Journal of the American Medical Informatics Association, 29(2), 2022, 321–328
doi: 10.1093/jamia/ocab181
Advance Access Publication Date: 24 September 2021
Research and Applications

Research and Applications

A comprehensive digital phenotype for postpartum hemorrhage

Amanda B. Zheutlin [1,*] Luciana Vieira,[2,*] Ryan A. Shewcraft,[1,*] Shilong Li,[1] Zichen Wang,[1] Emilio Schadt,[1] Yu-Han Kao,[1] Susan Gross,[1,3] Siobhan M. Dolan,[2,3] Joanne Stone,[2] Eric Schadt,[1,3,*] and Li Li [1,*]

[1]Sema4, Stamford, Connecticut, USA, [2]Division of Maternal Fetal Medicine, Department of Obstetrics, Gynecology, and Reproductive Science, Icahn School of Medicine at Mount Sinai, New York, New York, USA, and [3]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York, USA

*Equal contribution .

Corresponding Author: Li Li, MD, 333 Ludlow St., North Tower, 7th Floor, Stamford, CT 06902, USA; (li.li@sema4.com)

Received 25 February 2021; Revised 17 May 2021; Editorial Decision 4 August 2021; Accepted 11 August 2021

ABSTRACT

Objective: We aimed to establish a comprehensive digital phenotype for postpartum hemorrhage (PPH). Current guidelines rely primarily on estimates of blood loss, which can be inaccurate and biased and ignore complementary information readily available in electronic medical records (EMR). Inaccurate and incomplete phenotyping contributes to ongoing challenges in tracking PPH outcomes, developing more accurate risk assessments, and identifying novel interventions.

Materials and Methods: We constructed a cohort of 71 944 deliveries from the Mount Sinai Health System. Estimates of postpartum blood loss, shifts in hematocrit, administration of uterotonics, surgical interventions, and diagnostic codes were combined to identify PPH, retrospectively. Clinical features were extracted from EMRs and mapped to common data models for maximum interoperability across hospitals. Blinded chart review was done by a physician on a subset of PPH and non-PPH patients and performance was compared to alternate PPH phenotypes. PPH was defined as clinical diagnosis of postpartum hemorrhage documented in the patient's chart upon chart review.

Results: We identified 6639 PPH deliveries (9% prevalence) using our phenotype—more than 3 times as many as using blood loss alone (N = 1,747), supporting the need to incorporate other diagnostic and intervention data. Chart review revealed our phenotype had 89% accuracy and an F1-score of 0.92. Alternate phenotypes were less accurate, including a common blood loss-based definition (67%) and a previously published digital phenotype (74%).

Conclusion: We have developed a scalable, accurate, and valid digital phenotype that may be of significant use for tracking outcomes and ongoing clinical research to deliver better preventative interventions for PPH.

Key words: postpartum hemorrhage, digital phenotype, electronic medical records, maternal morbidity

INTRODUCTION

Postpartum hemorrhage (PPH) is a leading cause of maternal mortality in the United States (US).[1,2] The majority of these deaths are preventable, and a primary cause is error or delay in diagnosis and treatment.[2–6] Though mortality rates due to PPH have remained stable over the past 15 years,[7,8] the prevalence has increased consider-

ably,[9] ranging from 3% to 9% depending on the definition,[7,10–14] as well as the need for critical interventions to treat severe cases.[1,9,10,15,16] There is a critical need for optimization of preventative care and treatment modalities to reduce morbidity and mortality.[17]

Historically, postpartum hemorrhage has lacked a single, consistent definition and has relied heavily on visual estimates of blood loss, which can be inaccurate, biased, and unreliable.[6,17–22] More recently, in an effort to standardize clinical obstetric definitions, the American College of Obstetricians and Gynecologists (ACOG) developed the reVITALize program, which defines postpartum hemorrhage as a cumulative blood loss of greater than or equal 1000mL or blood loss accompanied by signs or symptoms of hypovolemia within 24 hours following delivery (including intrapartum blood loss).[23] Currently, there is no "gold standard" definition for postpartum hemorrhage that encompasses objective vital sign changes and clinical data.[24–27] Therefore, estimation of blood loss of at 1000 mL visually or quantitatively remains central to diagnosis and initiation of treatment.[4,6] The overreliance on estimated blood loss (EBL) alone has contributed to underestimation of hemorrhage[14,18,19,21,22,28] and this defines a key point for improvement in PPH prevention efforts .

Due to the limitations of visual or estimated blood loss, practice is shifting towards more objective measurements, such as quantitative blood loss (QBL), with obstetric hemorrhage toolkits like the California Maternal Quality Care Collaborative.[29] The use of QBL has not been widely adopted since it requires hospitals to have specialized equipment and training for providers.[6,20] Furthermore, the threshold of blood loss utilizing 1000mL is somewhat arbitrary and some women needing care for hemorrhage may ultimately lose less than 1000mL.[22]

To mitigate limitations of using blood loss alone to definite PPH, additional measures have been proposed to refine case definitions. A proxy measure of blood loss is change in hematocrit values, with a 10% drop indicating PPH,[1] although this may have low specificity and is affected by global changes in fluids like dehydration or any infusions.[20] Alternatively, PPH can also be indexed with diagnostic codes (which have low sensitivity[13]) or with indications of severe outcomes like blood transfusions or surgical interventions.[12,30] Some efforts combined multiple retrospective diagnostic codes with medications to manage uterine atony to improve detection of PPH.[12,20] Uterotonics, including oxytocin, carboprost tromethamine, misoprostol, and methylergonovine, are the first-line interventions for acute medical management of PPH,[1] so they may be useful markers for identifying PPH.

Here, we aimed to establish a physician-validated, comprehensive digital phenotype for PPH using information extracted from electronic medical records (EMR) in a large US health system. We used several sources to identify deliveries with significant blood loss, as well as deliveries where medical or surgical interventions for treating PPH were given. Through careful extraction of medication dosing and timing, mapping of fluctuations in lab values during labor and delivery, and synthesis of medical observations across labor and delivery admission, we aimed to develop a comprehensive phenotype to retrospectively identify deliveries with PPH. To validate this phenotype, blinded chart reviews by a physician were conducted on a subset of patients to confirm clinical diagnosis of PPH. Performance of our phenotype was then compared to both a common definition of PPH ($\geq$ 1000mL blood loss)[1] and the most comprehensive previously developed digital phenotype.[12] Inaccurate phenotyping remains a significant barrier to tracking incidence and management of PPH in hospitals, developing more accurate risk stratification tools, and identifying novel interventions. Our goal was to provide a robust digital phenotype that can be readily implemented retrospectively for both quality improvement initiatives and clinical research.

MATERIALS AND METHODS

Patient population

We used deidentified EMR data provided by the Mount Sinai Data Warehouse (MSDW) from the Mount Sinai Health System (MSHS), one of the largest and most comprehensive EMR systems in New York City. MSHS includes 5 member hospitals with EMR from 2000–2020, which draw from a racially and ethnically diverse patient population. Clinical variables including patient demographics, medical histories, or visit details were available for 9 million unique patients. These deidentified data were used to construct a delivery cohort and develop a digital phenotyping algorithm. A diagram of our workflow is presented in Figure 1A.

We received approval from the Icahn School of Medicine at Mount Sinai Institutional Review Board (IRB-17-01245) to conduct this study.

Delivery cohort

To identify all deliveries, we used 3 sources: (1) a standardized delivery summary completed by delivery staff on Labor & Delivery, (2) procedure records for vaginal or Cesarean deliveries (identified using CPT-4 and ICD-10-PCS billing codes), and (3) linked mother–infant hospital visit records time-stamped to the infant's day of birth. For all deliveries, we identified gestational weeks at delivery, delivery time, delivery method, parity, and hospital admission time. When gestational weeks at delivery was not recorded, it was estimated using gestational age reported for prenatal visits (admit reason). When delivery time was not available, we used the final procedure time stamp or the time stamp at which the mother received 5 or more units of oxytocin (a prophylactic dose given immediately after delivery of the anterior shoulder).[11] Delivery method was labeled using delivery procedure records and ICD-9-CM or ICD-10-CM diagnostic codes for delivery (Supplementary Table S1) given to either the mother or infant. Parity was estimated by assuming the first delivery for each woman was the earliest one included in our cohort and that all following deliveries were also at MSHS. Finally, hospital inpatient admission and unit transfer times were extracted to create admission–delivery journeys. Deliveries without any gestational age information or without admission time were excluded. We also limited the cohort to deliveries from January 1, 2011 through December 31, 2019 to ensure records were complete (prior to 2011, data availability through EMR is limited).

Clinical feature cleaning and normalization

All available demographic information, lab tests, vital signs, diagnoses, medications, and procedures were extracted for all women in our delivery cohort. We standardized these data by mapping native coding systems to common frameworks that are part of the Unified Medical Language System—a process that also increases interoperability between healthcare systems. All available observations were cleaned and normalized within data type.

For patient demographics, we extracted patient's age at delivery, race, ethnicity, and insurance carried during the current pregnancy journey. When there were inconsistencies within a patient's history of self-reported race or ethnicity, we assigned the most common self-

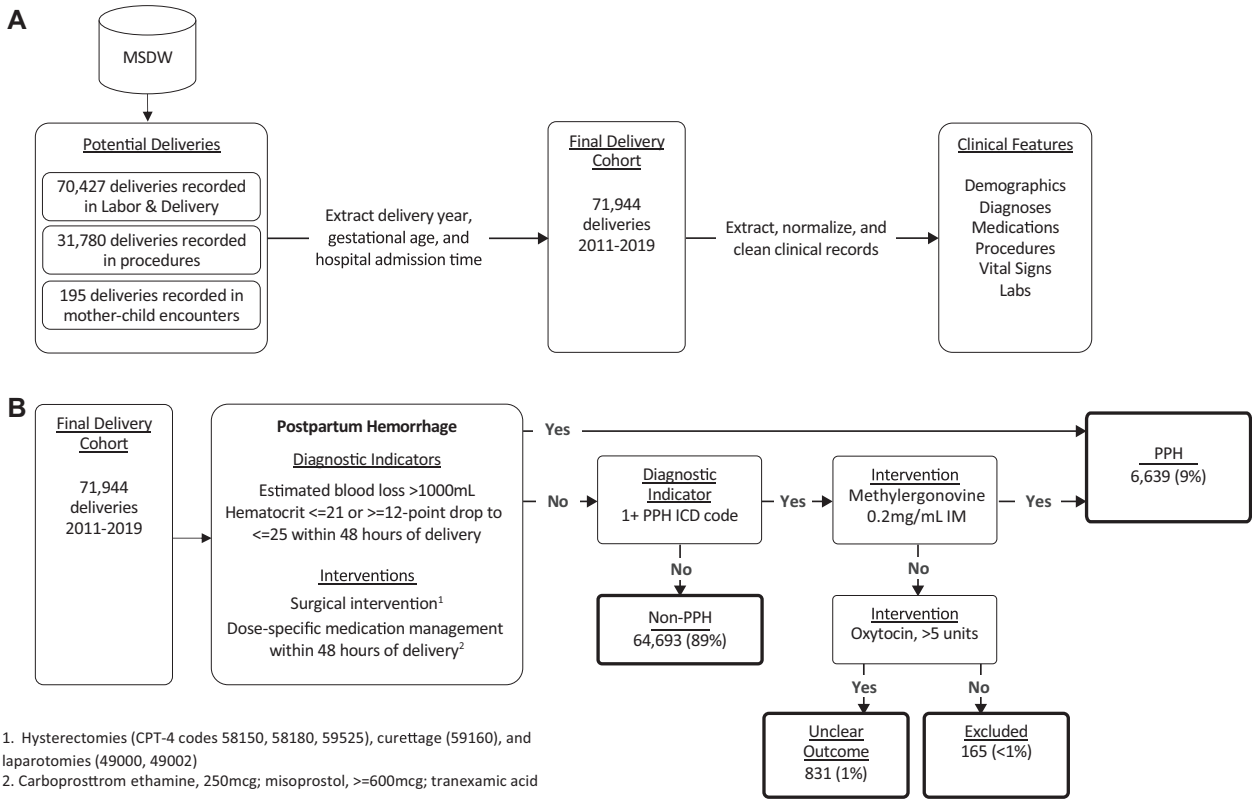

Figure 1. Workflow for data extraction (A) and digital phenotype (B).

report. Lab test names and units were mapped to logical observation identifiers names and codes (LOINC). Values were cleaned (invalid results and text removed) and converted to numeric values or standardized to non-numeric scales as appropriate. Duplicate results (eg, "preliminary" and "final" results from the same test with the same values) were filtered to retain the earliest result. Vital signs, including weight, height, temperature, respirations, pulse, diastolic blood pressure, and systolic blood pressure were standardized to common names and unit scales. Diagnoses from ICD-9-CM and ICD-10-CM were combined via mapping to broader categories using the Clinical Classifications Software.[31] We filtered medications to those administered to patients and mapped all medication names to RxNorm ingredients. Procedures were recorded through CompuRecord, an anesthesia information management program,[32] using CPT-4 codes. When procedures included multiple time points (eg, procedure start, anesthesia given, fluid given), only the earliest one was retained.

Clinical and obstetric characteristics

The latest hematocrit (LOINC 20570-8) test result and vital signs given within 48 hours prior to delivery were used as baseline measures. Oxytocin or misoprostol administered after hospital admission and prior to delivery was considered evidence of labor induction or augmentation. Sixty-one ICD-9/ICD-10 codes given during pregnancy or within 30 days post-delivery were used to index pregnancies with multiple gestation (Supplementary Table S2).

Digital phenotyping algorithm for PPH

We aimed to identify women who had significant blood loss, as well as those who had PPH-specific interventions. A diagram of our workflow is presented in Figure 1B.

Diagnostic indicators of PPH

We used multiple sources to detect substantial blood loss postpartum. First, we considered EBL or QBL by clinicians post-delivery to have exceeded 1000mL as evidence of PPH. MSHS adopted quantification practices in 2017; blood loss values were estimated prior to then. Since EBL is biased towards underreporting PPH,[6,18,19,23] and quantified blood loss is not always reliable,[12,33] we also included women with critically low hematocrit ($\leq$ 21) or a greater than 12-point drop from baseline—a proxy measure for blood loss[1]—that resulted in a minimum value at or below 25 within 48 hours of delivery. Finally, we included women given 1 or more of 30 diagnostic codes selected by a Maternal Fetal Medicine specialist as indicators of PPH (Supplementary Table S3). Since ICD codes can be assigned to a visit after care was given, we considered codes given on delivery day or within the following 14 days to reflect events during delivery. ICD diagnoses are often inaccurate when used on their own,[34] so we additionally required administration of any uterotonic medication except oxytocin (carboprost tromethamine, misoprostol, methylergonovine). Oxytocin was excluded due to its routine use in the active management of the third stage of labor.[1]

Interventions to prevent or treat PPH

We also identified deliveries where uterotonics or surgical interventions intra- or postpartum were given to prevent or treat maternal hemorrhage. Because methylergonovine can be given prophylactically or as treatment, we included women given methylergonovine intramuscularly only if they were also given an ICD code for PPH (as described above). If 1 first-line uterotonic does not sufficiently control bleeding, acute medical management of PPH requires combination use of misoprostol, carboprost tromethamine, and/or tra-

nexamic acid (an antifibrinolytic drug that promotes blood clotting).[1] As such, we included women administered 250mcg of carboprost tromethamine, ≥ 600mcg of misoprostol, or any dose of tranexamic acid within 48 hours of delivery.[1,11]

When hemorrhage continues despite medical management, Bakri balloon placement and surgical interventions may be required. This can include procedures occurring during laparotomies (CPT-4 codes 49000, 49002) like placement of compression sutures and uterine artery ligation or embolization, curettage (59160), or, typically as a last result, hysterectomies (58150, 58180, 59525).[1,11,35] We included all women that underwent any of these procedures postpartum.

Deliveries with unclear outcomes and non-PPH deliveries

There were 831 deliveries given 1 or more of 30 PPH ICD codes with a postpartum dose of oxytocin (10 or more units) and had no other indication of PPH. Since this is a plausible scenario for PPH, and we wanted to maximize our detection of true cases, we selected some of these charts for review, but labeled them as "deliveries with unclear outcomes" (DUO). Deliveries with an ICD code, but no other indication were excluded (N = 165); all remaining deliveries neither classified as a PPH delivery nor a DUO were classified as non-PPH.

Descriptive statistics

We used univariate logistic regression to assess statistical differences between PPH and non-PPH deliveries in age and gestational weeks at delivery, parity, baseline labs and vital signs, and hours from admission to delivery. Proportional differences in race, ethnicity, insurance, delivery method, labor induction or augmentation, and multiple gestation by PPH status were assessed using chi-square tests for independence. EBL/QBL differences were assessed separately for vaginal and Cesarean deliveries. We used a Bonferroni correction to conservatively control for multiple comparisons. With this adjustment, the significance threshold was set to alpha < 0.001, two-tailed. All statistical analyses were done using the stats package in R (version 3.6.0).[36]

Chart review

To assess accuracy of our digital phenotyping algorithm, an obstetrician with access to fully identified patient charts including labor and delivery flow charts, progress notes, nursing notes, delivery summary, and discharge summaries conducted manual chart review to ascertain presence or absence of a diagnosis of PPH in clinical documentation while blinded to the label assigned to each chart using the digital phenotyping algorithm. Therefore, the PPH definition was clinical documentation within the medical record. We randomly selected 45 charts consisting of PPH (N = 26), non-PPH (N = 11), and DUO (N = 6) defined by our digital phenotyping algorithm. Within PPH phenotype, the chart selection covers each rule that we used to identify phenotype to ensure representation including EBL/QBL (N = 7), ICDs and methylergonovine (N = 3), hematocrit drop (N = 6), intervention management (N = 4), and more than 1 rule (N = 6) (Figure 1, Table 2).

We sought to review a sufficient number of charts such that the 95% confidence interval (CI) of the accuracy estimate exceeded 80%. We generated a bootstrap estimate of the lower bound of the 95% CI as a function of sample size by calculating the accuracy for a randomly selected (with replacement) sub- or superset of the digital phenotype-clinical diagnosis pairs established by chart review. We performed 1000 permutations to generate an empirical distribution for chart review accuracy and used this to estimate the 95% CI.

Our first aim for this review was to verify that our data-mining techniques were accurate. Our data was deidentified to ensure patient privacy, so it did not include all information recorded for a delivery, including any information from physician or surgical notes. Thus, we wanted to verify data accuracy by comparing information extracted from EMR including lab results, medications, procedures, delivery time and method, and blood loss values to what was available in native clinical charts accessible by physicians. We summarized data accuracy by overall rates of exact matching values between our data and chart review.

Our second aim was to assess the validity of our digital phenotyping algorithm. Charts were reviewed for evidence of significant blood loss as indicated in notes or lab tests, explicit indication of PPH in delivery summary or visit notes, or evidence of interventions specifically for managing PPH clearly beyond standard care. A judgement for each chart of "yes," "no," or "unclear" was made based on this evidence. We summarized the performance of our digital phenotype by calculating specificity, sensitivity, positive and negative precision, accuracy, and F1 score for our labels compared to chart review labels.

For comparison, we also calculated these same metrics for each criterion individually (ie, by using presence or absence of that criterion as the PPH label and comparing it to chart review labels) in order to assess the value of combining criteria relative to each indicator on its own. Finally, to compare our digital phenotype with alternate phenotypes, we generated PPH labels based on the EBL criterion from the ACOG guidelines[1] (PPH defined as any delivery with EBL/QBL ≥ 1000mL) and the most comprehensive previously published digital phenotype, which was not evaluated with chart review in the original report.[12] ACOG additionally defines any blood loss followed by signs or symptoms of hypovolemia to be PPH, however, this criterion is difficult to ascertain from EHR data without discrete definitions. Because there is significant variability in the clinical and vital sign changes that are associated with blood loss, there are no established cutoff points to trigger clinical interventions.[37,38] The latter phenotype was originally proposed as 4 mutually exclusive levels of risk, which we have combined here into a single phenotype for simplicity. PPH deliveries were defined as having any of the following criteria: administration of any uterotonic (except oxytocin), 1 or more of 12 ICD-9 or ICD-10 codes for PPH[12] (Supplementary Table S3), transfusion of blood intra- or postpartum, receipt of intrauterine tamponade device, or hysterectomy. All patient charts were used for comparison performance metrics, regardless of their label using our digital phenotype.

RESULTS

Demographic, clinical, and obstetric characteristics for the pregnancy cohort

We identified 73 025 deliveries occurring between January 1, 2011 and December 31, 2019. We excluded 1081 deliveries in which a hospital admission time could not be identified, leaving 71 944 deliveries from 57 151 mothers in our final cohort. Summary statistics for demographic, clinical, and obstetric characteristics for the entire cohort, as well as for PPH and non-PPH deliveries, are provided in Table 1.

Table 1. Demographic, clinical, and obstetric characteristics for pregnancy cohort

	Delivery cohort N (%) | Mean±SD	PPH N (%) | Mean±SD	Non-PPH N (%) | Mean±SD
Demographics			
Number of patients	71 944 (100%)	6639 (9%)	64 309 (89%)
Age, years +	32 ± 6	33 ± 6	32 ± 6
Race +			
White	40 542 (56%)	3176 (48%)	36 801 (57%)
African American	7418 (10%)	911 (14%)	6407 (10%)
Asian	5819 (8%)	622 (9%)	5106 (8%)
Native American	284 (<1%)	25 (<1%)	253 (<1%)
Other	13 433 (19%)	1495 (22%)	11 761 (18%)
Unknown	4448 (6%)	410 (6%)	3981 (6%)
Ethnicity +			
Non-Hispanic	40 693 (57%)	3444 (55%)	36 629 (57%)
Hispanic	11 470 (16%)	1269 (19%)	10 044 (16%)
Unknown	19 825 (28%)	15 686 (25%)	17 891 (28%)
Insurance +			
Private	42 073 (59%)	3633 (55%)	37 810 (59%)
Medicaid or Medicare	20 827 (32%)	20 474 (37%)	23 644 (33%)
Uninsured	418 (<1%)	46 (1%)	471 (1%)
Other or missing	5254 (8%)	486 (7%)	5756 (8%)
Clinical baseline (last measurement prior to delivery)			
Body-mass index, kg/m^2 +	29 ± 5	30 ± 6	29 ± 5
DBP, mmHg +	73 ± 11	75 ± 12	72 ± 11
SBP, mmHg +	121 ± 14	125 ± 16	121 ± 14
Hematocrit, % +	36 ± 3	35 ± 4	36 ± 3
Obstetric characteristics			
Cesarean delivery +	25 434 (35%)	3132 (47%)	21 942 (34%)
Admission to delivery, hr +	9.5 ± 7.3	11.5 ± 8.3	9.3 ± 7.2
Gestational wks at delivery +	39 ± 2	38 ± 3	39 ± 2
Labor induction or augmentation +	47 697 (66%)	4783 (72%)	42 248 (66%)
Multiple gestation +	4122 (6%)	661 (10%)	3388 (5%)
Parity +	1.5 ± 0.6	1.4 ± 0.9	1.5 ± 0.9
EBL/QBL, mL			
Vaginal delivery +	322 ± 180	581 ± 414	297 ± 116
Cesarean delivery +	730 ± 342	1080 ± 606	677 ± 243

+Significant difference between PPH and non-PPH deliveries, $P < .001$.

Digital phenotype for PPH

We classified 6639 deliveries (9% prevalence) as PPH deliveries using diagnostic indicators of significant blood loss combined with any evidence of PPH-specific interventions. Frequencies for each criterion we considered for inclusion, as well as overall PPH prevalence are listed in Table 2. We also classified PPH deliveries based on their indication source (Figure 1B). We found that 71% of PPH deliveries had evidence from 1 source of data (Medications: 28%, Hematocrit: 16%, ICDs: 14%, EBL: 13%, Surgeries: <1%), including those with multiple indicators within a source (eg, a woman given carboprost and misoprostol, but no nonpharmaceutical interventions), while 29% of PPH deliveries had more than 1 type of indication in their medical record (eg, at least a 12-point drop in hematocrit to at or below 25 and EBL >1000 mL).

Validation of digital phenotype for PPH

We selected 45 charts to review for data accuracy and digital phenotype validation. Two charts were excluded due to restricted access, resulting in 43 charts for review. Delivery method labels, delivery times, estimates of blood loss, baseline and follow-up hematocrit lab values, and uterotonics administration exactly matched those found in chart review, with the exception of 1 missing medication for 2 deliveries. For these deliveries, methylergonovine was listed only in the patient's delivery summary, but not in the patient's medication records, so these events were not found in the structured data extracted from EMR.

We also evaluated the accuracy of our PPH, non-PPH, and DUO labels relative to chart review. Among deliveries labeled as having PPH or non-PPH by our digital phenotype (N = 37), 24 had PPH, and 13 did not according to physician review. One true PPH case and 3 non-PPH cases were misclassified by our phenotype algorithm, yielding 89% accuracy (Table 3) (95% CI lower bound = 81.6%; bootstrap method; Supplementary Figure S1) and an F1-score of 0.92. The overlap between the digital phenotype and chart review definitions of PPH status was statistically significant ($P = 8.1\text{x}10^{-6}$, odds ratio = 76.7; Fisher's exact test). Among patients with unclear outcomes (DUO group), 3 had PPH, and 3 did not. Overall performance and accuracy by PPH criterion were detailed in Table 4. We also compared accuracy of our digital phenotype to labels using only 1 of the criteria we considered (rather than combining them), and 2 alternate phenotypes: ACOG's EBL PPH guideline, and the most comprehensive EHR-based digital phenotype previously proposed. All were less accurate than our digital phenotype (Table 3).

Table 2. Frequencies of clinical features used to assess medically actionable risk for PPH

Criteria	PPH[a]	N	Percent[b]
EBL/QBL			
Any record		64 240	89%
>1000 mL	X	1747	2%
Hematocrit			
Baseline measurement		68 807	96%
Any follow-up measurement		29 899	42%
≤ 21 or ≥ 12-point drop from baseline to ≤ 25	X	1949	3%
Billing codes			
PPH ICD code		4219	6%
PPH ICD code + oxytocin, ≥ 10 units	DUO	3829	5%
PPH ICD code + methylergonovine, 0.2mg/mL IM	X	2060	3%
Medication management			
Oxytocin, ≥ 10 units		61 476	85%
Methylergonovine, 0.2mg/mL IM		8094	11%
Carboprost tromethamine, 250 mcg	X	1806	3%
Misoprostol, ≥ 600mcg	X	2000	3%
Tranexamic acid	X	111	<1%
Surgical interventions			
Laparotomy	X	18	<1%
Curettage	X	51	<1%
Hysterectomy	X	74	<1%
Any PPH indication	X	6639	9%

[a]Indicates delivery inclusion in digital phenotype (x) or deliveries with unclear outcome (DUO).

[b]As a percent of entire cohort N = 71 944.

DISCUSSION

We have developed a comprehensive digital phenotype for PPH using a large and diverse EMR system from New York City. Chart review confirmed its validity with 96% sensitivity, 77% specificity, and 89% precision. Considering this performance relative to any individual criterion (accuracy ranging from 40%–74%), the most comprehensive previously proposed digital phenotype (74% accuracy), and a common definition based on the EBL criterion from ACOG's guidelines (67% accuracy), our digital phenotype affords the highest accuracy (89%). Concretely, this increase in accuracy allows us to identify more than 3 times as many PPH deliveries (N = 6639) than would be identified using blood loss alone (N = 1747). While additional clinical criteria may be considered by a clinician when they are determining patient care, for retrospective PPH outcomes assessments and research, EBL is often the objective measure that is consistently documented. Our digital phenotype suggests that this approach may substantially underestimate the incidence.

We also confirmed the data we extracted were highly consistent with clinical notes, highlighting the reliability of our approach. To increase accessibility across healthcare systems, our digital phenotype used only structured data mapped to common data models and did not require the use of advanced methods to mine notes (eg, natural language processing) or individual chart review to extract data, both of which are variable and time-consuming, but are commonly included in phenotyping algorithms.[34,39] Together, we suggest this digital phenotype is a scalable, accurate, and valid research tool that could be used to improve tracking of PPH incidence and management, as well as facilitate research to enhance risk assessment and intervention.

Interestingly, we found that overlap between PPH-inclusion criteria was only moderate. While 29% of PPH deliveries had more than 1 indication (eg, medication and blood loss >1000 mL), most had indications from only 1 category, underlining the need to incorporate multiple sources of information for identification of PPH deliveries (Figure 1B). Our most accurate categories on their own were EBL/QBL and hematocrit, which were both 100% accurate (equally accurate as having more than 1 indication) (Tables 3 and 4). Though drop in hematocrit has been noted to have low specificity[20] and can also reflect changes in volume status, our definition for hematocrit drop was more stringent. Additionally, we included low hematocrit (≤ 21), which can indicate blood loss anemia related to delivery, which is supported by our data. EBL measures are known to be biased towards underreporting, which is consistent with our findings.[6,18,19,22] Deliveries where blood loss was estimated to be >1000 mL were highly likely to be PPH deliveries, but deliveries with estimated blood loss at or below 1000 mL were not always non-PPH deliveries (Tables 3 and 4). One metric for capturing PPH deliveries with blood loss estimates less than 1000 mL may be to use a conservative threshold for hematocrit (12-point drop to ≤ 25). While hematocrit measures can reflect other changes besides blood loss (eg, administration of intravenous fluids or blood transfusions),[15] we found this threshold to be a good discriminator of PPH deliveries from non-PPH deliveries. Finally, we found that the accu-

Table 3. Performance for digital phenotype, individual criteria, and alternate phenotypes

	Sensitivity (Recall)	Specificity	Precision (PPV)	NPV	Accuracy
Digital phenotype					
PPH vs non-PPH	96%	77%	88%	91%	89%
PPH and DUO vs non-PPH	96%	63%	81%	91%	84%
PPH vs non-PPH and DUO	85%	81%	89%	76%	84%
Phenotype criterion alone					
EBL/QBL >1000mL	44%	100%	100%	48%	65%
Hematocrit ≤ 21 or 12-point drop to ≤ 25	30%	100%	100%	46%	56%
ICD	74%	75%	83%	63%	74%
Medication (carboprost, misoprostol, tranexamic acid)	19%	88%	71%	39%	44%
Procedure	4%	100%	100%	38%	40%
Alternate phenotype					
Goffman et al, levels 1–4	85%	56%	77%	69%	74%
ACOG EBL	59%	81%	84%	54%	67%

Table 4. Digital phenotype performance by PPH criterion

Digital phenotype			Physician label from chart review		
Label	N	Criterion	Case	Control	Unclear
PPH	26	All indications	23	3	0
	7	EBL/QBL	7	0	0
	3	ICD + methylergonovine	2	1	0
	6	Hematocrit drop or critically low	6	0	0
	4	Medication management	2	2	0
	6	More than 1 indication	6	0	0
Non-PPH	11	No case indications	1	10	0
PPH and Non-PPH	37	All PPH and non-PPH	24	13	0
DUO	6	ICD + oxytocin	3	3	0

racy of diagnostic billing codes depended on the context. While they were frequently assigned correctly (83% precision; Table 3), precision dropped to 67% with methylergonovine only and to 50% when paired exclusively with oxytocin (Table 4).

This phenotype should be considered in the context of several limitations. We reported 3 false positives and 1 false negative in chart review. Of the false positives, 2 of the 3 were deliveries where PPH treatment was applied prophylactically, despite no evidence of hemorrhage, because they had significant risk factors at hospital admission. Considering this, the use of our phenotype may be best suited for identifying women who needed interventions for PPH (using this as our definition, precision would be 96%), although its precision for PPH is still high (89%; Table 3). In general, the use of medications for uterine atony relies on visual cues and clinical judgement which are not uniform and may vary by clinician, adding another limitation.

The false negative was a delivery with an unanticipated surgical complication, which is a less common cause of PPH (and 1 not treated with uterotonics), and whose hematocrit values narrowly missed our threshold. While 70%–80% of PPH cases are caused by uterine atony, uterine trauma (eg, lacerations or other tissue tears), retained tissue (eg, invasive placenta), and acquired or chronic coagulopathies can also cause PPH.[1,11] It is possible that information pertaining to these causes is more readily available through physician notes or blood bank records, which we could not access with deidentified data. However, in general, these causes are less likely to be preventable with improved risk prediction than uterine atony—eg, intraoperative complications can be unforeseen and thus harder to prevent, whereas atony can be targeted prophylactically with uterotonics—so, again, this suggests the phenotype may be best suited for identifying women likely to benefit from interventions for PPH.

The benefit of a comprehensive definition that encompasses both mild and severe cases is the early identification and time to intervene prior to further decompensation, particularly since many women without risk factors experience PPH.[40,41] Early identification of clinically actionable postpartum hemorrhage would also allow for further planning and allocation of resources that may prevent further clinical decompensation and adverse maternal outcomes.[41]

Additionally, accurate phenotyping is a cornerstone of high caliber clinical research and hospital quality improvement. Here, we offer a robust, portable, physician-validated digital phenotype for PPH that captures more than 3 times as many deliveries as the most commonly used approach by leveraging a suite of complementary information available in EMR. This research tool may be of significant use in designing patient safety initiatives in addition to ongoing clinical research to deliver better preventative interventions for the leading cause of maternal morbidity worldwide.[42]

FUNDING

This project was performed in collaboration with Sema4. Sema4 is a company that integrates genetic testing and data analytics to improve diagnosis, treatment, and prevention of disease. The Icahn School of Medicine at Mount Sinai holds equity in this for-profit company.

AUTHOR CONTRIBUTIONS

Concept and design: LL and ErS. Data acquisition, cleaning, and interpretation of data: ABZ, RAS, SL, ZW, and EmS. Data analysis: ABZ, RAS. Drafting of the manuscript: ABZ, LV, RAS, and LL. Critical revision of the manuscript for important intellectual content: ABZ, LV, RAS, SL, ZW, EmS, SG, SMD, JS, ErS and LL. Supervision: LL and ErS.

SUPPLEMENTARY MATERIAL

Supplementary material is available *at Journal of the American Medical Informatics Association* online.

ACKNOWLEDGMENTS

We would like to thank the Mount Sinai Data Warehouse physician team for validating data accuracy and facilitating the chart review process. We would also thank the Sema4 IT team for infrastructural and computational support.

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY STATEMENT

Data cannot be shared for ethical/privacy reasons.

REFERENCES

1. Shields LE, Goffman DC; Committee on Practice Bulletins-Obstetrics. Practice Bulletin No. 183: postpartum hemorrhage. *Obstet Gynecol* 2017; 130 (4): e168–86.

2. Building U.S. Capacity to review and prevent maternal deaths. 2018. http://reviewtoaction.org/2018_Report_from_MMRCs. Accessed December 12, 2020.
3. Della Torre M, Kilpatrick S, Hibbard J, *et al.* Assessing preventability for obstetric hemorrhage. *Am J Perinatol* 2011; 28 (10): 753–60.
4. Seacrist MJ, Morton CH, VanOtterloo LR, *et al.* Quality improvement opportunities identified through case review of pregnancy-related deaths from sepsis. *J Obstet Gynecol Neonatal Nurs* 2019; 48 (3): 311–20.
5. Bingham D, Jones R. Maternal death from obstetric hemorrhage. *J Obstet Gynecol Neonatal Nurs* 2012; 41 (4): 531–9.
6. Quantitative blood loss in obstetric hemorrhage. *Obstet Gynecol* 2019; 134 (6): 1368–9.
7. Marshall AL, Durani U, Bartley A, *et al.* The impact of postpartum hemorrhage on hospital length of stay and inpatient mortality: a national inpatient sample–based analysis. *Am J Obstet Gynecol* 2017; 217 (3): 344.e1–e6.
8. Creanga AA, Syverson C, Seed K, *et al.* Pregnancy-related mortality in the United States, 2011–2013. *Obstet Gynecol* 2017; 130 (2): 366–73.
9. Data on Selected Pregnancy Complications in the United States. Atlanta, GA: Centers Disease Control and Prevention. https://www.cdc.gov/reproductivehealth/maternalinfanthealth/pregnancy-complications-data.htm. Accessed December 10, 2020.
10. Mehrabadi A, Hutcheon J, Lee L, *et al.* Epidemiological investigation of a temporal increase in atonic postpartum haemorrhage: a population-based retrospective cohort study. *BJOG* 2013; 120 (7): 853–62.
11. Evensen A, Anderson JM, Fontaine P. Postpartum hemorrhage: prevention and treatment. *Am Fam Physician* 2017; 95 (7): 442–9.
12. Goffman D, Friedman AM, Sheen J-J, *et al.* A framework for improving characterization of obstetric hemorrhage using informatics data. *Obstet Gynecol* 2019; 134 (6): 1317–25.
13. Butwick AJ, Walsh EM, Kuzniewicz M, *et al.* Accuracy of international classification of diseases, ninth revision, codes for postpartum hemorrhage among women undergoing cesarean delivery. *Transfusion* 2018; 58 (4): 998–1005.
14. Bell SF, Watkins A, John M, *et al.* Incidence of postpartum haemorrhage defined by quantitative blood loss measurement: a national cohort. *BMC Pregnancy Childbirth* 2020; 20 (1): 271.
15. Bateman BT, Berman MF, Riley LE, *et al.* The epidemiology of postpartum hemorrhage in a large, nationwide sample of deliveries. *Anesth Analg* 2010; 110: 1368–73.
16. Knight M, Callaghan WM, Berg C, *et al.* Trends in postpartum hemorrhage in high resource countries: a review and recommendations from the International Postpartum Hemorrhage Collaborative Group. *BMC Pregnancy Childbirth* 2009; 9: 55.
17. Andrikopoulou M, D'Alton ME. Postpartum hemorrhage: early identification challenges. *Semin Perinatol* 2019; 43 (1): 11–7.
18. Conner SN, Tuuli MG, Colvin R, Shanks AL, Macones GA, Cahill AG. Accuracy of estimated blood loss in predicting need for transfusion after delivery. *Am J Perinatol* 2015; 32 (13): 1225–30.
19. Stafford I, Dildy GA, Clark SL, Belfort MA. Visually estimated and calculated blood loss in vaginal and cesarean delivery. *Am J Obstet Gynecol* 2008; 199 (5): 519.e1–e7.
20. Oberhardt M, Friedman AM, Perotte R. A principled framework for phenotyping postpartum hemorrhage across multiple levels of severity. In: *AMIA. Annual Symposium Proceedings AMIA Symposium*; Washington, DC; 2019; 2019: 691–8.http://www.ncbi.nlm.nih.gov/pubmed/32308864.
21. Hancock A, Weeks AD, Lavender DT. Is accurate and reliable blood loss estimation the 'crucial step' in early detection of postpartum haemorrhage: an integrative review of the literature. *BMC Pregnancy Childbirth* 2015; 15: 230.
22. Kerr R, Weeks A. Postpartum haemorrhage: a single definition is no longer enough. *BJOG* 2017; 124 (5): 723–6.
23. Menard MK, Main EK, Currigan SM. Executive summary of the reVITALize initiative. *Obstet Gynecol* 2014; 124 (1): 150–3.
24. Venkatesh KK, Strauss RA, Grotegut CA, *et al.* Machine learning and statistical models to predict postpartum hemorrhage. *Obstet Gynecol* 2020; 135 (4): 935–44.
25. Kerr R, Eckert LO, Winikoff B, *et al.*; Brighton Collaboration Primary Postpartum Haemorrhage Working Group. Postpartum haemorrhage: Case definition and guidelines for data collection, analysis, and presentation of immunization safety data. *Vaccine* 2016; 34 (49): 6102–9.
26. Mavrides E, Allard S, Chandraharan E, *et al.*; on behalf of the RC of O and G. Prevention and management of postpartum haemorrhage. *BJOG* 2017; 124 (5): e106–49.
27. Maternal WHO, Unit CH, Initiative SM. *The Prevention and Management of Postpartum Haemorrhage: Report of a Technical Working Group*. Geneva, Switzerland; July 3–6, 1989.
28. Larsson C, Saltvedt S, Wiklund I, *et al.* Estimation of blood loss after cesarean section and vaginal delivery has low validity with a tendency to exaggeration. *Acta Obstet Gynecol Scand* 2006; 85 (12): 1448–52.
29. Lyndon A, Lagrew D, Shields L, *et al. Improving Health Care Response to Obstetric Hemorrhage Version 2.0*. California Maternal Quality Care Collaborative; 2015.
30. Kawakita T, Mokhtari N, Huang JC, *et al.* Evaluation of risk-assessment tools for severe postpartum hemorrhage in women undergoing cesarean delivery. *Obstet Gynecol* 2019; 134 (6): 1308–16.
31. Cowen ME, Dusseau DJ, Toth BG, *et al.* Casemix adjustment of managed care claims data using the clinical classification for health policy research method. *Med Care* 1998; 36 (7): 1108–13.
32. Martin J, Ederle D, Milewski P. [CompuRecord–A perioperative information management-system for anesthesia]. *Anasthesiol Intensivmed Notfallmed Schmerzther* 2002; 37 (8): 488–91.
33. Bingham D, Scheich B, Bateman BT. Structure, process, and outcome data of AWHONN's postpartum hemorrhage quality improvement project. *J Obstet Gynecol Neonatal Nurs* 2018; 47 (5): 707–18.
34. Liao KP, Sun J, Cai TA, *et al.* High-throughput multimodal automated phenotyping (MAP) with application to PheWAS. *J Am Med Inform Assoc* 2019; 26 (11): 1255–62.
35. Chandraharan E, Arulkumaran S. Surgical aspects of postpartum haemorrhage. *Best Pract Res Clin Obstet Gynaecol* 2008; 22 (6): 1089–102.
36. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2021.
37. Borovac-Pinheiro A, Pacagnella RC, Cecatti JG, *et al.* Postpartum hemorrhage: new insights for definition and diagnosis. *Am J Obstet Gynecol* 2018; 219 (2): 162–8.
38. Pacagnella RC, Souza JP, Durocher J, *et al.* A systematic review of the relationship between blood loss and clinical signs. *PLoS One* 2013; 8 (3): e57594.
39. Kirby JC, Speltz P, Rasmussen LV, *et al.* PheKB: a catalog and workflow for creating electronic phenotype algorithms for transportability. *J Am Med Inform Assoc* 2016; 23 (6): 1046–52.
40. Giannella L, Mfuta K, Pedroni D, *et al.* Delays in the delivery room of a primary maternity unit: a retrospective analysis of obstetric outcomes. *J Matern Fetal Neonatal Med* 2013; 26 (6): 593–7.
41. Dilla AJ, Waters JH, Yazer MH. Clinical validation of risk stratification criteria for peripartum hemorrhage. *Obstet Gynecol* 2013; 122 (1): 120–6.
42. Geller SE, Koch AR, Garland CE, *et al.* A global view of severe maternal morbidity: moving beyond maternal mortality. *Reprod Health* 2018; 15 (Suppl 1): 98.