Implementation Challenges and Approaches for Rule-Based and Machine Learning-Based Sepsis Risk Prediction Tools: A Qualitative Study

Scholarly Report submitted in partial fulfillment of the MD Degree at Harvard Medical School

Date: 25 January 2020

Student Name: Mugdha Joshi, BS, BA

Scholarly Report Title: Implementation Challenges and Approaches for Rule-Based and Machine Learning-Based Sepsis Risk Prediction Tools: A Qualitative Study

Mentor Name(s) and Affiliations: Lipika Samal, MD, Dept of General Internal Medicine, Brigham and Women's Hospital

Collaborators, with Affiliations: Keizra Mecklai, Harvard Medical School Class of 2022

ABSTRACT

Title: Machine Learning-Based and Rule-Based Sepsis Risk Prediction Tools: A Qualitative Study of Implementation Challenges and Approaches

Background: Mandated reporting of sepsis outcomes have led many institutions to implement surveillance software to improve sepsis outcomes. Commercial EMRs, external vendors, and home grown risk prediction tools offer a variety of approaches. Traditional rule-based models draw on the Systemic Inflammatory Response Syndrome (SIRS) criteria while newer predictive models utilize machine-learning (ML) based algorithms to predict sepsis risk. The purpose of this study is to identify challenges and approaches for successful implementation of sepsis surveillance tools.

Methods: Semi-structured interviews were conducted with hospital leaders overseeing sepsis clinical decision support implementation at U.S. medical centers (n=14). Participants were recruited via purposive sampling. Interviews probed implementation process, challenges faced, and recommended approaches. Responses were independently coded by two coders with consensus approach and inductively analyzed for themes.

Results: Challenges shared by institutions with both SIRS and ML models categorize to technical build, optimization of alerts, workflow integration, tool validation, implementation time, and working with external vendors. Institutions using ML models reported greater difficulty with clinician acceptance of these tools due to user expectation management, limited tool intuitiveness, distrust in the technology, and confusion. Successful institutions report multiple approaches to improving acceptance including user education, expert support, and practitioner-led efforts.

Conclusion: In this small but diverse set of hospitals, we found that in addition to the known socio-technical challenges of implementing clinical decision support, less clinically intuitive ML models may require additional attention to user education, support, and expectation management.

TABLE OF CONTENTS

GLOSSARY OF ABBREVIATIONS

ML: Machine-learning, a subset of artificial intelligence based algorithms that rely on statistical inferences derived directly from a data set rather than direct programming using rules established by a human.

RB: Rule-based, algorithms that perform a task by applying a set of rules that have been programmed by a human.

SEP-1: Sepsis bundled payment implemented by Centers for Medicare and Medicaid Services in October 2015

SIRS: Systemic Inflammatory Response Syndrome, includes a set of clinically defined criteria used to identify the presence of sepsis in the definition of sepsis preceding the Sepsis 3 guidelines published in 2016.

INTRODUCTION

Sepsis affects 1.7 million adults in the United States each year, potentially contributes to more than 100,000 deaths, and is one of the most expensive causes for hospitalization[1]. It is well established that early resuscitation and antibiotics administration improve mortality[2]. It is also well established that sepsis is challenging to diagnose, with wide variability between clinicians[3]. To add to the challenge, clinical definitions for sepsis have been under review and changed most recently in 2016[4].

As of October 2015, sepsis recognition and care has become a Centers for Medicare and Medicaid Services priority with the implementation of a sepsis bundled payment (SEP-1). The SEP-1 bundle expects clinicians to measure serum lactate, obtain blood cultures, and administer broad spectrum antibiotics within 3 hours of sepsis onset or hospital presentation[5]. The bundle has led many hospitals to invest significant resources into sepsis care and data collection[6]

Health Information Technology (IT) tools have been an essential part of hospital efforts to comply with bundle requirements, offering real time monitoring and decision support. These tools have traditionally been rule-based, meaning that they rely on logic that comes from human clinical experts and clinical guidelines. The change in clinical definitions for sepsis in 2016 left many of these rule-based surveillance systems outdated[4]. Additionally, traditional early warning systems using SIRS criteria have been criticized for producing too many false positive alerts[2].

In contrast to traditional rule-based models, machine learning (ML) algorithms use associations derived from large data sets to make predictions and can often incorporate many more variables. ML models have demonstrated improved diagnostic accuracy using a variety of proprietary algorithms. For example, SIRS and other clinical sepsis assessment tools have a sharp drop off between sensitivity and specificity (0.72, 0.44) with the ML algorithm showing more balance between these constraints (0.8, 0.8)[7]. There is potential for such algorithms to reduce false positives, identify sepsis earlier to enable earlier intervention, and to improve outcomes. However, ML algorithms are new to clinical practice and raise practical, ethical, and trust based concerns.

Given the challenge of the clinical problem at hand, the policy incentives, and the potential for machine learning to add value to the status quo, it is not surprising that the market for clinical decision support options is diverse and growing. Leaders looking to health IT for improvement in their sepsis outcomes can choose from 3rd party ML applications such as

Dascena InSight, Hiteks Vigilant QA, Rothman Index, and POC Advisor, ML models produced by commercial EMR vendors including Epic and Cerner, rule-based best practice advisory toolkits from commercial vendors, and home grown ML and rule based solutions alike.

Sepsis is a significant source of mortality in this country. Financial incentives and quality improvement culture are aligning to encourage all hospitals to improve their sepsis outcomes. The market is ripe with health IT tools with potential to improve capability for proven clinical interventions such as early resuscitation and antibiotics administration. In order to understand the current state of health IT utilization for sepsis prediction and the barriers to institutional uptake, we aim to explore the implementation challenges faced by leaders working on operationalizing sepsis prediction health IT tools and the strategies that have allowed for success.

STUDENT ROLE

My role consisted of developing the research question, researching qualitative research methodology with mentoring from Dr. Samal, proposing a methodology, obtaining IRB exemption for the study, conducting interviews, recruiting a medical student assistant to help with data coding, conducting a consensus conference, developing a code book, completing coding of all interviews, analyzing the data for the themes presented in this work, presenting a poster representing this work at the SGIM North East Regional meeting, preparing a manuscript for submission, and leading ongoing work to gather additional data to support publication of this work.

METHODS

Choice of Study Design

An exploratory qualitative design was chosen for this study because little is known about this topic. A valid survey instrument can only be developed once there is a body of knowledge from which to craft questions. This study was considered to be IRB exempt by the Harvard Longwood IRB.

Participant Recruitment

Institutions were recruited purposively. We conducted semi-structured phone interviews with 14 hospital leaders overseeing the implementation of a sepsis or clinical deterioration predictive tool at 11 hospitals across the country. The first round of recruitment involved contacting hospitals who were mentioned in publicly available press releases for three third party sepsis prediction tools. 4 out of 18 institutions responded to the invitation to interview. The remaining 7 institutions were recruited via calls for participation sent on informatics email lists. Institutions were asked to suggest the person directly overseeing the implementation of the tool for interview. 3 out of 11 interviews were simultaneously conducted with more than one individual for a total of 14 participants.

Interview Guide and Survey Development

Participants were interviewed about their role at their institution, background of the institution's sepsis prediction efforts, the stakeholders involved in the process, process challenges, and successful approaches. All participants were asked to complete a survey about their institutional characteristics developed and distributed via Qualtrics.

Analysis

Interviews were recorded and transcribed. The first 5 transcripts were coded line-by-line independently by two coders. We applied grounded theory consensus approach to develop a thematic codebook from these 5 initial interviews with operational definitions for 14 codes (motivation, technical build, alert optimization, workflow integration, tool validation, interdisciplinary collaboration, clinician acceptance, contracting logistics, data privacy, expert support, implementation time, institutional culture, motivation, persistence) and 3 designations (challenge, approach, quote). Once developed, codes were assigned to complete thoughts for all interview transcripts independently by both coders. Interrater discrepancies were reconciled every 2 interviews with iterative adjustment to operational definitions of codes.

RESULTS

Characterizing the Study Sample

The study captured a set of institutions heterogenous in size, location, patient population, EMR vendor, and current sepsis prediction solution in use (Table 1). Furthermore, the individuals interviewed, designated by their institutions as the primary overseer for the implementation effort, carried a wide variety of titles. Of note, 8 out of 11 institutions tried a machine learning approach of some type. Of those 8, 5 are using the tool in practice, 2 are running the tool only in the background, and one abandoned the tool entirely.

Motivations for Implementing Sepsis Prediction Tools

Quality improvement for septic patients is the primary driver for institutions deciding to undertake a sepsis prediction tool initiative. Institutions describe turning to prediction tools in order to "enable earlier intervention", be able to "follow trajectories", and improve how sepsis can be "identified in data-driven ways", given how "clinically complex it is to define". Multiple institutions described being specifically motivated by sepsis outcomes that were inferior to other regional hospitals.

Hospitals for the most part preferred to use tools supplied by their EMR vendor or home grown tools that are integrated into their EMR over third party tools. Several who experimented with a third party tool ultimately chose not to use it due to difficulties with external contracting logistics, distrust of vendors, lack of customizability, and difficulty with integration into existing workflows. One participant described the choice as follows, "Either you purchase a program through your EHR vendor, or you try to build it yourself, or you purchase a third-party solution and hope that they are not lying to you. Or you know putting lipstick on the pig. Or you know just making it sound better, oh we're going to do this for you and it's going to be amazing right."

Challenges Shared by Institutions Implementing Rule Based and ML Models

Implementation challenges that emerged from the interviews clustered to technical build, optimization of alerts, workflow integration, tool validation, implementation time, working with external vendors, and clinician acceptance (Table 2). These were shared by rule based and ML models alike.

The Challenge of Clinician Acceptance

Within clinician acceptance, both groups reported difficulties stemming from overburdening of clinicians. However, subthemes of user expectation management, limited tool intuitiveness, distrust in the technology, and confusion emerged in discussions of ML models that were not shared by discussions of rule-based models (Table 3).

Successful Approaches to Promote Clinician Acceptance

Institutions fostered acceptance of ML models among their clinicians by using data demonstrating successful outcomes to increase buy-in, creating accessible support structures for users, creatively educating users about what to do with outputs of models, and incorporating practitioners in implementation efforts (Table 4).

DISCUSSION

In response to national policy shift and QI initiatives many institutions have started implementation efforts. Our study shows that these efforts are broad and heterogeneous, each requiring significant activation energy on part of the institution in order to implement. While ML models are entering the landscape, rule-based models remain tried and true. Not all places that have tried a rule-based model have tried a machine learning model, but in general those that have tried ML models have previously tried rule-based models. Many have returned to rule-based models after having failed to successfully implement a third party or commercial vendor offered ML tool.

The major themes that have emerged as challenges for implementation are technical build, alert optimization, workflow integration, external contracting, clinician acceptance, implementation time, and tool validation. There appears to be a difference between how people talk about the challenges of clinician acceptance in rule-based models and ML models. Both discuss challenges with alert optimization and workflow integration in order to not overburden physicians, however institutions with ML models contend with challenges of confusion, trust, expectation management, and intuitiveness. Institutions have employed creative methods to try and address these challenges through use of data, providing accessible support, educating users, involving providers, and explicit management of expectations.

The challenges emerging in this study align with the known challenges proposed by Sittig and Singh in their sociotechnical model[8]. Each challenge is individually rich with sub-themes to explore for challenges that exist and the solutions teams have used to overcome them. Despite work building on this framework for close to a decade, there continues to be considerable decentralized effort at the institutional level to overcome these challenges at each individual hospital.

The clinician acceptance challenges that emerged with the machine learning models have been predicted in the literature with descriptions of the unacceptability of the "black box" of machine learning in clinical medicine.[9] With rule based models, explanations are more straightforward since the rules are derived from expert clinicians. However, in the ML models, demystifying the "black box" is critical for establishing trust in the model[10]. There is a growing body of work in machine learning literature that explores the components of effective explanations of models and predictions so that they can be used easily by users without in depth machine learning backgrounds[10–14]. Importantly, there is a difference between explaining a model by naming the variables involved in a prediction and explaining a prediction such that it is readily intuitive for the user[13]. In the sites that we studied, the onus of packaging and displaying that score in a usable way to providers is done on an institutional level. While unavoidable for sites with home grown models, this required reinvention for each new implementation even for sites with a model supplied by a commercial vendor.

Current literature supports that in order to trust a predictive model, a user must trust an individual prediction enough to act on it and trust in the model to work with fidelity[10]. However, this study also suggests that separate from trusting the output of a model, users require knowledge of how to make that prediction clinically meaningful. While rule based models are drawn from criteria that are taught in clinical training, a prediction from a machine learning model that forecasts that a patient may have sepsis is a new type of data that clinicians have not been trained to work with. The strategies focusing on user education, accessible support, and management of expectations have likely been successful because they address this need.

Per diffusion of innovation theory, the more complex something is to use, the less likely it is to be accepted.[15] Unfamiliarity with the new data and associated new processes makes individual acceptance of this innovation challenging. With the onus for intuitive packaging of model output, user support, and user education falling to individual institutions, there is complexity on the institutional level potentially hindering effective diffusion of this innovation.

One limitation of this study is the small sample size of participants that were interviewed. While size was limited by the number of individuals who responded to the invitation to participate, a wide net of institutions was captured by the recruitment strategy that were heterogeneous in geography, size, patient population, and academic orientation. The data captured in this study provides the insight that people are turning to predictive analytics for quality improvement even in community hospitals without academic informatics departments. Generalizability of this study is also potentially limited by nonresponse bias if struggling or successful institutions differed in their likelihood to respond to the invitation to participate. The variety of model types and stages of implementation suggests that even within the institutions sampled, there was difference in ease of implementation based on circumstances of individual institutions, nevertheless the challenges that emerged were consistent throughout the interviews.

As complex as these undertakings have been, it has not been established that use of ML predictive models compared to rule-based models improves sepsis outcomes. Without successful implementations, such a comparison cannot be established. The question should be asked, is this worth it? And if the answer is yes, there is much work to be done in facilitating organizations to overcome many of these prohibitive barriers.

Other next steps stemming from this work are to better characterize the elements contributing to confusion, distrust, and expectation management among providers and propose a framework that institutions can use to develop user support and education that can address these elements. Further study could examine the vendor perspective on offering user support and education with products that they design. Paranjape et al recently proposed inclusion of critical appraisal of AI technology in pre-clinical curricula, relevant application of AI technologies in clinical training, and incorporation of fundamentals of AI and data science into licensing examinations and continuing medical education.[16] The findings from the interviews conducted in this study suggest that the merit of considering education reform recommendations like these is not only theoretically valuable, but relevant to actual needs of current practice.

Summary of Recommendations:

- Assessment of outcomes with implementation of ML tools for sepsis prediction to determine if the complexity of this undertaking is justified

- Emphasis on explanation of the prediction as it applies to an individual patient
- Centralized suggestions from vendors for how model outputs should be packaged and delivered to users to facilitate diffusion to new sites
- Consideration of curricular reforms that emphasize data science and utilization of predictive tools at all levels of medical education.

CONCLUSION

In this small but diverse set of hospitals, we find broad heterogeneity in institutional application of predictive analytics to improve sepsis outcomes. Implementation of predictive models is time consuming and complicated, with the job of making predictive output of ML models clinically usable falling largely to individual institutions. While there are shared socio-technical challenges of implementing clinical decision support for both rule-based and ML models, attention to user education, support, and expectation management and dissemination of best practices related to these areas may improve feasibility and effectiveness of ML models being used in quality improvement efforts.

ACKNOWLEDGEMENTS

A huge thank you to Dr. Lipika Samal who has been my primary mentor throughout this entire project, has believed in it when I did not, and pushed me to feel like I could take something from start to finish and figure things out along the way. I'd also like to acknowledge Keizra Mecklai a student in the class of 2022 for her enthusiasm, help with coding and data analysis, and for her many insights along the way.

REFERENCES

1. Rhee C, Dantes R, Epstein L, et al. Incidence and Trends of Sepsis in US Hospitals Using Clinical vs Claims Data, 2009-2014. *JAMA*. 2017;318(13):1241-1249. doi:10.1001/jama.2017.13836
2. Semler MW, Weavind L, Hooper MH, et al. An Electronic Tool for the Evaluation and Treatment of Sepsis in the ICU: A Randomized Controlled Trial. 2016;43(8):1595-1602. doi:10.1097/CCM.0000000000001020.An
3. Rhee C, Kadri SS, Danner RL, et al. Diagnosing sepsis is subjective and highly variable: a survey of intensivists using case vignettes. *Crit Care*. 2016;20:89. doi:10.1186/s13054-016-1266-9
4. Abraham E. New Definitions for Sepsis and Septic Shock: Continuing Evolution but With Much Still to Be DoneNew Definitions for Sepsis and Septic ShockEditorial. *JAMA*. 2016;315(8):757-759. doi:10.1001/jama.2016.0290
5. Klompas M, Rhee C. The CMS Sepsis Mandate: Right Disease, Wrong Measure. *Ann Intern Med*. 2016;165(7):517-519. doi:10.1111/j
6. Walkey AJ, Lindenauer PK. Keeping It Simple in Sepsis Measures. *J Hosp Med*. 2017;12(12):1019-1020. doi:10.12788/jhm.2873
7. Desautels T, Calvert J, Hoffman J, et al. Prediction of Sepsis in the Intensive Care Unit With Minimal Electronic Health Record Data: A Machine Learning Approach. *JMIR Med informatics*. 2016;4(3):e28. doi:10.2196/medinform.5909
8. Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. *Qual Saf Heal Care*. 2010;19(Suppl 3):i68-i74. doi:10.1136/qshc.2010.042085
9. Braunstein ML. *Health Informatics on FHIR: How HL7's New API Is Transforming Healthcare*.; 2018. doi:10.1007/978-3-319-93414-3
10. Tulio Ribeiro M, Singh S, Guestrin C. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *arXiv e-prints*. February 2016:arXiv:1602.04938. https://ui.adsabs.harvard.edu/abs/2016arXiv160204938T.
11. Lundberg SM, Nair B, Vavilala MS, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nat Biomed Eng*. 2018;2(10):749-760. doi:10.1038/s41551-018-0304-0

12. Lundberg SM, Lee S-I. A Unified Approach to Interpreting Model Predictions. In: Guyon I, Luxburg U V, Bengio S, et al., eds. *Advances in Neural Information Processing Systems 30*. Curran Associates, Inc.; 2017:4765-4774. http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf.
13. Tajgardoon M, Samayamuthu MJ, Calzoni L, Visweswaran S. Patient-Specific Explanations for Predictions of Clinical Outcomes. *ACI Open*. 2019;03(02):e88-e97. doi:10.1055/s-0039-1697907
14. Caruana R, Kangarloo H, Dionisio JD, Sinha U, Johnson D. Case-based explanation of non-case-based learning methods. *Proceedings AMIA Symp*. 1999:212-215.
15. Rogers EM, Singhal A, Quinlan MM. Diffusion of innovations. In: *An Integrated Approach to Communication Theory and Research, Third Edition*. ; 2019. doi:10.4324/9780203710753-35
16. Paranjape K, Schinkel M, Nannan Panday R, Car J, Nanayakkara P. Introducing Artificial Intelligence Training in Medical Education. *JMIR Med Educ*. 2019;5(2):e16048. doi:10.2196/16048

TABLES AND FIGURES

Table 1: Heterogeneity of Samples

	Size	**Type**	**Region**	**Adult/Pediatric**	**EMR Vendor**	**Current Solution**	**Title of Survey Respondent**
1	<300	Community	West	Adult	VistA	3rd party ML	CMO
2	<300	Community	NE	Adult	MEDITECH	Rule Based	CNO
3	>500	Community	MW	Adult	Epic	EMR ML	1. System Epidemiologist 2. Business Intelligence Manager
4	300-500	Academic	NE	Adult	Epic	Rule Based	Associate CMIO
5	>500	Community	MW	Adult	Epic	EMR ML	CMIO
6	>500	Academic	MW	Adult	Epic	3rd party ML	Executive Director of Clinical Operations
7	>500	Academic	NE	Adult	Epic	Home Grown ML	Senior Director, Clinical Operations
8	>500	Academic	MW	Adult	Epic	EMR ML	1. Director of Clinical Effectiveness

							and Informatics 2. Clinical Informatics Lead
9	>500	Community	SE	Adult	Epic	Rule Based	CMIO
10	300-500	Academic	NE	Pediatric	Cerner	Rule Based	1. CMIO 2. Director of Critical Care
11	300-500	Academic	NE	Pediatric	Epic	Home Grown ML	Emergency Department Director of Clinical Care

Table 2: Challenges Common to ML and Rule Based Algorithms

Theme	Representative Quotation
Alert Optimization	"We also decided how much of a wide or a narrow net...for looking at the surveillance. Would you want to catch…the widest net possible which would mean you'd have a lot of false positives or did you want to meet more specific in which case you would have less false positives but you might also have more missed opportunities." (RB) "So it would be great, you know everybody wants this screening test that is perfectly sensitive and a confirmatory test that is perfectly specific but that is not always possible." (ML)
Technical Build	"It's kind of an iterative process because this particular model…I don't know there's probably 30 parameters or so, I have to double-check, but it requires configuration, so at a high level the model is using demographic information, medications, lab results" (ML)

	"The other difficulty in general is the definition or gold standard for sepsis makes it hard for…any standard but machine learning in particular because it is much easier…if you have true positives and true negatives. And then the ambiguous cases can be used for learning but here the ambiguous cases, often nobody really knows what to do with them."(ML) "And I think the biggest issue was at the discrete data points…you want to know if you have a pneumonia…there's no discrete data field that says pneumonia in it… so it is a little bit more difficult to pull some of those discrete data elements when you're doing the data mining unless you have some sort of artificial intelligence that can read words and recognize terms...synonymous with pneumonia. So that's where I think everybody's biggest challenge is in trying to figure out how to make that work right." (RB)
Workflow Integration	"But I don't think any of these, so you can have the technology, but you still have to build the workflow around the technology. I don't think any of them are totally plug and play. That play is going to depend on a lot of other factors."(ML) "I think a real challenge has been figuring out who…get(s) the alert and how they were going to get it and what we were going to do in cases where a patient triggered an alert and one hadn't yet assigned themselves to the patient so I think some of those logistical challenges of making sure the alert has gotten the attention of the provider but didn't disrupt the provider. That at least for me was what we found the most challenging." (RB)
External Contracting	"We generally like to do things as much as possible within our EHR without involving third-party vendors and so that is what we looked to do." (RB) "And because we have Epic, because there was no additional cost to implement their method, this is in all honesty, it was determined that that could be where we could start." (ML)

Tool Validation	"I think we are challenged generally in evaluating our decision support efficacy…where we struggle the most is looking at the patient outcome once the system's triggered just because it's hard to track patients like that. But we're certainly able to track when it triggered, what did someone do. What did someone order, what's the turnaround time from the order to the administration of whatever they ordered, and then the part where we have trouble is figuring out did the patient actually become septic, did they get transferred to the ICU, did they walk out of the hospital, or did they die, things like that." (RB) "The biggest delay is usually anytime we have to make a change to the model, like recently we had to make a change to one of the parameters in this model because there are no values, and we didn't have the correct values for the parameter. So every time we make a change, we have to let it run for another 6 weeks in order to start using the data and to validate it." (ML)
Implementation Time	"We had a 6 month project which ultimately took like 3 years. Yeah, it was years on the design, and then I still remember the email saying our first "go live" date was going to be August 8, 2017 and our actual go live date ended up being June 7, 2018 so it's almost an entire year between when we thought we were going to be ready to go live and when we actually did." (RB) "When we were doing the homegrown model, that process actually took about a year-and-a-half. Because we would create the model, our own internal model, put it into the system, have it run silently in the background. We would notice some odd behavior or something that we wanted to tweak, we would tweak it put it back in the system, watch its behavior for a few weeks to months and then keep adjusting. It took a long time for us to get to the point where we felt comfortable enough to put it out." (ML)

Clinician Acceptance	"if you overwhelm providers with warnings, they'll ignore them. So many alerts are false positives so we're trying to find the correct balance..." (RB) "If you have a lot of false positives you are at risk for alert fatigue and if a positive alert is embedded in a workflow that involves some amount of time and attention, doing too many of those false ones, then people are going to get very frustrated and abandon the effort entirely so I think the biggest challenge is getting to that sweet spot of the sensitivity and specificity of the alert and aligning it with a workflow that is practical and that bedside nurses and providers are actually going to be able to do."(ML)

Table 3: Sub-Themes Related to Clinician Acceptance in Machine Learning Models

Sub-Theme	Representative Quote
Expectation Management	"I think if you sort of set the expectation that this tool is going to come in and find sepsis for you, you are going to set yourself up for significant disappointment."
Tool Intuitiveness	"Well my major, major bone of contention with the algorithm is that the endpoint that they are predicting is not clinical. I understand that it's easier for them…for me the buyer and the clinician, for this to work I need a clinical endpoint."
Distrust in Technology	"I am not so thrilled with the predictive model. I don't understand…it's a big black box…. I don't know who built it, I don't know what state, the four hospitals are mystery hospitals."
Confusion	"People didn't really understand it. They were touting it as an artificial intelligence and people didn't understand what that was."

Table 4: Successful Approaches to Improve Clinician Acceptance of Machine Learning Models

Approach	Representative Quote
Use of Data	"Just showing people…giving them visual patterns of…succeeding or failing is a powerful tool"
Accessible Support	"we created a resource through the virtual care team that allowed nursing staff, provider staff to call anytime 24/7"
User Education	"[We created a video] comparing predictive models to a weather forecast. It doesn't mean you're going to put the rainboots on now because it's not raining right now."
Practitioner led efforts	"I think the fact that we as the clinical effectiveness team are clinicians, I think really helps."

Managing Expectations	"[We] have to manage expectations that we are not yet at a point where these rules are going to be able to define sepsis and without help from humans…"