Human Factors and Systems Interaction, Vol. 192, 2025, 130–135
https://doi.org/10.54941/ahfe1006635

Human Factors Methods in Developing AI and Machine Learning High-Risk Prediction Models in Obstetric Care

Prithima Reddy Mosaly[1,2], Gregg Tracton[3], Medha Reddy Thummala[3], Karl Shieh[1], and Alison Stuebe[1]

[1]Department of Obstetrics and Gynecology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27514, North Carolina
[2]Ben Allegretti Consulting Inc, Stafford, VA, United State
[3]Division of Healthcare Engineering, University of North Carolina Healthcare, Chapel Hill, NC 27514, North Carolina

ABSTRACT

This study investigates the application of human factors (HF) methods to the development of artificial intelligence (AI) and machine learning (ML) models for high-risk obstetric (OB) care, focusing on integration within the Epic electronic health record (EHR) system across two hospital systems. A systematic scoping review of 39 AI/ML techniques revealed that none had achieved clinician acceptance. To address this issue, we propose a human-centered design approach, emphasizing clinical decision-making, workflow alignment, and potential maternal morbidity. Our multi-phase strategy actively engages stakeholders, including OB care providers, to refine system prototypes while considering usability, explainability, trust, and cultural sensitivity. The research aims to establish a roadmap for the future development of high-risk maternal health prediction models.

Keywords: Human systems integration, Systems engineering, Systems modelling language, Human-centered design

INTRODUCTION

OB care providers face challenges in stratifying and treating high-risk pregnancies and in predicting maternal outcomes due to complex, variable patient data (Commonwealth Fund 2019, The Pew Charitable Trusts 2020, National Partnership for Women & Families 2018). The American College of Obstetricians and Gynecologists (ACOG) standards suffer from interpretation due to drug interactions and Social Drivers of Health (SDOH) such as lack of transportation, illegal drug use, or poor finances (Howell 2018; Nattell, 2024; ACOG Committee Statement No. 11, 2024). Epic's poor usability metrics further deepen outcome issues. Artificial Intelligence (AI) and Machine Learning (ML) offers promising tools for both risk prediction and workflow optimization, but only if aligned with clinical practices (Vyas, 2020; Friedman, 2018; Ryan, 2017; Arora, 2016; Tucker, 2015). We argue that HF methods are essential to ensure these tools can be produced.

BACKGROUND

Of 734 studies that were identified from a Scoping Review of ML models predicting risk in maternal morbidity and mortality (Vasudevan, 2024), only 39 studies published sufficient details to review the quality of the predictions, and none of these studies described or evaluated clinical applications of machine learning models for providers or associated implementation factors. Therefore, even an ML algorithm that predicts risk factors perfectly is insufficient to result in a clinically acceptable tool.

To address this issue, we propose a human-centered design (HCD) approach, emphasizing clinical decision-making, workflow alignment, and potential maternal morbidity. Our multi-phase strategy actively engages stakeholders, including OB care providers, to refine system prototypes while considering usability, explainability, trust, and cultural sensitivity.

The research aims to establish a roadmap for the future development of high-risk maternal health prediction models.

METHODS

A multi-method and a multi-phase approach was employed to ensure that the tools being developed are not only effective in terms of AI/ML performance but also well-integrated into clinical workflows and widely accepted by stakeholders. The team was divided into 3 cores: Machine Learning, Clinical Translation, and Data (to harmonize OB interpretations of data originating from 3 different hospital systems). This paper focuses only on the methods and findings from phases 1 and 2 of the Clinical Translation team, not on the ML methods or data manipulations.

RESULTS

Phase 1

The literature review focused on evaluating AI/ML designs and visualizations in obstetrics (OB) care using the primary metrics: usability, trust, decision-making, mental model support, and ethics within AI systems in OB care. The review identified several studies exploring the integration of AI/ML into clinical workflows, with a particular focus on balancing trust and transparency. However, gaps were noted in the literature, especially concerning risk prediction and the evaluation of trust in AI/ML systems in the context of OB care.

Further, the stakeholder analysis highlighted key challenges such as delayed lab results, miscommunication of high-risk conditions, unexpected patient changes, and missed appointments, all of which should be integrated into AI risk prediction models to enhance accuracy and decision-making. Finally, contextual inquiry findings revealed pain points in clinical and EHR workflows, provider perceptions of AI, and areas where AI could add value, stressing the need for actionable insights, provider trust, and workload optimization.

Table 1: Phases: design architecture, data systems, and ML methods.

Phase	Objective	Methods
Phase 1: Current state assessment, and establishing system requirements and engaging Stakeholders	Assess current state of clinical workflows and identify system requirements. Engage stakeholders and review literature.	Literature review of AI/ML models in maternal health to inform tool development. Stakeholder engagement and contextual inquiry: Interviews, observations, and focus groups with OB providers to map workflows, identify pain points, and information flow issues.
Phase 2: Design prototypes of “seed ideas” for facilitating discussion and feedback	Align “seed” UI designs via human factors (HF) principles addressing the gaps and issues from Phase 1.	Participatory Co-Design: As part of part 1 work, OB care process mapping and interviews with OB providers were conducted to define workflow problems. As part of part 2, the team developed UI mockup designs (smart-SBAR, checklists, summaries) in Balsamiq.
Phase 3: Develop ML/AI Tools*	Develop low-fidelity prototypes and test AI/ML tools within clinical workflows.	Low-fidelity visualizations of synthetic patient cases for OB providers to interact with.
Phase 4: Conduct Formative and Summative Evaluation of ML/AI Tools**	Evaluate tool effectiveness and impact on decision-making, trust, and outcomes.	Conduct formative and summative evaluation of the designed tool

*Ongoing study; ** Future study.

Phase 2

Our interviews pointed to (potentially) universal truths for high-risk Obstetrics. [HF researchers should verify these in their own contexts.]

- Healthcare providers claim that there are very few rare, ambiguous, or hard-to-predict medical conditions in OB care.
- Providers recognize the importance of addressing Social Determinants of Health (SDOH) but have limited resources except for Medicaid patients.
- To reduce provider workload in group obstetrics practices, teams of clinical nurses may triage staff or patient messaging. This is not typically available in private provider clinics.
- Advance preparation is common for complex cases in clinics with high patient compliance; most providers work at home nightly to complete documentation and review charts for patients being seen the next day.

Clinicians ranked (via dot-voting) their most important daily goals (as collected from Phase 1) while using Epic. The goals with the highest importance were: communicating with the patient during an encounter (11 votes), evaluating a patient (9 votes), and documenting significant medical status (7 votes).

Clinicians then indicated specific pain points encountered within each goal. Key findings from phase 2 indicated: that providers seek streamlined documentation that aligns with clinical workflows; Epic's Notes section suffers from poor integration with other information sources, inconsistent charting, and a lack of policy-driven shared templates ("dot phrases"). The external records are often massive and disorganized, making it difficult to access critical patient data. Communication inefficiencies persist, with secure chat only useful during work hours and team-based triage lacking urgent message tracking. SDOH remain largely unaddressed due to limited social worker support, short appointment times, and lack of effective organizational support mechanisms.

Providers support AI for reducing administrative burden, summarizing medical records, and automating documentation, but they are skeptical of AI making medical decisions while still expressing a need for AI assistants which understand medical decisions. Alert fatigue is a concern, with unnecessary notifications disrupting workflow. AI could improve handoff processes and prioritize urgent communications, but clinicians resist AI involvement in bedside manner, preferring human judgment in patient interactions. While providers acknowledge AI's potential, they emphasize the need for transparency, control, and trust-building in clinical applications.

Phase 3

Our team is currently working on phase 3. The findings from phase 1 and phase 2 will guide the development of AI and ML tools designed for prototyping and testing within clinical workflows, emphasizing usability and seamless integration with existing practices. Early insights indicate that transparent and interpretable AI outputs, combined with workflow optimization, are key to enhancing clinician trust and maximizing the practical utility of these technologies. By addressing documentation inefficiencies, communication gaps, and administrative burdens, AI solutions can support clinicians in focusing on patient care while ensuring that risk prediction and decision-support tools align with real-world clinical needs. We plan to leverage an AI LLM to spur on our creativity to help design testable prototypes for phase 3.

DISCUSSION

This study, although still incomplete, suggests why even "perfect" risk predictors might still fail in clinical deployments: human factors were possibly ignored. The findings from this study emphasize the critical role of HCD in the development and implementation of AI and ML tools in OB care. Our multi-phase approach highlights the complexities of clinical workflows, provider expectations, and the need for AI solutions to be seamlessly integrated into existing EHR systems.

Phase 1 demonstrated that AI/ML tools must address specific clinical challenges, including delayed lab results, miscommunication of high-risk conditions, and inefficient documentation practices. Furthermore, providers expressed concerns regarding usability and trust, emphasizing the necessity for AI tools to be transparent, interpretable, and aligned with clinical decision-making processes.

Phase 2 suggests that as AI tools' functionality nears medical decision-making capabilities, clinicians express a decline in both interest and enthusiasm. This occurs despite multiple instances where they approve of AI demonstrating an understanding of the nuances of medical care. Through extrapolation, the authors identify a potential slippery slope: if AI can comprehend medicine, it is a small hop to imagine that AI might also be able to make medical decisions faster, more accurately, more consistently, and with greater agility (rapid learning) than humans, given sufficiently large training data. This fright of being replaced or losing control may be driving the conversation.

Additionally, Phase 2 reinforced these insights, particularly in the context of Epic EHR usability. Providers consistently identified issues such as poor integration of external records, inconsistent charting, and a lack of effective documentation tools, all of which contribute to workflow inefficiencies and frustrations. While team-based triage has improved communication, the absence of urgent message tracking in secure chat highlights the need for AI-driven enhancements in communication of messages whose urgency varies over time. Providers acknowledged the importance of addressing SDOH but cited significant barriers, including inadequate organizational support and limited access to social workers. These findings suggest that AI can play a role in mitigating administrative burdens and improving information accessibility but should not replace human oversight in complex patient care decisions.

A key takeaway from our research is the nuanced perception of AI's role in medical decision-making. While providers widely support AI for automating documentation, summarizing patient records, and prioritizing urgent messages, they remain skeptical of AI making independent medical decisions. Concerns about alert fatigue were also prominent, particularly regarding unnecessary notifications that disrupt workflow. These insights underscore the importance of designing AI solutions that prioritize usability, reduce cognitive load, and enhance—rather than hinder—clinical efficiency.

CONCLUSION

This study underscores the necessity of a HCD approach to AI and ML integration in OB care. By engaging stakeholders early in the design process, we identified critical pain points and opportunities for AI-driven improvements in documentation, communication, and workflow efficiency. Our findings highlight the need for AI tools that are not only technologically advanced but also practical, interpretable, and seamlessly integrated into clinical environments.

As we move into Phase 3, the development and prototyping of AI solutions will be guided by these insights, ensuring that tools address real-world clinical challenges while maintaining provider trust and usability. Future work will involve testing these prototypes in clinical settings, gathering provider feedback, and refining AI outputs to better support OB care providers. Ultimately, successful AI implementation in OB care hinges on designing systems that complement and enhance clinical expertise while preserving the human-centered nature of patient care.

ACKNOWLEDGMENT

The authors would like to acknowledge the AMMI grant, its members, and the dedicated clinicians who participated in interviews.

REFERENCES

Arora, K., Shields, L., Grobman W., D'Alton, M., Lappen, J., and Mercer, B. (2016). "Triggers, bundles, protocols, and checklists-what every maternal care provider needs to know," American Journal of Obstetrics and Gynecology, vol. 214, no. 4, pp. 444–451, doi: 10.1016/j.ajog.2015.10.011.

Commonwealth Fund (2019). "What is status of women's health? U. S. vs. 10 other countries," The Commonwealth fund. https://www.commonwealthfund.org/publications/issuebriefs/2018/dec/womens-health-us-compared-ten-other-countries (accessed Jul. 06, 2020).

Friedman, A., Campbell, M., Kline, C., Wiesner, S., D'Alton, M., and Shields, L. (2018). "Implementing Obstetric Early Warning Systems," American Journal of Perinatology Reports, vol. 08, no. 02, pp. e79–e84, 2018, doi: 10.1055/s-0038-1641569.

Howell, E. (2018). "Reducing Disparities in Severe Maternal Morbidity and Mortality," Clinical Obstetrics and Gynecology, vol. 61, no. 2, pp. 387–399, doi: 10.1097/GRF.0000000000000349.

National Partnership for Women & Families (2018). "Black Women's Maternal Health: A Multifaceted Approach to Addressing Persistent and Dire Health Disparities Background: Black maternal health disparities". Accessed: Jul. 06, 2020. [Online]. Available: www. NationalPartnership.org.

Nattell, N. (2024). Addressing Social and Structural Determinants of Health in the Delivery of Reproductive Health Care: ACOG Committee Statement No. 11. Obstetrics & Gynecology 144(5): p. e113–e120, November 2024. | doi: 10.1097/AOG.0000000000005721.

Ryan H., et al., (2017). "Validating the Performance of the Modified Early Obstetric Warning System Multivariable Model to Predict Maternal Intensive Care Unit Admission," Journal of Obstetrics and Gynaecology Canada, vol. 39, no. 9, pp. 728–733.e3, 2017, doi: 10.1016/j.jogc.2017.01.028.

The Pew Charitable Trusts (2020). "Preventable Maternal Deaths Continue to Occur in the U. S." https://www.pewtrusts.org/en/research-and-analysis/articles/2020/01/06/preventable-maternaldeaths-continue-to-occur-in-the-us (accessed Jul. 06, 2020).

Vasudevan, Lavanya, et. al (2024). "Machine Learning Models to Predict Risk of Maternal Morbidity and Mortality from Electronic Medical Record Data: A Scoping Review", Submitted to JBI Manual for Evidence Synthesis. https://doi.org/10.46658/JBIMES-24-01

Vyas, D., Eisenstein, L., and Jones, D. (2020). "Hidden in Plain Sight — Reconsidering the Use of Race Correction in Clinical Algorithms," New England Journal of Medicine, Jun. 2020, doi: 10.1056/nejmms2004740.