Automated early detection of obstetric complications: theoretic and methodologic considerations

Gabriel J. Escobar, MD; Neeru R. Gupta, MD; Eileen M. Walsh, RN, MPH; Lauren Soltesz, BS; Stephanie M. Terry, MD; Patricia Kipnis, PhD

Compared with adults who are admitted to general medical-surgical wards, women who are admitted to labor and delivery services are at much lower risk of experiencing unexpected critical illness. Nonetheless, critical illness and other complications that put either the mother or fetus at risk do occur. One potential approach to prevention is to use automated early warning systems, such as those used for nonpregnant adults. Predictive models that use data extracted in real time from electronic records constitute the cornerstone of such systems. This article addresses several issues that are involved in the development of such predictive models: specification of temporal characteristics, choice of denominator, selection of outcomes for model calibration, potential uses of existing adult severity of illness scores, approaches to data processing, statistical considerations, validation, and options for instantiation. These have not been addressed explicitly in the obstetrics literature, which has focused on the use of manually assigned scores. In addition, this article provides some results from work in progress to develop 2 obstetric predictive models with the use of data from 262,071 women who were admitted to a labor and delivery service at 15 Kaiser Permanente Northern California hospitals between 2010 and 2017.

Key words: early warning system, electronic medical record, obstetrics, predictive model, severity of illness

Most women in the labor and delivery (L&D) service are healthy. Compared with adults in general medical-surgical wards, they are younger, and their risk of death is extremely low. However, maternal death in developed nations persists and is rising in the United States.[1,2] Further, obstetric complications are not rare, which has prompted considerable interest in the use of obstetric early warning scores.[3–6] Some institutions have developed predictive models that can be used for automated early warning systems (EWSs) for adults at risk of deterioration and/or cardiac arrest outside intensive care units.[7–10] The evidence base that EWSs of any kind actually improve outcomes is limited; no study has shown conclusive outcomes improvements,[9,10] and some have shown no benefit.[11,12] However, work on these systems continues. The notion that automated EWSs based on electronic medical records (EMRs) could be used in L&D has started to be discussed in the research and commercial literature.[13]

A number of early warning tools for obstetrics (the Maternal Early Warning Criteria, Modified Early Obstetric Warning System, and the Maternal Early Warning Trigger) have been described in recent literature.[3–6] These tools use a small set of discrete physiologic measurements (abnormal maternal temperature, oximetry, heart rate, respiratory rate, and mental status; abnormal fetal heart rate) as triggers for more focused evaluation. Cutoffs for the triggers were defined by expert opinion, as opposed to being calibrated against defined populations with known outcomes rates. Consequently, their statistical performance characteristics (sensitivity, specificity, number needed to evaluate to detect 1 outcome) are not known, so it is not possible to address the problem of alert fatigue. These triggers are designed for either manual use (which could be associated with inconsistent documentation) or as simple electronic alarms (issue alert if heart rate is <50 or >110 beats/min). Importantly, they do not include 2 indicators of the progress of labor: cervical dilation and station of fetal descent. They cannot capture subtle relationships (eg, the rate of change of a physiologic parameter) or important interactions, such as the dynamic relationship between heart rate and systolic blood pressure or between cervical dilation and station of fetal descent. Finally, these tools do not capture differential risk because of the specific maternal attributes (eg, parity, body mass index) or conditions (eg, gestational diabetes mellitus).

In contrast, existing automated EWSs for nonpregnant adults have been developed with the use of explicitly defined populations and outcomes.[7–10] They incorporate information from static variables (eg, diabetes mellitus, heart rate at a discrete point in time) and dynamic ones (eg, rate of change of heart

From the Division of Research, Systems Research Initiative (Drs Escobar and Kipnis and Ms Soltesz) and Perinatal Research Unit (Ms Walsh), Kaiser Permanente Northern California, Oakland, CA; the Department of Obstetrics and Gynecology, Kaiser Permanente Medical Center, Oakland (Dr Gupta) and San Francisco (Dr Terry), CA; and Decision Support, Kaiser Foundation Hospitals, Inc, Oakland, CA (Dr Kipnis).

Received Oct. 12, 2018; revised Dec. 20, 2018; accepted Jan. 10, 2019.

Supported by The Permanente Medical Group, Inc, and Kaiser Foundation Hospitals, Inc, Oakland, CA.

The authors report no conflict of interest.

Corresponding author: Gabriel Escobar, MD. Gabriel.Escobar@kp.org

0002-9378/$36.00

https://doi.org/10.1016/j.ajog.2019.01.208

Click Supplemental Materials under article title in Contents at ajog.org

rate or cervical dilation). The weights or coefficients assigned to their components are based on statistical models. Performance characteristics can be calculated exactly, which permits adjustment of alert thresholds and balancing of available resources with the need to avoid alert fatigue. They also provide health systems with multiple options for how to provide clinicians with probability estimates, which can be displayed directly in the EMR, reported to trained nurses at physical or virtual command centers, or sent by smart phone text alerts to clinicians.

We recently described development and implementation of several EMR-based predictive models that generate probability estimates in real time for nonobstetric patients.[14–17] At present, 3 are in operational use in Kaiser Permanente Northern California (KPNC), an integrated healthcare delivery system that owns 21 hospitals. The first, known as Advanced Alert Monitor (AAM), in use in adult general medical-surgical wards, provides a real time severity of illness score (Laboratory-based Acute Physiology Score, version 2 [LAPS2]),[18] a longitudinal comorbidity score (COmorbidity Point Score, version 2),[18] and a discrete probability estimate.[15–18] This latter estimate, now provided hourly, is the probability that a ward patient will deteriorate within the next 12 hours. At present, AAM is operational in all 21 KPNC hospitals. The second, also operational in all KPNC hospitals, uses many of the same variables and provides a daily probability estimate of a patient's risk for nonelective readmission or death within 30 days after hospital discharge.[19] The third, which provides a quantitative estimate of neonatal early onset sepsis risk, is used in all KPNC nurseries and other health systems.[20–24] Currently, it does require some manual data entry, but work is in progress to instantiate it in the Epic (www.epicsystems.com) EMR. A joint KPNC and Epic team recently has instantiated the AAM and rehospitalization models into this commercial EMR as well.

Our team is developing EMR-based predictive models for early detection of obstetric complications, including those associated with or that could lead to adverse fetal and neonatal outcomes. If we are successful, our models would be embedded in the EMR and provide obstetric teams with probabilistic alerts with sufficient lead time to prevent or mitigate complications. In this article, we conduct a focused theoretic and methodologic reflection that aims to delineate key issues that must be addressed in the development of such models. We cover 3 areas: basic epidemiologic considerations (notably, selection of outcomes and predictors), methodologic considerations peculiar to the L&D setting, and the "nuts and bolts" of data processing and analytics. Because we have discussed these issues elsewhere,[14–18] we only touch instantiation and implementation lightly. A comprehensive treatment of these topics would merit separate articles.

This work has been approved by the KPNC Institutional Board for the Protection of Human Subjects, which has jurisdiction over all the hospitals described in this article. The data on which some of the examples provided herein are based come from 262,071 hospital encounters that took place between January 1, 2010, and March 31, 2017, at the 15 KPNC hospitals that have L&D services.

Defining an attainable desirable state (and how to get there)

Our goal is to develop EMR-based predictive models that could serve as core components for EWSs that are integrated into clinician workflows in L&D and postpartum wards. Such models should match the level of specification that has been reported for AAM and electronic cardiac arrest triage [eCART]).[8,17,18] Given current technology, it is highly desirable that the following be reported for obstetrics predictive models: temporal characteristics, nature and size of the denominator (population at risk), size and rate of the numerator (outcome), discrimination (typically, the area under the receiver operator characteristic curve [C-statistic]), explanatory power, calibration, and number needed to evaluate (or work-up to detection ratio).[14,25–27] Ideally, predictors should be described clearly and justified on both statistical and biologic grounds. Further, if specific patient subsets are excluded, the rationale for exclusion should be made explicit. Finally, the process used to validate the model(s) should also be described.[28]

Temporal characteristics

Figure 1 summarizes temporal characteristics that must be defined when a predictive model is developed. The T_0 is the time when the model issues a probability estimate that at least 1 undesirable event (X) will occur within some elapsed time, which is the event (or look forward) time frame. The time between the T_0 and X (lead time) should be sufficient for clinicians to mount a response. All predictive models are based on data that were available before the T_0, which we refer to as the "look back" period. Compared with adult EWSs, unique challenges are present. The first challenge is that one is not dealing with a single temporal frame: one must consider antepartum and postpartum phases as both physiologic changes and presence of a fetus can alter patient outcomes in different ways (Figure 1). Unlike scores such as AAM and eCART, in which the set of predictors can remain constant, the presence of 2 phases complicates analyses: the information value of predictors, such as dilation and station, changes dramatically after delivery. Second, time frames in L&D are often highly compressed. The "look forward" time frames for AAM and eCART are 12 and 24 hours, respectively, which makes sense in the context of adult medical-surgical wards where average lengths of stay range from 3–5 days. Most L&D and postpartum stays are shorter, and the antepartum phase may sometimes be extremely brief. Consequently, we aim to achieve a "look forward" time frame of 6 hours. A third problem involves the "look back" period. For a given physiologic parameter, all severity of illness scores currently in use select the worst (most highly deranged) value in the "look back" time frame. This is problematic because vital signs are often temporarily abnormal during the

most important time period in L&D: delivery. We expand on this problem in the section on severity scores.

What is the denominator?

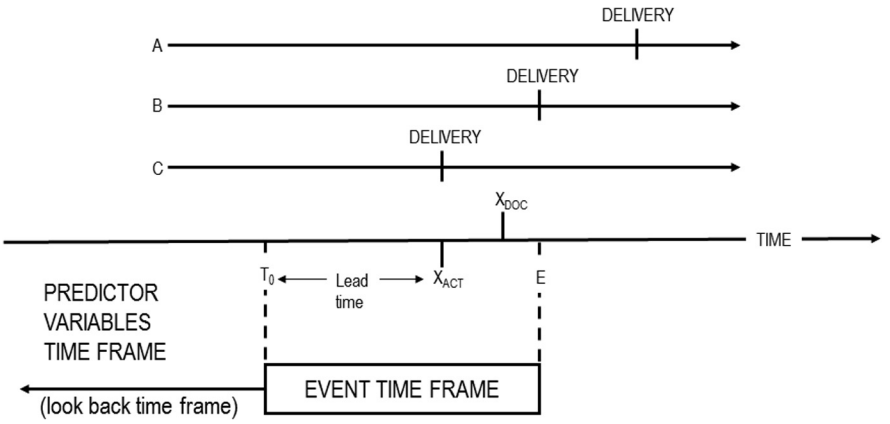
In this article, we focus on prevention and mitigation of obstetric complications in the L&D service using EWSs, rather than improving management of such complications in other settings, such as the intensive care unit (ICU). Given this focus, the proper denominator should be the largest one: all women who are admitted to L&D, independent of outcome or subsequent disposition. Note that this denominator precludes inclusion of women who arrive for triage, whether in the emergency department or some other location linked to L&D, but who are not admitted to the service. We are not including women who are admitted to other hospital units in our models because these women would be monitored by systems such as AAM or eCART. Women who are discharged home from triage are not necessarily at zero risk, and those with repeated visits are at higher risk of having a newborn infant who requires respiratory support.[29] However, monitoring women who are sent home after triage is outside the scope of our current work. Similarly, although the study of how scores may predict outcomes in pregnant women after they have become critically ill and/or have been transferred to the ICU[30,31] is of value to intensivists, it does not address the needs of routine obstetrics. Table 1 describes the cohort assembly for our predictive model development, and a graphic display of this process is in Appendix 1. Note that, although we are excluding women who were admitted to L&D after experiencing a fetal loss from our denominator for modeling purposes, such women would be monitored when the EWS goes live.

What is the numerator?

Any epidemiologic investigation, whether it supports the development of a predictive model or not, should also clearly define its numerator(s). Unlike the situation with neonatal, pediatric, or adult EWSs, selection of an outcome variable in obstetrics is difficult. This is because at least two patients are involved: the laboring mother and her offspring. Ideally, obstetrics providers should be provided with risk estimates for maternal and fetal harm. We have made an explicit decision to include 3 fetal/neonatal outcomes in the numerator. The rationale for this is that some of these outcomes might be linked to decisions that could be informed by a probability estimate. For example, it is possible that subtle changes in vital signs and the interactions of these with cervical dilation measurements could be associated with uterine rupture. Because some cases of uterine rupture are identified only at the time of cesarean delivery and documented subjectively in an operative note, the only indicator of such an event might be an ill newborn infant (which can be identified objectively with cord gases). Thus, it might be necessary to use 2 or more predictive models. These could run concurrently, and the fact that 2 models are in use need not be apparent to the end user, who need only see a risk estimate in the EMR graphic user interface. The risk estimate would be the probability that any adverse outcome would occur within a defined time frame (which we hope will be 6 hours).

Another problem that needs attention is how one defines and captures adverse outcomes. In our work, we prefer to use outcomes that can be defined objectively and also be tightly linked to a discrete EMR date and time stamp. For example, although it may be possible to link a vaginal laceration to the moment of birth, not all clinicians may record that these occurred, and they may not be captured in real time by hospital coders. Table 2 shows the outcomes we are using for model calibration, relevant time stamps, and preliminary incidence estimates. We have made an explicit decision to develop 2 models (1 for antepartum and 1 for postpartum events). Because of sample size limitations, discussed later, we will be pooling events into composite outcomes (eg,

FIGURE 1
Obstetric prediction time frames



Obstetric considerations that are relevant to preparing data for the development of an early warning system: at a T_0, a detection system issues a probability estimate that an undesirable event, X (which must be defined explicitly), will occur within some elapsed time (event time frame). This time should be sufficient to provide ample lead time for an adequate clinical response. Note that the time when X actually occurred (X_{ACT}) is not the same as when X was documented in the electronic record (X_{DOC}). Data used to populate the early warning system not only must be from the time preceding the T_0 (look back time frame) but also actually should be available at the T_0; data might not be available because of charting delays. Figure also shows that the event time frame can **A,** precede, **B,** coincide with, or **C,** come after the time of delivery. Consequently, predictive models that issue estimates may be predicting simultaneously for antepartum, intrapartum, and postpartum events.

Escobar. Development of methods for obstetric automated early warning systems. Am J Obstet Gynecol 2019.

TABLE 1
Inclusion and exclusion criteria for obstetrics early warning cohort dataset

Criteria	Description and rationale
Inclusion	
Admitted to labor and delivery service,[a] delivered	Core denominator for "bread and butter" obstetrics. Note that "delivered" includes both live births and fetal losses, because a woman who experienced a fetal loss may still have adverse outcomes before and after delivery.
Gestational age ≥22 wk	Pathophysiology of miscarriages before this gestation is not well understood, and it is unlikely that a physiology-based predictive model that uses currently available electronic medical records data elements can predict for such losses.
Arrived at labor and delivery service with fetal heart rate present	Early warning system is calibrated for women who had a living fetus on arrival.
Admitted to labor and delivery service, record not located using hospital service code	When constructing a dataset for predictive model development, some clerical errors may be present in the initial data extraction. The data for women with labor and delivery service records located via alternative linkage strategies (eg, erroneously listed as being ward patients, but maternal record was found by linkage to neonatal record) should be retained.
Exclusion	
Not pregnant	Not eligible for early warning system, even if record is listed (erroneously) as labor and delivery service.
Not admitted to labor and delivery service	Some women are admitted initially to other services (emergency department, ward). If a woman is transferred to the labor and delivery service, that record is included in the cohort, with T_0 for admission being the time stamp for admission to the labor and delivery service or labor and delivery triage room. Maternal data (eg, vital signs, laboratory test results) from these other services are included so long as they fall within the "look back" time frame of early warning system.

Escobar. Development of methods for obstetric automated early warning systems. Am J Obstet Gynecol 2019. *(continued)*

hemorrhage OR pulmonary embolus OR adverse birth event), which has operational implications that are discussed later. However, as noted by Zuckerwise and Lipkind[5] and Friedman et al[6], additional models for different outcomes using different denominators will be needed.

Definition of a time stamp may be challenging for some outcomes. For example, women in L&D seldom have spontaneous cardiac arrest; when asystole or ventricular fibrillation occurs, it is usually a distal event, typically preceded by proximal events such as hemorrhage or pulmonary embolus. Consequently, if one uses the time stamp for cardiac arrest or death, one may not be calibrating the model properly. For this reason, the records of all maternal deaths that are included in our model are being manually reviewed by a clinician panel. This panel is assisting us in determining a proximal cause (eg, hemorrhage) that will be used to define the time stamp. Similarly, the use of transfer to the intensive care or telemetry unit as an outcome is problematic because some women may be transferred there preventively because of a preexisting condition (eg, severe cardiac disease, severe obstetric sleep apnea because of narcotics) or for reasons unrelated to obstetric care (eg, diabetic ketoacidosis or after surgery for a malignancy discovered during pregnancy). Finally, assigning a time stamp for some outcomes may also be problematic because multiple time stamps exist. This is the case with severe preeclampsia and hemorrhage, which can be defined in multiple ways based on combinations of physiologic markers (Figure 2).

What are the predictors?

The fact that outcomes can occur in the antepartum, postpartum, or both time frames affects predictor selection. In theory, as has been suggested recently,[32] one could use all available data in the EMR for prediction using machine learning and neural networks. In practice, given limited computational resources and time, we will start with biologically plausible predictors and use machine learning and content expertise to define the final set included in the model. The same constraints we have for numerators also apply to predictors: they must be objective, and they must have a time stamp. A specific outcome with a discrete time stamp can occur only once, but many predictors vary across time. Table 3 shows the predictors that we will be evaluating in both of our models, with those that are associated with a single measurement (eg, body mass index) at the top and those that repeat at the bottom. The variables listed in Table 3 highlight the problematic nature of prediction in L&D. Some obstetric emergencies (eg, shoulder dystocia) can occur suddenly, without preceding vital signs or laboratory abnormalities, so "traditional" predictors that are used in many EWSs and severity of illness scores may not be useful (or they may be useful only in some cases). Others, such as

hemorrhage, usually accompanied by changes in vital signs, are likely to be predictable with the use of modeling approaches that are used for adult EWSs.

The predictor list shown in Table 3 is not exhaustive or definitive. In the course of the predictive modeling process, many variables may be transformed and/or combined with other variables (interaction terms). For example, 1 important variable for the prediction of antepartum complications may be the change in cervical dilation over time. The LAPS2 contains multiple interaction terms (eg, lactate-pH, heart rate divided by systolic blood pressure, blood urea nitrogen divided by creatinine), although the AAM model also includes instability terms (change in vital signs over time). Many of the candidate variables shown in Table 3 may be eliminated on statistical grounds during the modeling process.

TABLE 1
Inclusion and exclusion criteria for obstetrics early warning cohort dataset *(continued)*

Criteria	Description and rationale
Intrauterine fetal demise before labor and delivery service admission	Women who arrive to the labor and delivery service without a detectable fetal heart tone are not the primary focus of the early warning system.
Gestational age <22 wk, independent of outcome	Pathophysiologic condition of miscarriages before this gestation is not well-understood, and it is unlikely that a physiologic-based predictive model that uses currently available electronic medical records data elements can predict such losses. The early warning system is designed for women who have a delivery.
Unusual service pattern	This category includes women who have other conditions precluding admission to the labor and delivery service. For example, women admitted to the intensive care unit for preexisting conditions (eg, malignancy, uncommon cardiac conditions). Although these women clearly need special monitoring, their data might distort a predictive model targeting "bread and butter" obstetrics. Records of these women need manual review for determination of whether they should be included in main study cohort.

[a] Includes any designated hospital antepartum or postpartum unit.

Escobar. Development of methods for obstetric automated early warning systems. Am J Obstet Gynecol 2019.

Potential uses of existing adult severity of illness scores in obstetrics

Physiologic-based severity scores occupy a peculiar niche in this methodologic space, because they can serve as both predictors and outcomes. These scores, which include the Acute Physiology and Chronic Health Examination and Simplified Acute Physiology Score are based on multiple predictors that may include vital signs, pulse oximetry, mental or neurologic status measures, and laboratory test results.[33] The statistical weights assigned to specific abnormalities (eg, the number of points assigned for a heart rate of ≥120 beats/min) are based on multivariate models rather than the value of the predictor taken in isolation. A number of these scores exist, with most having been calibrated for death in patients in the ICU, which was among the earliest locations where these physiologic variables became widely available electronically.[33] Some have been evaluated for possible use in obstetric patients already in the ICU.[30,31] However, to our knowledge, formal investigation of the use of automated ICU severity scores in obstetrics patients outside the ICU has not yet been reported.

Using data from 391,584 KPNC adult hospitalizations, our team developed an acute physiology score (LAPS2[18,34]) that was calibrated for inpatient deaths among all hospitalized nonlaboring adults, not just those in intensive care. Although the LAPS2's statistical weights were derived using patient mortality, the score is not used to predict death directly. It is used in multiple KPNC predictive models as a dimensionless number that synthesizes information from a patient's vital signs, pulse oximetry, neurologic status, and 16 laboratory tests; it has a "look back" time frame of 72 hours. An earlier version of this score, the LAPS, has also been validated externally in Canada.[35] The higher the LAPS2, the greater the degree of acute physiologic derangement and the higher the death risk. Working with a team from Epic, our team has recently validated a real-time version of the LAPS2 that is embedded in the Epic electronic medical record. In KPNC, all adults in the hospital are now being assigned a LAPS2 score with a "look back" time frame of 72 hours at the time of hospital admission and a score with a "look back" time frame of 24 hours (OB24) every hour.

In theory, the LAPS2 (or scores like it, including eCART) could be used "as is" for use in an obstetric EWS. In practice, this may not be useful for 2 reasons. The first is that vital signs and laboratory test results in laboring and postpartum women do not have the same normal ranges as those of adults in general medical-surgical wards. The second, related to the "look back" time frame and the tachycardia and tachypnea that are seen around the time of delivery, was mentioned previously. Figure 3 shows how the distribution of hourly LAPS2 scores can be affected by delivery and choice of "look back" time frame. Our approach will be to evaluate 3 variants of the LAPS2 as potential predictors: the original score (72-hour "look back" time frame), the LAPS2 OB24, and the LAPS2 postpartum score (PP02). This latter score is a postpartum score, that is

TABLE 2
Antepartum and postpartum outcomes for obstetric early warning system

Outcome	Description and how confirmed in electronic medical record	Frequency, n (rate per 1000 deliveries)
Fetal death[a]	Only includes fetal deaths that occurred after a woman was admitted to labor and delivery service. Time stamp is obtained from nursing flow sheet: if fetal heart rate of 0 is documented, that is used; otherwise, time stamp is time of delivery.	55 (0.21)
Hypoxic-ischemic encephalopathy[a]	Outcome is ascertained based on research registry that captures medical record numbers of all newborn infants who are eligible for therapeutic hypothermia protocol. Time stamp used is time of birth.	315 (1.20)
Neonatal acidosis[a]	Defined as any blood gas (cord or infant) with a base deficit of minus 12 or more in the initial hour of life, and either of (a) intensive care nursery admission for ≥24 hours or (b) neonatal disposition of transport or death. Overlap exists with hypoxic-ischemic encephalopathy. Relevant blood gases are obtained from laboratory database. The time stamp used is time of birth. Note that the newborn infant outcome is assigned to the mother.	904 (3.45)
Eclampsia[a,b]	100% manual ascertainment (all records with International Classification of Diseases code(s) for eclampsia or documented seizure(s) during the encounter, are manually reviewed); time stamp assigned based on first documented seizure in seizure flowsheet.	1 (0.05)[a] 5 (0.02)[b]
Severe preeclampsia[a,b]	Defined as meeting criteria for preeclampsia **and** relevant biochemical abnormalities (eg, elevated liver function tests) **and** ever had Laboratory-based Acute Physiology Score, version 2 score ≥60[a] or 80.[b] Time stamp is time when severity threshold was reached.	1134 (4.33)[a] 107 (0.41)[b]
Hemorrhage[a,b]	Defined as any 1 of the following events: (a) patient was ever transfused with ≥4 units packed red blood cells; (b) patient was transfused with 1−3 units packed red blood cells **and** had a documented hematocrit level <18%; or (c) patient had documented estimated blood loss >1500 mL and a hematocrit level <18%. Time stamp is the earliest of 1 of the following events: (a) transfusion time, (b) hematocrit time, (c) estimated blood loss time, or (d) time when patient had a Laboratory-based Acute Physiology Score, version 2 score ≥60 (for antepartum hemorrhage) or 80 (for postpartum hemorrhage).	83 (0.32)[a] 844 (3.22)[b]
Emboli[a,b]	100% manual ascertainment (all records with relevant International Classification of Diseases code are reviewed manually; only those that were not preexisting and that have evidence of new anticoagulation treatment in electronic medical records medication administration record are retained). Includes pulmonary, air, fat, and amniotic fluid emboli. Time stamp is defined as either (a) the first time patient reached a Laboratory-based Acute Physiology Score, version 2 score ≥70[a] or 90[b] or (b) the time of the highest score in the antepartum or postpartum phase. Other deep venous thromboses are not included, because they are not amenable to detection by the vital signs-based early warning system.	11 (0.04)[a] 22 (0.08)[b]
Transfer to intensive care[a,b]	Ascertained from bed history. Patients admitted to intensive care preventively will not be considered to have had this outcome, as will patients whose admission was due to nonobstetrics issues (eg, surgery for colon cancer, severe influenza). Exclusion may be algorithmic (ie, patients admitted to intensive care with low severity of illness may be considered to have been preventive admissions). Some patients with unusual clinical conditions may need to be excluded from denominator altogether.	42 (0.16)[a] 595 (2.27)[b]
Major deterioration without transfer to intensive care[a,b]	Assigned algorithmically based on presence of very high Laboratory-based Acute Physiology Score, version 2 score (≥110[a], 120[b]), with time stamp when patient first reached a score of 90[a] or 100.[b] Outcome is intended to capture major physiologic derangement not captured by the other outcomes listed earlier.	29 (0.11)[a] 254 (0.97)[b]
Uterine rupture[a]	Ascertained from International Classification of Diseases codes. If patient had an elevated laboratory-based acute physiology score, version 2, score before delivery, rupture time stamp is time of delivery minus 20 minutes. If no elevated score, then use time of delivery as time stamp.	249 (0.95)
Maternal death[b]	After decedents are identified in electronic medical records, records are reviewed manually by an expert panel to ascertain time of proximal event (eg, hemorrhage, embolus), because distal event (eg, cardiac arrest) may have occurred after the patient was discharged home. The time stamp of the proximal event will be used for modeling.	16 (0.06)

[a] Antepartum; [b] postpartum. (See text and Escobar et al[18] for details on the Laboratory-based Acute Physiology Score, version 2).

Escobar. Development of methods for obstetric automated early warning systems. Am J Obstet Gynecol 2019.

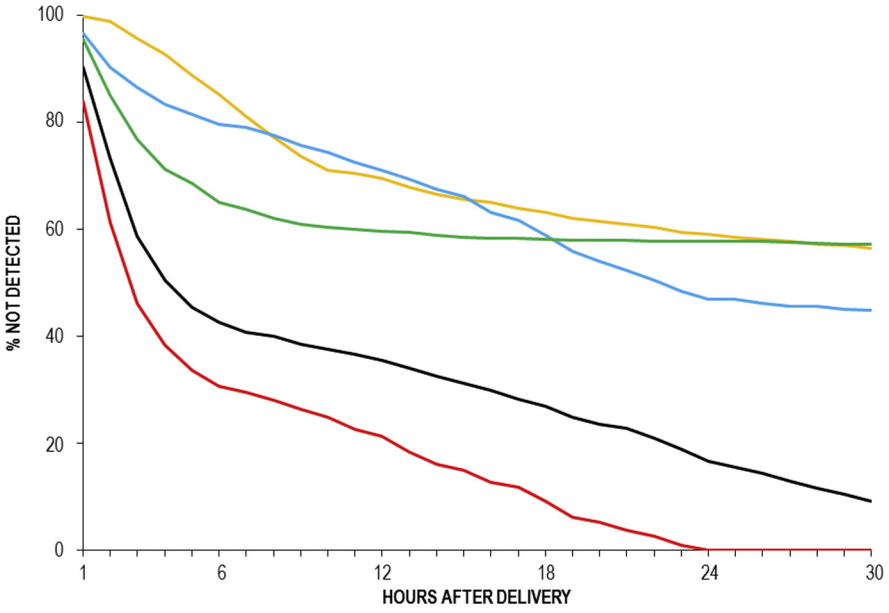
available only after delivery, and has these characteristics: the T_0 begins immediately after delivery, no vital signs are included from the time period up to and including delivery, and (subject to the 2 preceding constraints) the "look back" period for vital signs, neurologic status, and pulse oximetry is set to 2 hours, with this time frame set to 24 hours for laboratory tests. The rationale for testing the LAPS2 PP02 as a potential predictor is that, by explicitly excluding antepartum vital signs (which may be elevated and thus result in high scores), it may prove better at the detection of postpartum abnormalities. The potential value of combining severity scores with other physiologic and process markers is shown in Figure 2. Note that use of these scores (which are based on multivariate models that synthesize interactions between physiologic predictors) does not preclude concurrent use of individual vital signs and laboratory tests as well (Table 3).

Adult severity scores can function as outcomes, not just predictors. For example, during AAM's development, our team found that adult ward patients with an admission LAPS2 ≥110 are at much higher risk of requiring unplanned transfer to the ICU, and analyses of real-time hourly LAPS2 scores among nonlaboring adults who were admitted to KPNC ICUs have found that the median LAPS2 ranges from 111−120. Because it is undesirable that patients reach a high degree of physiologic instability and because it is now possible for us to assign such scores electronically both retrospectively and in real time, it stands to reason that one could pick a high LAPS2 value (eg, 110 or 120) and treat it as an outcome. Alternatively, it is also possible to combine the LAPS2 with the time stamp for ICU transfer, which would permit an algorithmic approach to distinguishing between preventive ICU admissions and ICU admissions "for cause."

What is the role of electronic fetal monitoring?

Extensive discussion of the role of electronic fetal monitoring (EFM) in obstetric patient safety is beyond this article's scope, but we can make some methodologic observations. Currently, important limitations exist with respect to being able to make accurate predictions with the use of EFM.[36−39] It is likely that novel approaches to the analysis of streaming data (eg, as described by Cahill et al[40]), including those employing machine learning, eventually will permit more consistent use and interpretation of EFM data. From the perspective of our current work, there are 2 important limitations of EFM. The first is that, in current obstetric practice, not all women are monitored continuously and/or have inconsistent interpretation of EFM tracings. Thus, the actual risk (prior probability of an adverse event) at the time EFM is initiated is not known. Second, because we are attempting to predict both maternal and fetal/neonatal outcomes, the fact that EFM does not predict for other antepartum outcomes (eg, hemorrhage) and is of no use in the postpartum period limits its utility for our current work. However, we suspect that, in the future, it may be possible to combine data from EWSs, such as the ones we are developing with those from EFM.

FIGURE 2
Detection and electronic time stamps



Timing of postpartum hemorrhage as a function of which electronic marker was used among 665 women who experienced this outcome within 24 hours after delivery. Each *line* shows the proportion of women whose condition remained undetected after delivery based on the time stamp of a given electronic marker: *yellow line*, time of first Laboratory-based Acute Physiology Score, version 2 (*LAPS2*) ≥80; *blue line*, hematocrit <18%; *green line*, time of first entry for estimated blood loss >1500 mL; *black line*, time of first transfusion order; and *red line*, the earliest of any of these. The LAPS2 used a 24-hour "look back" time frame. (See text and Escobar et al[18] for details on the LAPS2.)

Escobar. Development of methods for obstetric automated early warning systems. Am J Obstet Gynecol 2019.

The nuts and bolts of predictive modeling

In this article, we will not go into detail on how one actually conducts predictive modeling once one has a properly assembled dataset, because this topic is covered extensively in the statistical and machine learning literature. Instead, we will focus on 2 critical topics that have not received attention in the obstetrics literature: structuring predictor and outcomes data (data processing) and sample size considerations.

TABLE 3
Predictors to be used in obstetric early warning system

Variable	Description and rationale
Single instance predictors	
Maternal age in years	Captured from demographic databases. Extremes are known to be associated with increased rates of adverse outcomes.
Gestational age in weeks	Located by electronic scanning; if missing, calculated algorithmically based on data that would be available in real time.
Parity	Located by electronic scanning; if missing, default to primigravida status.
Multiple gestation	Identified from the presence of fetal heart tones for >1 infant during the delivery encounter.
COmorbidity Point Score, version 2	12-month longitudinal open source comorbidity score calculated based on Centers for Medicare and Medicaid Services Hierarchical Condition Categories; the higher the COmorbidity Point Score, version 2, the greater the preexisting comorbid illness burden. (See text and Escobar et al[18] for details.) Score can be calculated in real time.
Gestational diabetes mellitus	Located by electronic scanning for relevant International Classification of Diseases codes and laboratory tests (hemoglobin A1c, glucose tolerance tests).
Diabetes mellitus	Located by electronic scanning for relevant International Classification of Diseases codes and laboratory tests (hemoglobin A1c, glucose tolerance tests).
Time of rupture of membranes	Discrete time stamp exists in electronic medical records. If missing, default to time of delivery.
Amniotic fluid characteristics	Located by electronic scanning; defined algorithmically (all entries that do not state fluid as clear are classified as "not clear"; missing defaults to "clear").
Fetal scalp electrode, intrauterine fetal monitoring	Discrete electronic medical records time stamp for placement exists. (See text for discussion on fetal heart rate monitoring.)
Body mass index	Discrete field exists in electronic medical records; if not available there, outpatient record up to 30 days before arrival to labor and delivery used; otherwise default to normal.
History of emboli	Electronic scanning for International Classification of Diseases codes and evidence of anticoagulation treatment.
History of hypertension	Electronic scanning for International Classification of Diseases codes.
Repeating predictors to be captured or assigned hourly	
Individual maternal vital signs	Temperature, heart rate, respiratory rate, systolic blood pressure, diastolic blood pressure; all are found in nursing flowsheets.
Pulse oximetry	Found in nursing flowsheets.
Neurologic status	Determined algorithmically based on free text entries from nursing flowsheets (see Escobar et al[18,34] for details).

Escobar. Development of methods for obstetric automated early warning systems. Am J Obstet Gynecol 2019. *(continued)*

Data processing

Most predictive models currently in use in medicine are static, by this we mean that one starts with a set of predictors defined at some discrete point in time (eg, various risk factors such as serum triglycerides, age in years, and family history) that one then uses to predict some outcome (eg, risk of myocardial infarction within X years). Static predictive models usually use a simple "flat file" data structure, with 1 row per observation, a simple yes/no outcome (0 for observations without the outcome; 1 for those with the outcome), and individual predictors as columns. In our case, a static prediction model would have 262,071 rows and, assuming 50 predictors and a single outcome for each phase, a total of 53 columns: 1 for a patient identification number (blinded study ID), 1 each for a composite dichotomous antepartum/postpartum outcome, and 50 for predictors.

In contrast, if one aims to issue regularly updated probability estimates for patients in the L&D and postpartum services, a very different data structure is required. This data structure must take the dynamic nature of predictors and the existence of 2 phases (antepartum and postpartum) into account. For example, if one uses cervical dilation, it is important to keep in mind that a value of 7 cm has a very different meaning if present for 1 hour than if present for 8 hours; put differently, the data structure must take elapsed time and time to the outcome into account. Similarly, it is critical that time to delivery (which will not be known in real time) or time from delivery (which will be known) be considered in all analyses. This requires a very different data structure, in order to account for the changing value of information. For this approach, the dataset can be structured to add a row each time a new piece of information arrives (as is the case with eCART, which updates estimates every time a new laboratory test or vital sign is entered) or, in the approach we are taking, a row is added for each patient hour in the dataset. This results in much larger datasets. For example, in our case, with a dataset of 262,071 hospital records with average hospital length of

stay of approximately 63 hours, the dataset we have created consists of approximately 16.6 million (63×262,071) rows and approximately 100 columns (1 for the study identifier, 1 each for the antepartum/postpartum outcome flags, with remaining columns for individual predictors). A detailed description of how one structures data from actual obstetric emergencies into such a dataset is provided in Appendix 2. These approaches to data processing are common in predictive modeling and are used not only for AAM and eCART but also for models such as those developed by Simon et al[41] to predict suicide attempts.

TABLE 3
Predictors to be used in obstetric early warning system *(continued)*

Variable	Description and rationale
LAPS2 (3 variants)	The LAPS2 includes neurologic status, pulse oximetry, and all vital signs; it also includes 16 laboratory tests. (See text and Escobar et al[18,34] for details.)
Individual laboratory tests	All individual laboratory tests used in LAPS2 severity score, plus: magnesium, aspartate aminotransferase, alanine aminotransferase, lactate dehydrogenase, uric acid, urine creatinine, urine protein, urine protein creatinine ratio; hemoglobin, hematocrit, platelet count. All are found in electronic medical records or laboratory database, with discrete time stamps.
Station of fetal descent, cervical dilation in centimeters	Found in nursing flowsheets.

Escobar. Development of methods for obstetric automated early warning systems. Am J Obstet Gynecol 2019.

Sample size considerations

When predictive models are being developed, a large population is desirable but not sufficient; what are most critical are the number of outcomes and the proportion of outcomes in the population. A general rule of thumb for regression models is that one should have 10 outcomes for each predictor in the model.[42−44] However, this rule primarily applies to static regression models. The situation is different for models that aim to predict in real time because the outcome rate falls dramatically when the event time frame is extremely brief (eg, 6 hours). These difficulties are exacerbated by the "class imbalance" problem, which occurs when the number of nonevents is much larger than the number of events and usual modeling metrics have poor accuracy.[45,46] For example, when the outcome rate is very low, the C-statistic is much less valuable; greater attention must be paid to the number needed to evaluate (work-up to detection ratio).[27]

When reporting on model performance, it is also important to distinguish whether one is reporting rates based on a single time frame (eg, the sensitivity when one only uses the exact initial prediction time) or on multiple time frames (eg, the sensitivity when one uses all time periods after an initial prediction). Suppose a predictive model with a 6-hour "look forward" time frame issues an alert at 3:15 AM on June 27, 2019, and an adverse outcome occurs at 11:50 AM on the next day. If one reports on performance based on the exact prediction time, then the model failed; on the other hand, if one reports on whether the model "ever" detected an outcome, then the model succeeded.

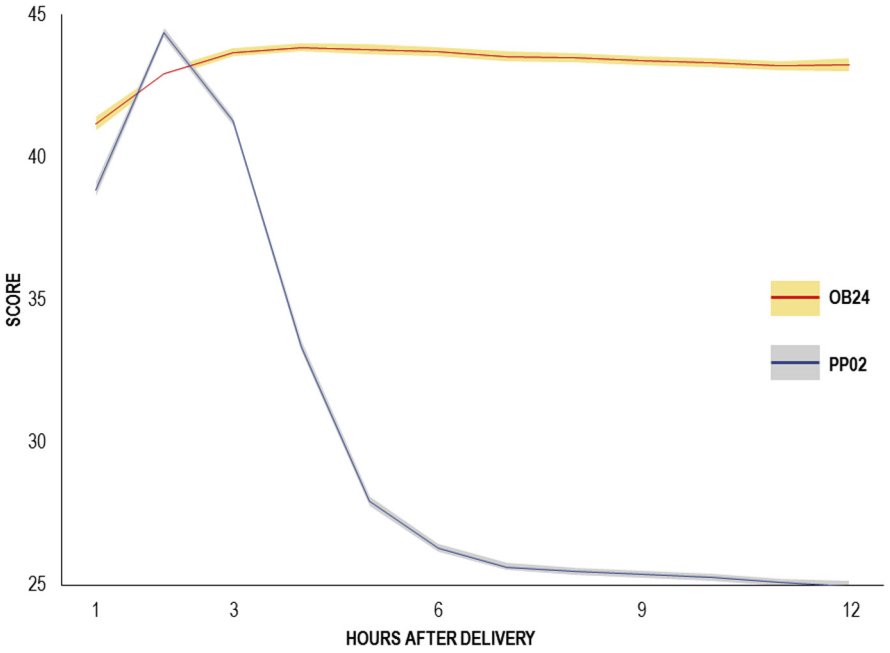
Table 2 shows that event rates in obstetrics are very low, which would make individual models for each outcome very difficult to develop. Because of this, we will need to pool all study events into 2 global outcomes (antepartum and postpartum). It is instructive to consider that, to develop the AAM model, which was statistically challenging, we had access to a dataset of 649,418 hospitalization episodes that involved 374,838 patients, with 19,153 of the episodes (2.9%) having at least 1 outcome.[17] Thus, in this work, we must face the possibility that our models may not be successful or that they will be predictive but have extremely high numbers needed to evaluate, which would make them difficult to use operationally.

Instantiation

Once a predictive model is developed, considerable work remains. We will touch only on these issues briefly, because we and others have discussed them elsewhere[9,10,15] and because the topics would merit separate articles. Generally speaking, 3 possible approaches exist for generating automated probability estimates from an EMR. In the first, known as a "web service," data (eg, vital signs, indicators of the progress of labor) are exported out of the EMR to an external application that applies the algorithm and then "writes" the result in the EMR or in an external web page. In the second, a real-time EMR "mirror" or "shadow" server provides data for the algorithm (ie, data that are a 100% match for the EMR but with a small delay, typically less than a few minutes); algorithm output is then "written" to the EMR or in an external web page. The last option, which is the one that KPNC will be transitioning to for the aforementioned models, is to have all algorithms run directly within the EMR (although it would seem, intuitively, that this is the best option, in actual fact such an approach, depending on the type of EMR, can cause significant transaction delays and slow down EMR function for others, so it is not always an option). In addition to requiring full cooperation from a hospital or hospital system's information technology department, all 3 options would also require changes in clinician work flow and formal approval by a hospital's Executive Committee.

Another important instantiation issue is how to alert clinicians. When first piloted, AAM probability estimates were displayed directly in the Epic EMR hospitalist and rapid response team dashboard every 6 hours.[14-17] Subsequently, when the decision was made to deploy systemwide with hourly data scans, KPNC clinical leaders elected to stop displaying alerts directly in the EMR. Instead, a command center

FIGURE 3
Look back time frame effects on severity scores



Distribution of values of 2 versions of an adult severity of illness score (Laboratory-based Acute Physiology Score, version 2 [*LAPS2*]) in a cohort of 21,868 randomly selected healthy delivery encounters. The *upper line* shows the LAPS2 OB24, which has a "look back" time frame of 24 hours; the *bottom line* shows the LAPS2 PP02, which is assigned only after delivery and which selects the worst laboratory test results from the preceding 24 hours but (1) does not include any vital signs from the time of delivery or earlier and (2) uses a very restricted (2-hour) "look back" time frame for vital signs. The Figure highlights the importance of the "look back" time frame; the distribution of LAPS2 OB24 scores (which include vital signs during the delivery process) is much higher than that of the LAPS PP02, which explicitly excludes delivery vital signs, and could offer some mathematic advantages for the detection of postpartum outcomes. Note that, because of the large sample size, confidence intervals are very narrow. (See text and Escobar et al[18] for details on the LAPS2.)

OB24, 24-hour "look back" time frame; *PP02,* postpartum score with 2-hour look back frame for vital signs.

Escobar. Development of methods for obstetric automated early warning systems. Am J Obstet Gynecol 2019.

approach is now in use: alerts are displayed in a separate website where trained nurses review scores remotely and act as first responders. These trained nurses conduct a preliminary chart review before notifying the rapid response team. They also serve as a buffer against alert fatigue, in that they can "snooze" an alert while clinicians respond to an alert.

Implementation

Simply having a score or probability estimate does not guarantee practice change; clinicians must use the new information in meaningful ways. A full description of the challenges that are involved in this process is outside the scope of this article. Implementation of EWSs or protocols, automated or not, requires not only having clinician "buy in" but also substantial organizational investment is necessary, which has been described in the obstetric literature[4] and in the adult setting.[16] One issue that affects adult EWSs, which is the fact that many patients who meet the alert threshold may not desire rescue because they are near the end of life,[14,47] has not yet been addressed in the obstetric literature. Given that maternal death is rare, this is reasonable, but we do need to start considering the implications of EWSs that could predict fetal loss or early neonatal death.

One important limitation of existing EMRs and predictive models is that it is not always possible to pinpoint exactly which variables led to a probability estimate. The major reason for this is that models may require the use of multiple interaction terms, which makes it difficult to "tease out" the contribution of an individual variable. In the case of the models that we are developing, an additional problem is the use of a composite outcome, which is a limitation imposed on us because of the sample size considerations described earlier.

Conclusions

Given the growing availability of data in comprehensive EMRs and the increasing number of integrated healthcare delivery systems, automated EWSs for obstetrics are going to be developed. As the scientific community starts working on and evaluating these systems, the issues raised in this article will need further discussion. In addition, novel collaborative structures may be needed for the development of predictive models because the incidence of these events is quite low. ■

ACKNOWLEDGMENTS

We thank our executive sponsors, Nancy Goler, MD; Barbara Crawford, MS, RN, NEA-BC; and Robin Betts, MBA-HM, RN, CPHQ, for securing funding for this work and providing administrative assistance; the Division of Research Strategic Programming Group (Jamila Gul, Wei Tao, Mei Lee, and Jonathan Lontok) for their assistance in developing the study datasets; Drs Mara Greenberg and Michael Kuzniewicz for methodologic advice; Drs Stephen Parodi and Tracy Flanagan for administrative support; Hamid Niki for programming assistance; Kirby McKinnon for project management support, and Kathleen Daly for her help with the graphics.

REFERENCES

1. MacDorman MF, Declercq E, Cabral H, et al. Is the United States maternal mortality rate increasing? Disentangling trends from measurement issues. Obstet Gynecol 2016;128:447.

2. UNICEF. Trends in estimates of maternal mortality ratio (maternal deaths per 100,000 live births) 1990-2015. February 2017 ed, 2017 https://data.

unicef.org/resources/dataset/maternal-mortality-data/.
3. Isaacs RA, Wee MY, Bick DE, et al. A national survey of obstetric early warning systems in the United Kingdom: five years on. Anaesthesia 2014;69:687–92.
4. Shields LE, Wiesner S, Klein C, et al. Use of maternal early warning trigger tool reduces maternal morbidity. Am J Obstet Gynecol 2016;214:527.e1–6.
5. Zuckerwise LC, Lipkind HS. Maternal early warning systems: towards reducing preventable maternal mortality and severe maternal morbidity through improved clinical surveillance and responsiveness. Semin Perinatol 2017;41:161–5.
6. Friedman AM, Campbell ML, Kline CR, et al. Implementing obstetric early warning systems. AJP Reo 2018;8:e79.
7. Rothman MJ, Rothman SI, Beals JT. Development and validation of a continuous measure of patient condition using the electronic medical record. J Biomed Inform 2013;46:837–48.
8. Churpek M, Yuen T, Winslow C, et al. Multicenter development and validation of a risk stratification tool for ward patients. Am J Respir Crit Care Med 2014;190:649–55.
9. Kollef MH, Chen Y, Heard K, et al. A randomized trial of real-time automated clinical deterioration alerts sent to a rapid response team. J Hosp Med 2014;9:424–9.
10. Evans RS, Kuttler KG, Simpson KJ, et al. Automated detection of physiologic deterioration in hospitalized patients. J Am Med Inform Assoc 2015;22:350–60.
11. Parshuram CS, Dryden-Palmer K, Farrell C, et al. Effect of a pediatric early warning system on all-cause mortality in hospitalized pediatric patients: the EPOCH randomized clinical trial. JAMA 2018;319:1002–12.
12. Halpern NA. Early warning systems for hospitalized pediatric patients. JAMA 2018;319:981–2.
13. Behling DJ, Renaud M. Development of an obstetric vital sign alert to improve outcomes in acute care obstetrics. Nurs Womens Health 2015;19:128–41.
14. Escobar GJ, Dellinger RP. Early detection, prevention, and mitigation of critical illness outside intensive care settings. J Hosp Med 2016;11(suppl1):S5–10.
15. Escobar GJ, Turk BJ, Ragins A, et al. Piloting electronic medical record-based early detection of inpatient deterioration in community hospitals. J Hosp Med 2016;11(suppl1):S18–24.
16. Dummett BA, Adams C, Scruth E, et al. Incorporating an early detection system into routine clinical practice in two community hospitals. J Hosp Med 2016;11(suppl1):S25–31.
17. Kipnis P, Turk BJ, Wulf DA, et al. Development and validation of an electronic medical record-based alert score for detection of inpatient deterioration outside the ICU. J Biomed Inform 2016;64:10–9.
18. Escobar GJ, Gardner MN, Greene JD, et al. Risk-adjusting hospital mortality using a comprehensive electronic record in an integrated health care delivery system. Med Care 2013;51:446–53.
19. Escobar GJ, Ragins A, Scheirer P, et al. Nonelective rehospitalizations and postdischarge mortality: predictive models suitable for use in real time. Med Care 2015;53:916–23.
20. Puopolo KM, Draper D, Wi S, et al. Estimating the probability of neonatal early-onset infection on the basis of maternal risk factors. Pediatrics 2011;128:e1155–63.
21. Escobar GJ, Puopolo KM, Wi S, et al. Stratification of risk of early-onset sepsis in newborns ≥34 weeks' gestation. Pediatrics 2014;133:30–6.
22. Kuzniewicz MW, Walsh EM, Li S, et al. Development and implementation of an early-onset sepsis calculator to guide antibiotic management in late preterm and term neonates. Jt Comm J Qual Patient Saf 2016;42:232–9.
23. Kuzniewicz MW, Puopolo KM, Fischer A, et al. A quantitative, risk-based approach to the management of neonatal early-onset sepsis. JAMA Pediatr 2017;171:365–71.
24. Strunk T, Buchiboyina A, Sharp M, et al. Implementation of the neonatal sepsis calculator in an Australian tertiary perinatal centre. Neonatology 2018;113:379–82.
25. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007;115:928–35.
26. Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. New York: Springer; 2009.
27. Romero-Brufau S, Huddleston JM, Escobar GJ, et al. Why the C-statistic is not informative to evaluate early warning scores and what metrics to use. Crit Care 2015;19:285.
28. Moons KG, Altman DG, Reitsma JB, 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.
29. Escobar G, Folck B, Gardner M, et al. Looking for trouble in all the right places: the legal implications associated with "electronic signatures" and high-risk clinical situation. In: Henriksen K, Battles JB, Marks ES, Lewin DI, editors. Advances in Patient Safety: From Research to Implementation (Volume 3: Implementation Issues). Rockville (MD): Agency for Healthcare Research and Quality (US); 2005 Feb.
30. Lapinsky SE, Hallett D, Collop N, et al. Evaluation of standard and modified severity of illness scores in the obstetric patient. J Crit Care 2011;26:535.e1–7.
31. Paternina-Caicedo A, Miranda J, Bourjeily G, et al. Performance of the obstetric early warning score in critically ill patients for the prediction of maternal death. Am J Obstet Gynecol 2017;216:58.e1–8.
32. Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. npj Digital Medicine 2018;1:18.
33. Vincent JL, Moreno R. Clinical review: scoring systems in the critically ill. Crit Care 2010;14:207.
34. Escobar GJ, LaGuardia J, Turk BJ, et al. Early detection of impending physiologic deterioration among patients who are not in intensive care: development of predictive models using data from an automated electronic medical record. J Hosp Med 2012;7:388–95.
35. Van Walraven C, Escobar GJ, Greene JD, et al. The Kaiser Permanente inpatient risk adjustment methodology was valid in an external patient population. J Clin Epidemiol 2010;63:798–803.
36. Macones GA, Hankins GD, Spong CY, et al. The 2008 National Institute of Child Health and Human Development workshop report on electronic fetal monitoring: update on definitions, interpretation, and research guidelines. J Obstet Gynecol Neonatal Nurs 2008;37:510–5.
37. Elliott C, Warrick PA, Graham E, et al. Graded classification of fetal heart rate tracings: association with neonatal metabolic acidosis and neurologic morbidity. Am J Obstet Gynecol 2010;202:258.e1–8.
38. Clark SL, Meyers JA, Frye DK, et al. Recognition and response to electronic fetal heart rate patterns: impact on newborn outcomes and primary cesarean delivery rate in women undergoing induction of labor. Am J Obstet Gynecol 2015;212:494.e1–6.
39. Clark SL, Hamilton EF, Garite TJ, et al. The limits of electronic fetal heart rate monitoring in the prevention of neonatal metabolic acidemia. Am J Obstet Gynecol 2017;216:163.e1–6.
40. Cahill AG, Tuuli MG, Stout MJ, et al. A prospective cohort study of fetal heart rate monitoring: deceleration area is predictive of fetal acidemia. Am J Obstet Gynecol 2018;218:523.e1–12.
41. Simon GE, Johnson E, Lawrence JM, et al. Predicting suicide attempts and suicide deaths following outpatient visits using electronic health records. Am J Psychiatry 2018;175:951–60.
42. Wasson JH, Sox HC, Neff RK, et al. Clinical prediction rules: applications and methodological standards. N Engl J Med 1985;313:793–9.
43. Peduzzi P, Concato J, Feinstein AR, et al. Importance of events per independent variable in proportional hazards regression analysis: II, accuracy and precision of regression estimates. J Clin Epidemiol 1995;48:1503–10.
44. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49:1373–9.
45. Japkowicz N, Shaju S. The class imbalance problem: a systematic study. Journal of Intelligent Data Analysis 2002;6:429–49.
46. Galar M, Fernandez A, Barrenechea E, et al. A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 2012;42:463–84.
47. Granich R, Sutton Z, Kim YS, et al. Early detection of critical illness outside the intensive care unit: clarifying treatment plans and honoring goals of care using a supportive care team. J Hosp Med 2016;11(suppl1):S40–7.