm birth antepartum	7,880 (3.4)	3,798 (3.3)	935 (2.5)
Missing n (%)	0 (--)	0 (--)	0 (--)
Antepartum hospital admission	17,500 (7.6)	7,424 (6.5)	1,625 (4.3)
Missing n (%)	84,603 (37.0)	55,696 (48.6)	17,264 (46.0)
Gestational age at delivery, weeks Mean (SD) Median (IQR)	 38.2 (2.47) 39 (38, 40)	 38.1 (2.53) 39 (38, 40)	 38.2 (2.41) 39 (38, 40)
Missing n (%)	0 (--)	0 (--)	0 (--)
Characteristics on L&D admission			
Temperature (F) Mean (SD)	 98.0 (0.81)	 98.0 (0.81)	 98.0 (0.79)
Missing n (%)	47,977 (21.0)	23,172 (20.2)	7,606 (20.3)
Systolic blood pressure Mean (SD)	 124.1 (14.89)	 124.1 (14.83)	 124.0 (14.9)

Missing n (%)	52,766 (23.0)	25,885 (22.6)	9,263 (23.0)
Diastolic blood pressure Mean (SD)	74.3 (11.6)	74.1 (11.9)	74.3 (11.6)
Missing n (%)	52,800 (23.1)	25,898 (22.6)	9,268 (24.7)
Trial of labor	192,074 (84.0)	96,533 (84.1)	31,636 (84.2)
Missing n (%)	0 (--)	0 (--)	0 (--)
Spontaneous labor	122,673 (53.7)	63,154 (55.1)	19,973 (53.2)
Missing n (%)	0 (--)	0 (--)	0 (--)
Premature rupture of membranes	16,219 (7.1)	9,369 (8.1)	3,017 (8.0)
Missing n (%)	0 (--)	0 (--)	0 (--)
Chorioamnionitis on admission	1,527 (1.0)	841 (0.8)	202 (0.7)
Missing n (%)	83,009 (36.3)	19,858 (19.8)	10,220 (32.9)
*N if 47 less than the final study cohort as 47 participants were missing year of delivery.			

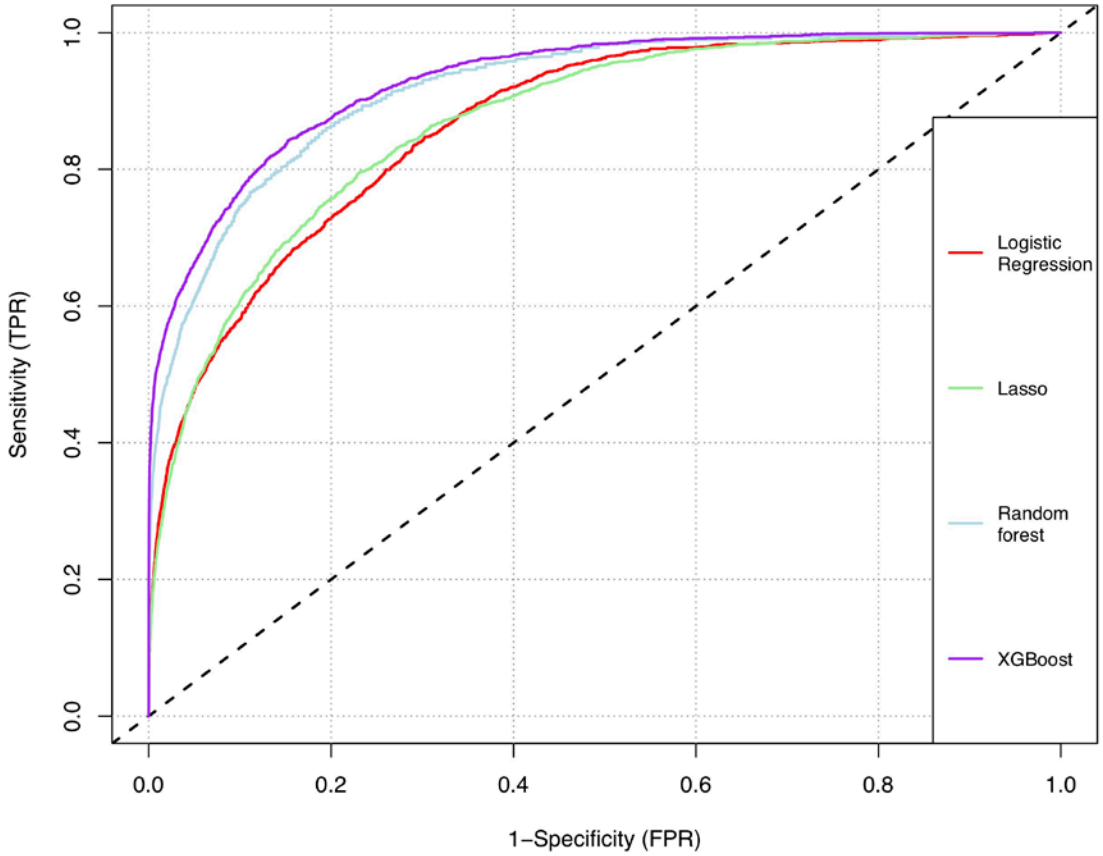
Appendix 4. Frequency of postpartum hemorrhage overall and by mode of delivery			
Overall	**Overall**	**Vaginal delivery**	**Cesarean delivery**
Mean EBL (SD), ml[1]	444.6 (2327.10)	315.0 (226.80)	769.9 (4,326.84)
Median EBL (IQR), ml[1]	350.0 (300, 600)	300.0 (250, 350)	700.0 (600, 800)
Primary outcome			
EBL ≥1,000 ml[1]			
N	7,279/152,279	654/108,872	6,625/43,407
% (95% CI)	4.7 (4.67 to 4.88)	0.6 (0.55 to 0.64)	15.2 (14.92 to 15.60)

[1]EBL data available from 9 of 12 sites: 152,279/228,438 (66.7%).
Abbreviations: EBL (estimated blood loss), SD (standard deviation), IQR (interquartile range), CI (confidence interval).

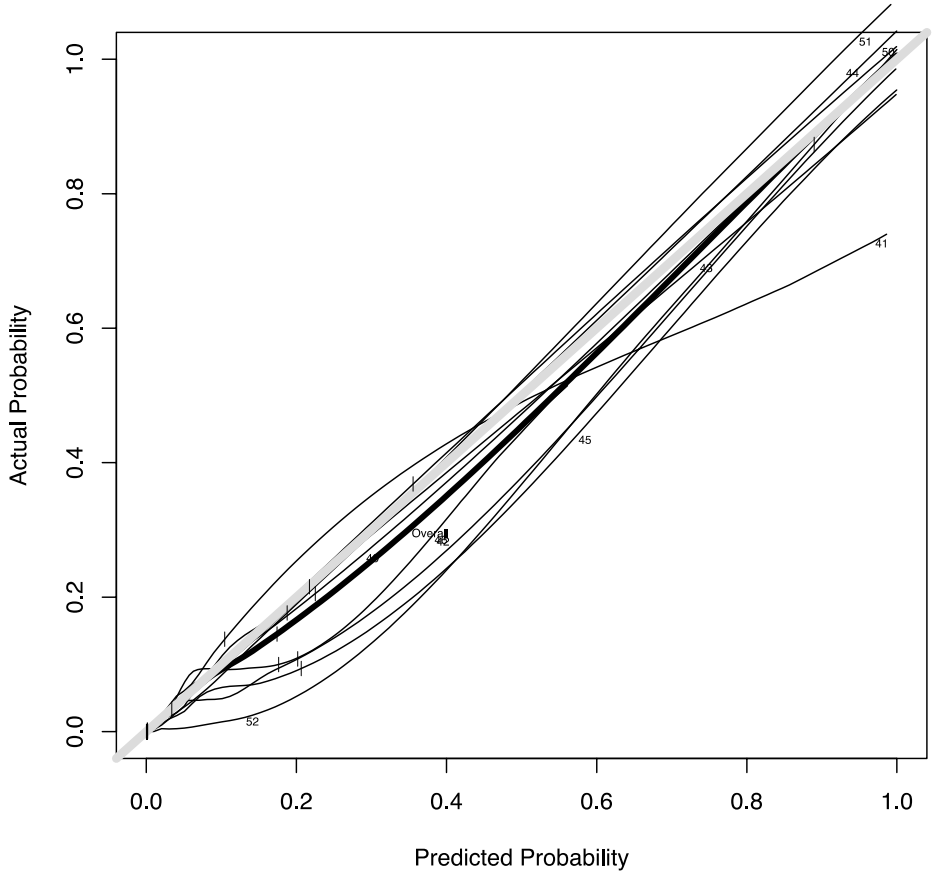
Appendix 5. Variables Included in Statistical and Machine Learning Models				
	2 statistical models		2 machine learning models	
Variable	**Lasso Regression**	**Logistic Regression**	**Extreme Gradient Boosted**	**Random Forest**
Age	X	X	X	X
Pre-pregnancy weight	X		X	X
Pre-pregnancy BMI		X	X	X
Admission weight		X	X	X
Chronic hypertension		X	X	X
Parity		X	X	X
Anemia	X		X	X
Assisted reproductive technology	X	X	X	X
Breech presentation	X	X	X	X
Fetal macrosomia	X	X	X	X
Preeclampsia without severe features	X	X	X	X
Placental abruption	X	X	X	X
Small for gestational age, antenatal diagnosis	X	X	X	X
Large for gestational age, antenatal diagnosis	X	X	X	X
Antepartum vaginal bleeding	X	X	X	X
Placenta accreta	X	X	X	X
Gestational age at delivery	X	X	X	X
Admission systolic blood pressure	X	X	X	X
Trial of labor	X	X	X	X
Chorioamnionitis on admission	X	X	X	X
Insurance status	X	X		X
Maternal race	X	X	X	X
Gestational diabetes	X		X	X
Seizure disorder	X	X	X	X
Multiple gestation	X		X	X
Polyhydramnios	X		X	X
History of prior cesarean delivery	X		X	X
Preeclampsia with severe features	X		X	X
Antenatal steroids	X	X	X	X
Placenta previa	X	X	X	X
Threatened preterm labor	X	X	X	X
Admission temperature	X		X	X

Admission diastolic blood pressure	X	X	X	X
Spontaneous labor	X	X	X	X
Education status	X		X	X
Prior antepartum hospitalization		X	X	X
Admission temperature		X		
Premature rupture of membranes		X	X	X
Marital status		X	X	X
Tobacco use			X	X
Drug use			X	X
Non-gestational diabetes			X	X
Gestational hypertension			X	X
Heart disease			X	X
Thyroid disease			X	X
Renal disease			X	X
Asthma			X	X
Depression			X	X
Gastrointestinal disease			X	X
History of seizures			X	X
History of prior preterm birth			X	X
Superimposed preeclampsia			X	
Ecclampsia			X	X
Fetal demise			X	X
Maternal GBS colonization			X	X
Magnesium sulfate			X	X

Appendix 6. Receiver operating characteristic curves of postpartum hemorrhage using two machine learning and two statistical models. Lasso, logistic regression with lasso regularization; XGBoost, extreme gradient boosting.



Appendix 7. Calibration curve demonstrating the Extreme Gradient Boosting model's performance of predicting postpartum hemorrhage by hospital site. The figure demonstrates the variation in gradient boosted model performance at each site in the cohort. The *grey line* indicates perfect agreement between the predicted probability of the model and the actual probability. Each *thin black line* labelled by a number is the calibration of the model in women from a single site. The *thicker black line* is the overall model calibration in the total cohort. The model overpredicts the actual risk at some sites when predicted probabilities are less than approximately 0.5. At site 41, the model underpredicts actual risk above 0.5.



Venkatesh KK, Strauss RA, Grotegut C, Heine RP, Chescheir NC, Stringer JSA, et al. Machine learning and statistical models to predict postpartum hemorrhage. Obstet Gynecol 2020;135.
The authors provided this information as a supplement to their article.
