



Investigating Why Clinicians Deviate from Standards of Care: Liberating Patients from Mechanical Ventilation in the ICU

Nur Yildirim
Carnegie Mellon University
Pittsburgh, PA, USA
yildirim@cmu.edu

Susanna Zlotnikov
Carnegie Mellon University
Pittsburgh, PA, USA
susanna.zlotnikov@gmail.com

Aradhana Venkat
Carnegie Mellon University
Pittsburgh, PA, USA
aradie.v@gmail.com

Gursimran Chawla
Carnegie Mellon University
Pittsburgh, PA, USA
simchawla11@gmail.com

Jennifer Kim
Carnegie Mellon University
Pittsburgh, PA, USA
jk4@andrew.cmu.edu

Leigh A. Bukowski
University of Pittsburgh
Pittsburgh, PA, USA
lab108@pitt.edu

Jeremy M. Kahn
University of Pittsburgh
Pittsburgh, PA, USA
jeremykahn@pitt.edu

James McCann
Carnegie Mellon University
Pittsburgh, PA, USA
jmccann@cs.cmu.edu

John Zimmerman
Carnegie Mellon University
Pittsburgh, PA, USA
johnz@cs.cmu.edu

ABSTRACT

Clinical practice guidelines, care pathways, and protocols are designed to support evidence-based practices for clinicians; however, their adoption remains a challenge. We set out to investigate why clinicians deviate from the “Wake Up and Breathe” protocol, an evidence-based guideline for liberating patients from mechanical ventilation in the intensive care unit (ICU). We conducted over 40 hours of direct observations of live clinical workflows, 17 interviews with frontline care providers, and 4 co-design workshops at three different medical intensive care units. Our findings indicate that unlike prior literature suggests, disagreement with the protocol is not a substantial barrier to adoption. Instead, the uncertainty surrounding the application of the protocol for individual patients leads clinicians to deprioritize adoption in favor of tasks where they have high certainty. Reflecting on these insights, we identify opportunities for technical systems to help clinicians in effectively executing the protocol and discuss future directions for HCI research to support the integration of protocols into clinical practice in complex, team-based healthcare settings.

CCS CONCEPTS

• Human-centered computing → Field studies.

KEYWORDS

healthcare, field study, standards of care, adherence, ventilation

ACM Reference Format:

Nur Yildirim, Susanna Zlotnikov, Aradhana Venkat, Gursimran Chawla, Jennifer Kim, Leigh A. Bukowski, Jeremy M. Kahn, James McCann, and John

Zimmerman. 2024. Investigating Why Clinicians Deviate from Standards of Care: Liberating Patients from Mechanical Ventilation in the ICU. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3613904.3641982>

1 INTRODUCTION

Healthcare advances result from carefully controlled, randomized, double-blind studies. Actions and interventions that prove successful in these trials get captured as clinical guidelines, care pathways, and protocols. These structures provide evidence-based scaffolding for clinicians in support of their practice, enabling the results of clinical trials to improve health outcomes at the population level. While protocols improve patient outcomes and quality of patient care, their implementation and adoption by frontline care providers remains a persistent challenge [34, 99]. Protocols are produced in carefully controlled environments; hence, they often do not account for the messy complexity of day-to-day clinical practice [34], creating barriers to adoption. Specifically, protocols often cannot account for the innumerable complexities of healthcare in the real world in which care needs to adapt to specific contexts and available resources [53]. Recent research shows that changes to information technology (IT) systems may offer one way to increase adherence to protocols [4].

In this study, we set out to investigate the adherence to clinical protocols with a specific focus on opportunities where technology might aid protocol adoption. We used mechanical ventilation in the intensive care unit (ICU) as a model system, enabling us to integrate our complementary perspectives as human-computer interaction (HCI) researchers and critical care medicine practitioners. Mechanical ventilators save lives by taking over the work of breathing, so a patient in the ICU can rest and heal [42]. It is important to liberate patients from the ventilator as soon as possible, while also taking care not to liberate them too soon, before their lungs have had a chance to heal. Achieving this balance is a complicated and high-risk endeavor, requiring the careful consideration and coordinated



This work is licensed under a Creative Commons Attribution-NoDerivs International 4.0 License.

CHI '24, May 11–16, 2024, Honolulu, HI, USA

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0330-0/24/05

<https://doi.org/10.1145/3613904.3641982>

actions of a team of clinicians, including a critical care physician, an ICU nurse, and a respiratory therapist (RT).

To aid this process, medical researchers created the *Wake Up and Breathe* (WUB) protocol, an evidence-based guideline for assessing mechanically ventilated patients and deciding when they should be liberated [55]. The protocol involves two sequential, interdependent actions. First, the ICU nurse assigned to a specific patient performs a “Spontaneous Awakening Trial” (SAT). They cut off a patient’s sedation to observe if they can tolerate being awake. Next, while the patient is free from sedating medications, an RT performs a “Spontaneous Breathing Trial” (SBT). They cut off the ventilator’s breathing support to observe if the patient can successfully breathe on their own without mechanical support. Based on the results of these two assessments, the physician, RT, and nurse collectively decide whether a patient should be extubated (have a breathing tube removed), liberating them from the ventilator [109].

Clinical trials show that application of WUB dramatically improves patient outcomes, reducing the average number of days a patient remains on a ventilator and decreasing mortality [37, 55]. Unfortunately, the protocol is often not followed in clinical practice. We chose to conduct a field study to investigate the current workflows and practices for mechanically ventilated patients. Our goal was twofold:

- (1) To gain a deeper understanding of the barriers to WUB adherence; to help explain how, when, and why clinicians deviate.
- (2) To identify opportunities where technology might effectively remove barriers and increase adherence to standards of care.

We conducted observations and interviewed clinicians who work at three different medical ICUs within a large, integrated health system. Our findings indicate that unlike what prior literature suggests, clinicians do not seem to consciously avoid executing WUB in relation to specific barriers. Instead, the uncertainty surrounding individual patients’ eligibility and the role of individual providers within the context of the interprofessional care team seems to make clinicians ‘procrastinate’. They unconsciously deprioritize WUB by focusing on other tasks where they have high certainty, and this causes them to miss the window for performing WUB. Specific sources of uncertainty included which patients might be eligible, when an RT might arrive to perform SBT, if the physician wants WUB to happen, and what counts as a sedative. Reflecting on these insights, we identified several opportunities where new technical systems might help. In addition, we observe that the challenge of uncertainty and adherence appears to be a problem that goes beyond liberation from mechanical ventilation and beyond clinical practice in the ICU.

This study makes two novel contributions to the HCI literature. First, our field study provides a rare, first-hand description of nurses’ and RTs’ clinical workflows for providing care to mechanically ventilated patients in the ICU, a topic of immense public health importance, particularly in the wake of the COVID-19 pandemic. We deepen the understanding of the barriers to WUB adherence with particular attention to clinicians’ information needs, coordination needs, and overall uncertainty. Second, we identify novel opportunities for computational and artificial intelligence (AI) approaches to help clinicians effectively execute the patient liberation protocol by emphasizing the ‘how’ of protocol implementation,

rather than simply the ‘what’. Based on these advances, we discuss design implications and present future research directions for better integrating protocols and guidelines into clinical practice beyond the context of mechanical ventilation in the ICU.

2 RELATED WORK

2.1 HCI Research on Intensive Care

As an information-rich environment, the ICU has attracted the attention of HCI researchers over the past few decades. Early ethnographic research investigating the clinical practice in the ICU noted the cyclic and temporal nature of activities: daily routines, such as patient rounds, helped clinicians plan and coordinate their tasks and overall patient care [81]. Several studies highlighted the importance of presenting information in a temporal context, showing not only current activities, but also past and future activities that present the overall patient trajectory [61, 84]. Other research studied workflows around specific tasks, such as how physicians and nurses form volume therapy decisions [50] or how nurses document patient notes and their workarounds to understand the limitations of electronic health records (EHR) [29]. Based on this understanding of clinical practice, researchers designed and evaluated the use of new technical systems to support documentation and note-taking [108], reduce alert fatigue and interruptions [67, 97], and support patient handovers [93]. Our prior research broadly explored the AI innovation opportunities in intensive care [120], revealing a need for intelligent systems that support clinicians in performing evidence-based practices for patient liberation. Building on this line of work, this paper seeks to understand the clinical workflows of nurses and RTs around WUB to inform the design of future technologies.

2.2 Coordination in Clinical Settings

Supporting collaboration and teamwork within complex health-care settings has been a major research interest within the HCI and *Computer-Supported Collaborative Work* (CSCW) literature [33]. Patient care is often delivered by multiple care providers that coordinate across time and space [6, 21]; however, coordination between care team members remains challenging [2]. Prior work studied coordination in time- and safety-critical settings, such as surgical ward [6, 7, 92], trauma resuscitation room [90], emergency department [63], and intensive care [50]. Researchers highlighted the role of artifacts and *common information spaces* [21, 45], such as paper schedules, electronic records, whiteboards and interactive displays, in facilitating coordination and collaborative sensemaking [8, 14, 77, 80, 101, 112].

Early research viewed the shift from paper-based medical records to electronic records as an opportunity for technology to support collaboration between care team members by providing role-based, shared representations of care plans [81–84]. However, recent studies report that EHR systems lack collaborative affordances, making it difficult for clinicians to obtain the type of holistic understanding of the patient that is necessary to coordinate their activities [9, 51, 98]. Researchers note that current systems do not support team-based, shared care plans: team members often cannot access all of the patient information, and care goals are often set separately rather than as a team [2]. As a result, many clinicians often view

the EHR as a hindrance to collaboration, rather than an enabler [10, 28, 75]. Bardram and Houben [9] highlight four design implications for medical records to facilitate collaboration: being portable across patient wards and the hospital, providing collocated access, providing a shared overview of medical data, and giving clinicians ways to maintain *mutual awareness* – the ability to perceive each others’ activities and relate them to a joint context [86]. We build on prior research on care coordination by investigating the use of EHR and other systems in the ICU for performing WUB – a non-emergency protocol that requires the sequential and coordinated actions of care providers.

2.3 Adherence to Clinical Guidelines

The effective translation of clinically validated protocols and guidelines into daily clinical practice is considered one of the greatest challenges in evidence-based medicine [34, 99]. A large body of healthcare literature has investigated the barriers to adherence in the context of ICU, including the use of guidelines for sepsis [99], patient nutrition [25], patient liberation from ventilators [24], and others (e.g., [17, 40, 107]). These studies use a variety of methods, but typically approach the subject through a psychological lens, exploring barriers to behavior change on the part of the individual providers. Interestingly, the reported barriers using these methods are highly consistent across the literature, regardless of the specifics of each guideline. The barriers mainly include: lack of time and staffing, unfamiliarity with guidelines, disagreement with guidelines, and lack of coordination across clinical roles [40, 99]. Researchers also report that guidelines often do not account for the complexity of the ICU environment, and can be perceived as a “top down” approach that does not consider the expertise or judgment of care providers [34].

Based on these challenges, medical researchers have proposed solutions to increase the uptake of clinical guidelines. These efforts typically fall into two broad approaches. One approach characterizes care providers’ individual beliefs, attitudes, and lack of awareness as the root cause of noncompliance [99]. This strand of research proposes interventions such as formal education and hospital-wide training programs to align provider attitudes and beliefs toward evidence-based practice. While studies show that these interventions can help shift attitudes, researchers note that they are insufficient to change actual clinician behavior [16, 91]. A second, less popular approach characterizes adherence as a systems problem rather than solely a human problem [40, 65]. This line of work suggests that noncompliance is a consequence of the interactions between care providers across time, locations, and system touchpoints along with physical and cultural components of an ICU. Gurses et al. frames current barriers as ambiguity around tasks, methods, and exceptions, and recommends interventions that reduce ambiguity (e.g., use of visual cues to indicate the status of patients with respect to a particular guideline) [40].

Prior HCI research has explored how EHR systems influence protocol adherence [52, 79], and developed systems to support the use of clinical guidelines in the form of digital checklists (e.g., trauma resuscitation [59, 60, 68]), treatment recommendations (e.g., providing clinical guidelines for pneumonia [49], sepsis [95], and cancer [11]), and guideline-specific dashboards (e.g., displaying a list of patients

that qualify for a protocol [4]). While helpful in operationalizing broad best-practice guidelines, these approaches are often limited at the bedside as they do not account for patient-level variation [58]. Recent work highlighted the potential of AI technologies to combine clinical guidelines with patient-level variables to deliver more specific and personalized treatment recommendations [95]. We build on this line of research to understand information needs for WUB to discover opportunities for EHR-based and AI-based interventions that make it easier to consider and perform clinical guidelines.

2.4 AI Systems in Healthcare and ICU

A large body of research has explored data-driven and AI applications in healthcare, often in the form of clinical decision support systems (CDDS) (e.g., [26, 95, 113]). Within the context of ICU, the majority of AI applications has focused on automating documentation related tasks to save time (e.g., transcribing ICU rounds [57]) or providing diagnostic or prognostic insights to help with decision making (e.g., predicting the onset of conditions such as sepsis [41, 71], tachycardia [64] or hypotension [121]). Recently, with the availability of high-density EHR data on large numbers of mechanically ventilated ICU patients (e.g., MIMIC [48]), researchers created AI systems that predict if a patient will need a ventilator [100], predict optimal ventilator settings for a patient [78], and predict the risk of patient extubation failure [105, 124].

While proof-of-concept predictive models showcase initial feasibility, AI systems often fail when moving from research labs to clinical practice [43, 76, 110, 114]. HCI researchers point out that the clinical utility and actionability – the specific actions clinicians can take based on a prediction – of predictive models often remain unclear [36, 102, 106, 119]; and that seamless integration into current workflows is critical for clinician acceptance [74, 103, 114]. In response, an increasing body of HCI literature has called for socio-technical, participatory approaches for understanding clinical workflows and engaging healthcare stakeholders early in the AI system development [3, 27, 36, 47, 87, 114, 117, 118]. In the context of intensive care, a recent interview study explored *what predictions would be useful* for ICU physicians and nurses [32]. Interestingly, clinicians expressed desires for predictions around patient trajectory and prioritization, mainly to reduce the high cognitive load caused by tracking the status of multiple highly dynamic patients rather than aiding decision making. Our research builds on this line of work by investigating current workflows for mechanically ventilated patient care with an eye for opportunities for clinically relevant AI prediction tasks to support the use of WUB in clinical practice.

3 BACKGROUND ON THE WAKE UP AND BREATHE PROTOCOL

ICUs are specialized hospital wards designed to care for patients at extremely high risk for death and disability, particularly patients experiencing acute organ dysfunction necessitating artificial life support. Mechanical ventilation provided via a tube inserted through the trachea (known as “endotracheal intubation”) is among the most common forms of life-support provided in the ICU. Over

700,000 patients in the United States undergo mechanical ventilation in an ICU each year, typically for severe pneumonia or other types of acute respiratory failure, and overall mortality for these patients approaches 30% [46, 111]. Recently, the COVID-19 pandemic caused an enormous influx of patients requiring mechanical ventilation in ICUs worldwide, highlighting the need for approaches to improve outcomes for this high-risk patient population [12].

Mechanical ventilators offer a life-saving intervention by both protecting a patient's airway and by taking over the work of breathing [42]. This treatment helps manage a patient's oxygen level and provides support so their body can rest and heal. Ventilators manage the inflow of oxygen and outflow of carbon dioxide, allowing clinicians to control air volume, pressure, flow, oxygen levels, and respiratory rate. They can take over the entire work of breathing while patients are sedated to tolerate the endotracheal tube running through their throat or they can be set to assist, to react and support a patient's effort to breathe [70]. In the US, the use of a ventilator requires the collective actions of a team of specialists including the respiratory therapist (RT), who makes changes to the ventilator's modes and settings, the attending physician, who oversees the totality of a patient's care, and an ICU nurse, who monitors a patient's changing condition, adjusts sedation in real time [109].

Standards of care for ventilators have considerably changed over the last six decades; best practices suggest reducing time spent on ventilators, reducing sedation, and reducing breathing support [35]. Continued sedation can lead to delirium [70], and remaining on a ventilator increases the chances of pneumonia [42]. Air pressure and volume forced into the lungs can cause damage, and it can negatively impact heart function [35]. Too much oxygen can lead to toxicity in the blood [35]. Clinical teams work to get patients off the ventilator as quickly as possible to reduce the negative effects of their use. However, liberating patients from a ventilator too soon will frequently lead to reintubation, which is associated with greater morbidity and mortality [39, 85, 88].

Medical researchers have created '*the Wake Up and Breathe*' protocol that describes evidenced-based practices for patient liberation [55]. First, an ICU nurse will perform a spontaneous awakening trial (SAT); they cut sedation and then assess if a patient can adjust to being awake. Next, an RT will perform a spontaneous breathing trial (SBT); they remove the ventilator's support and monitor if the patient can successfully breathe on their own. Based on the results of these two assessments, the nurse, RT, and physician will decide if a patient can be extubated (i.e., have the breathing tube removed). To make this decision, the team monitors the patient's physiological status during the SAT and SBT – patients who are able to breathe comfortably during the trials are eligible for extubation, while patients who show signs of respiratory distress during the trial should be returned to the full support of the ventilator. The protocol is performed every 24 hours for mechanically ventilated patients that meet the criteria to evaluate whether a patient is ready to be liberated from the ventilator [109].

The protocol requires careful coordination between the nurse who executes the SAT and the RT who executes the SBT. Cutting sedation, done by the nurse, should occur shortly but not immediately before the RT starts the breathing trial. Cutting the sedation too early may lead patients to become agitated and experience extreme

discomfort from the breathing tube while waiting for the SBT to begin. However, too much sedation inhibits the drive to breathe and may cause the patient to fail an SBT. Typically, the assessment is initiated by the nurse in coordination with the RT. It does not require the direct involvement of physicians. The protocol is ideally carried out before the patient rounds, so that during the rounds, the RT and the nurse report for each patient, allowing the care team to make the final determination about patient extubation. Once the protocol is delayed after rounds, clinicians tend to miss 'the window of opportunity' for WUB, which often postpones patient extubations until the following day, leading patients to remain longer on the ventilator [5].

The use of WUB reduces the average number of days a patient remains on a ventilator and decreases mortality [55]. Unfortunately, the success of the protocol in controlled studies has been hard to reproduce in clinical practice. Often fewer than 50% of eligible patients receive the SAT and SBT [18, 56]. Researchers have used surveys and interviews of ICU clinicians to probe barriers to protocol adherence [31, 40, 56]. They note that nurses seem more reluctant to cut off sedation for patients receiving high doses [56] as well as for patients who suffer from conditions like diarrhea, fatigue, confusion, or agitation [31]. SAT and SBT are not performed when the workload is perceived as high or complex, or when the team is unfamiliar with the protocol [31]. In addition, clinicians sometimes disagree with patient eligibility criteria [31] or believe that SAT can cause the patient short-term harm [69] and that a resting patient is healing [31]. Finally, barriers also include a lack of coordination [18, 31] and a lack of supportive leadership [31].

Prior research explored how to overcome barriers by comparing the differences between ICUs that have better adherence to ICUs that have poorer adherence. They found the protocol is used more when teams can predict each other's actions [20, 30], and when the physician is involved in performing the daily breathing trials [30]. Additionally, discussion of sedation during daily patient rounds indicated a higher likelihood of the protocol being executed [69]. A recent study showed that presenting a dedicated dashboard displaying eligible patients to nurses and RTs may shorten the duration of mechanical ventilation, although the effect sizes were modest and rates of SAT/SBT use was unchanged, calling into question the robustness of the findings [4]. Building on this line of research, this paper aims to understand the clinical practice and information needs around patient liberation in the ICU in order to inform more effective use of machine intelligence to enhance collaborative practice.

4 METHOD

We wanted to gain a deeper understanding of the clinical workflow around the liberation of mechanically ventilated patients. Specifically, we sought to understand how nurses and respiratory therapists decide which patients should receive SAT and SBT, and situations where they deviate from the protocol. We wanted to gain a nuanced understanding of previously identified WUB barriers and identify opportunities where technology might help.

Table 1: 1:1 interview participants.

ID	Exp.	Site	ID	Exp.	Site
Nurse 1	6 yrs	Site 3	RT1	2 yrs	Site 2
Nurse 2	6 yrs	Site 3	RT2	12 yrs	Site 2
Nurse 3	4 yrs	Site 1	RT3	28 yrs	Site 1
Nurse 4	14 yrs	Site 1	RT4	7 yrs	Site 2
Nurse 5	1.5 yrs	Site 1	RT5	8 yrs	Site 2
Nurse 6	2 yrs	Site 1	RT6	5 yrs	Site 1
Nurse 7	10 yrs	Site 2	RT7	7 yrs	Site 1
Nurse 8	1 yrs	Site 1			
Nurse 9	1 yrs	Site 1			
Nurse 10	2.5 yrs	Site 2			

Table 2: Group interview participants.

Session	ID	Exp.	Site
Session 1	G1-Nurse 1	3 yrs	Site 1
Session 1	G1-RT1	28 yrs	Site 1
Session 1	G1-RT2	5 yrs	Site 3
Session 2	G2-Nurse 2	5 yrs	Site 1
Session 2	G2-Nurse 3	3 yrs	Site 2
Session 2	G2-RT3	22 yrs	Site 2
Session 3	G3-Nurse 4	1 yr	Site 1
Session 3	G3-RT4	20 yrs	Site 2
Session 4	G4-Nurse 5	1.5 yrs	Site 1
Session 4	G4-RT5	5 yrs	Site 3
Session 4	G4-RT6	19 yrs	Site 1

4.1 Study Design

Building on prior ethnographic approaches in HCI research for healthcare [50, 84, 114], we chose to conduct a qualitative field study to capture clinical workflows and context with an eye for design opportunities to improve clinical practice. We conducted observations and semi-structured interviews across three sites in the east coast of the United States. All sites were medical ICUs (MICU), which are general facilities that treat adult patients with any number of serious medical conditions (as opposed to specialized ICUs such as surgical ICUs or pediatric ICUs). To strengthen the generalizability of our work, we studied both community and academic ICUs. The study was approved by our Institutional Review Board. Below we provide an overview of the sites:

- **Site 1:** A community hospital with a 16-bed MICU. It has attending physicians who supervise residents, but it does not operate as a teaching hospital. The MICU includes step-down beds where patients receive a higher level of treatment than a typical inpatient ward but less intensive interventions and monitoring than a typical ICU.
- **Site 2:** An academic hospital attached to a cancer center. It has an 18-bed MICU and other specialized ICUs. The patients are seen by physicians and fellows who supervise residents and medical students. It primarily admits oncology patients, but it does treat other patients when needed.
- **Site 3:** An academic hospital with a 14-bed MICU. This site specializes in women's care, but it also provides care for all types of critically ill patients.

Observations. We observed clinical workflows over 17 sessions (47 hours in total). We observed care in patient rooms by reviewing the EHR and purposively selecting mechanically ventilated patients who are likely to be eligible for an SAT and SBT. We also followed the clinical team during rounds, and shadowed individual clinicians (i.e., attending physician, RT, nurse) in order to observe collaborative interactions from a variety of different perspectives. During observations, we recorded field notes, paying attention to start time and duration of clinical actions. We did not collect any personal identifiable information and did no recordings to protect patient privacy.

1:1 Interviews. We conducted one-on-one interviews with 10 nurses and 7 RTs. We used a semi-structured interview guide designed to elicit information about their role, practices and workflow

around mechanically ventilated patients. We asked them to walk us through a recent patient where they performed an SAT and/or SBT in order to ground their responses in specific patient experiences. The open ended section of the guide probed for details around collaboration and coordination with clinical care team members; timing of care; and other related tasks they performed. We asked them about cases where the protocol might not be followed, current barriers and their needs for support. Interviews lasted between 45-60 minutes, and were audio recorded and transcribed verbatim. Participants were compensated \$100 for their time. Table 1 provides an overview of the participants from the one-on-one interviews.

Group Interviews. Following the one-on-one interviews, we conducted group interviews to validate the insights we gained with nurses and RTs together. In each session, we had at least one nurse and one RT participant, with a total of 11 participants (i.e., 6 RTs, 5 nurses) across four group sessions. Except for three participants, none of the participants had been involved in our prior interview study. Each session lasted 45-60 minutes. Participants were compensated \$125 for their time. Table 2 presents an overview of the participants from the group interviews.

We recruited an initial set of participants through our collaborators at each hospital who shared our research study with their clinician colleagues. We then expanded this set through snowball sampling [94], asking participants to share any contacts with the relevant clinical experience.

4.2 Data Analysis

We analyzed our notes from observations using affinity diagrams [66]. We created a service blueprint [13] of the clinical workflow for mechanically ventilated patient care. We analyzed the interview transcripts through thematic analysis, following the analysis process outlined in [23]. First, all researchers familiarized themselves with the data. We then began generating initial codes and divided the transcripts between researchers. Each interview transcript was read and coded by at least two researchers. We frequently met to review, discuss and refine themes, and iteratively update the codes as we analyzed each interview. We generated 11 domain categories (e.g., SAT, SBT, timing of care, coordination) and 72 preliminary themes through several iterations and discussions, from which we constructed three main themes.

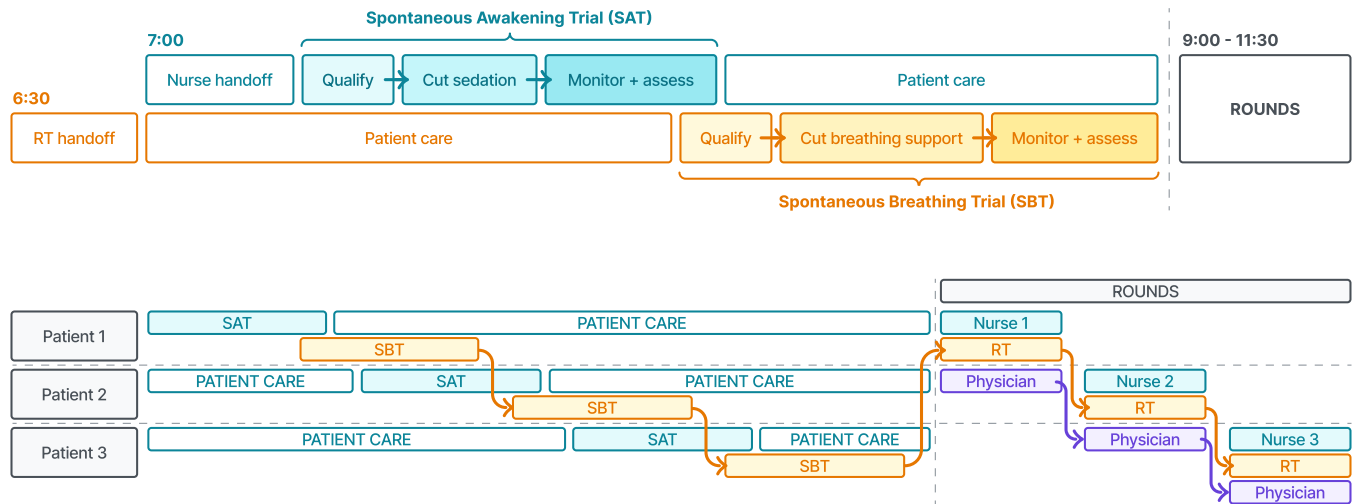


Figure 1: Ideal workflow for Spontaneous Awakening Trial (SAT) and Spontaneous Breathing Trial (SBT) for a single patient (top) and unit-level patients (bottom). Nurses (teal), RTs (orange), physicians and all other clinicians performing patient rounds (purple) are shown as swim lanes. Notice that the RT ideally arrives at each patient’s room shortly after the nurse caring for that patient has started the SAT.

5 FINDINGS

First, we provide a high-level overview of the clinical context for the patient liberation protocol in the ICU. Second, we provide evidence demonstrating the essential role of workflow uncertainty as a barrier to effective care delivery, capturing the various points of uncertainty that nurses and RTs face across the protocol steps, from assessing a patient’s eligibility to receive a protocol, to performing the protocol and making care decisions (e.g., whether to liberate the patient or not). Finally, we detail how care providers interact with the EHR with a focus on their unmet information needs.

5.1 Overview of the ICU Context

During our observations, we did not see a single instance where WUB was executed as the protocol is described in the literature. SAT often did not happen or happened after an SBT started. Echoing prior literature, several participants shared that less experienced nurses might not be aware of the protocol or might feel intimidated to cut sedation: “They’re afraid, it’s very intimidating. [The patient is] biting at the tube ... you’re trying to make sure their restraints are tied.” (Nurse 2) However, the resistance to SAT seemed much less than the literature indicates. The lack of adherence did not stem from a lack of implementation: SAT and SBT were often performed, albeit with deviations.

Instead, what we observed was a tension between the desire to perform the protocol and the obligation to do a whole bunch of other tasks – paired with the notion that the other tasks are more urgent or easier as they do not require collaboration: “You want to do it, you just sometimes can’t in reality ... It’s very busy, you’re pulled in a lot of different directions.” (Nurse 7) Below we present an overview of what a day looked like for a nurse and an RT with particular attention to the timeline for SAT and SBT.

5.1.1 Team Composition and Tasks. All ICU sites contained physicians, nurses, and RTs who were meant to collaborate when executing WUB. These different team members had varying schedules, assignments, and responsibilities. The ICU nurses worked three 12-hour shifts during a 7-day period. Each nurse has two patients, but they may have more if the ICU is short staffed. Nurses received patient assignments at the beginning of their shift and are not always assigned to the same patients day-to-day. The RT group was usually composed of 7-10 RTs – these care providers support the entire hospital and can be assigned to work on the ICU unit or elsewhere. When they are assigned to the ICU, typically one or two RTs are responsible for all patients in the unit (i.e., somewhere between 16 to 20 patients). The attending physicians worked 12-hour shifts over 7 consecutive days (though often staying longer). As a result, all care teams for individual patients, defined as the physician-nurse-RT triad responsible for a specific mechanically ventilated patient, were largely ad hoc in nature and would vary from patient-to-patient, day-to-day, and week-to-week.

The RT, nurse, and physician start the day shift respectively at 6.30am, 7am, and 8am (Figure 1). Shift change takes place with verbal handoffs within each discipline for about 30 minutes. Nurses hold a daily huddle, a 10-minute brief meeting before handoffs, where the charge nurse announces patient assignments. RTs and physicians typically do individual 10-minute rounds to quickly go through each patient before the handoff. At Site-1 and Site-2, physicians and nurses had a 30-minute meeting around 8.30am to go over patients “just to get to speed and be on the same page”. Clinicians shared that typically 30% of the patients in a unit are on mechanical ventilators, whereas 50% would be considered high acuity (e.g., 7-8 patients in a 16-bed unit). While we did not conduct observations during the COVID-19 pandemic waves, clinicians noted that during the pandemic ‘the whole unit was on ventilators’.

SAT or SBT was only one of the many tasks nurses and RTs perform on a daily basis. ICU nurses had a large number of tasks they had to complete before rounds, including head-to-toe patient assessment, delivering medications, tracking and documenting vitals, placing new orders, cleaning or mobilizing patients, and occasionally talking to family members. RTs provided care for patients who are on invasive ventilation (i.e., mechanical ventilation) or non-invasive ventilation (e.g., CPAP or BIPAP, nasal high flow cannulas). They spent at least 10 minutes going through each patient in the ICU, performing a variety of breathing related tasks before rounds started. All clinicians worked under time pressure with the need to react to unscheduled, high priority events. On several occasions we observed individual members of a care team being pulled away from a task to address a crisis, such as reviving a patient going into cardiac arrest. As an RT put it, the care team was constantly *'putting out fires'*.

5.1.2 Timeframe for SAT and SBT. All ICUs conducted daily multidisciplinary rounds, where the clinical team assembled outside of a patient's room to discuss the case and collaborate over a plan of action for individual patients. Rounds usually started around 9am and took about two hours or longer, which meant that SAT and SBT should ideally be completed before 9am. SAT took about half an hour, as patients needed time to wake up from the sedatives. SBT took about an hour to observe if the patient could successfully breathe on their own. Overall, WUB took about 1-1.5 hours to perform, which meant that the care team should have a plan on which patients to SAT and SBT, and initiate the protocol as soon as they start shift. This was rarely the case; as aforementioned, morning time before the rounds was incredibly busy with lots of tasks, leaving little time for effective coordination between the time they arrive and the time rounds start.

Our interviews revealed that the protocol was often delayed or deprioritized, especially when care providers felt uncertainty about whether a patient should receive the protocol, and uncertainty around each other's schedule. Below, we present nuanced details on how these uncertainties impede the protocol adherence for nurses and RTs.

5.2 Points of Uncertainty Surrounding the Execution of Wake Up and Breathe

Our service blueprint (mapping the use of protocol in clinical practice) revealed three phases for WUB: (1) assessing a patient's eligibility to receive SAT/SBT, (2) performing SAT/SBT, (3) informing care decisions for patient liberation. In each phase, care providers faced uncertainties that made it challenging to adhere to the protocol. These uncertainties cascaded across care providers and led to unintentional noncompliance, given the ease in which providers could default to care practices for which there was less uncertainty. Below, we present the points of uncertainty for each phase and detail how these caused deviations.

5.2.1 Uncertainty before performing SAT and SBT. Struggle to know which patients are eligible for SAT and SBT. Our interviews revealed that both nurses and RTs face challenges in assessing a patient's eligibility to receive an SAT or SBT. Several nurses indicated that the protocol provided only a high level guidance and

multiple opportunities to consider a patient non-eligible for vague and often subjective reasons (e.g., "hemodynamic instability"): *"We need a more concrete pathway as to when we're going to SAT and SBT somebody, and what are signs and symptoms to look for."* (Nurse 7) SAT eligibility criteria required nurses to evaluate both patient conditions (i.e. hypertension, open chest, upcoming surgery, etc.) and ventilator settings (i.e. concentration of oxygen, pressure levels). Interpreting ventilator settings created an additional barrier, especially for inexperienced nurses: *"the nurse may not know how to interpret vent settings into weanable or extubatable."* (RT4)

RTs were dependent on information from nurses in order to assess a patient's SBT eligibility. They do not have the full picture of a patient because their lens is narrowly focused on the patient's respiratory status rather than the whole patient: *"[The patient] could have chronically high or low blood pressure that I'm unaware of. [I might think] the heart rate is really high, maybe I shouldn't SBT them. [Nurse] can say, oh, that's normal for them."* (RT1) RTs also looked for information about a patient's daily care plan, and whether the patient is scheduled for a procedure that made them ineligible (e.g., dialysis, surgery). Some RTs shared that after they get their report, they try to listen in during the nurses' handoff to learn more about the patients. They expressed that it can be challenging to get this information from the EHR or by looking at the patient: *"I might not be able to look at the IV (intravenous bag) pole and be able to tell what [sedatives] they're on. That's where the communication with my nurse comes in."* (RT1)

Uncertainty combined with high workload under time pressure seemed to deprioritize WUB. Care providers were constantly *"pulled in a lot of different directions"* (Nurse 7). This drove them to prioritize simple actions they know they must take (e.g., administering medications or charting in the medical record), and delay complex actions with ambiguity and uncertainty: *"We usually wait until we complete rounds to ask about whether [physicians] want us to do a sedation interruption (SAT) or an SBT. Between nursing and respiratory, if there is a question, we feel safer addressing it in rounds and then doing those two tasks."* (Nurse 1) Both nurses and RTs wanted to know the physician's priority and goals for patient liberation to address uncertainty: *"If you're really lucky, you'll get the physician that comes through and catches you before you start your round. You say, Hey listen, this is who I've got, what do you want us to do today? Like, what is your goal, so that it's my goal, so that we can get more people extubated?"* (G1-RT2) Importantly, given the busy ICU environment, these otherwise well-intentioned "delays" often meant the patient would never receive an SAT and SBT at all, not because of any intentional decision but instead because of uncertainty around eligibility affecting overall care.

5.2.2 Uncertainty when performing SAT and SBT. What does it mean to cut sedation? Our interviews surfaced ambiguities in terms of whether sedation should be cut completely or reduced. Some nurses shared that it is common practice to have "a touch of sedation" during SBT for patient comfort; they did not consider this a deviation from the protocol: *"Just because they're on some sedation doesn't necessarily mean we're suppressing their whole respiratory drive. We can just make them a little bit more comfortable so that they're able to be more compliant with the ventilator."* (Nurse 6) Others raised concerns, noting that patients were unlikely to pass SBT

while on sedation: *“I ask the nurse giving me report, why did [the patient] fail their SBT? They did the SBT on fentanyl. It’s not like a true SBT in my eyes.”* (Nurse 7) RTs often commented that they would try an SBT for qualifying patients, regardless of sedation: *“You look at drips, look at the patient, look at the vials, look at the ventilator. And then you start a trial if you think they’re able to do it, regardless of whether there is some sedation or not.”* (RT6)

What counts as sedation? Related to cutting sedation, our conversations also revealed uncertainty around what counts as sedation. All nurses we interviewed shared that propofol and fentanyl were the most commonly used sedatives; however there were participants who did not consider these as sedatives for SAT: *“I was always told sedation isn’t a pain medication and fentanyl isn’t sedation.”* (Nurse 8) A few participants mentioned the use of additional sedatives that did not suppress the respiratory system, yet this did not seem to be common knowledge: *“There’s really only one that people can do a spontaneous breathing trial on ... Precedex, because that does not have any respiratory depression. It’s a bit infrequent, but we do occasionally use it.”* (Nurse 6) In this context, some providers might feel that they are successfully completing an SAT when they cut only some sedative medications, leading to well-intentioned non-adherence.

Which patients should have priority? Among our participants, there was no agreed upon approach for patient prioritization for SAT and SBT. During our fieldwork, we observed some RTs creating a prioritized patient list on a piece of paper. When asked whether and how they prioritize patients, every RT we interviewed shared their own way of prioritizing. Some started with the patients who were most likely to extubate: *“I will generally go ahead to see who is on the lightest settings, then I begin to do a spontaneous breathing trial on them. One of my first go-to’s was in the morning, who can I think about waking up and breathing to possibly get extubated by noon.”* (RT4) Others simply followed the geographic location order, going patient to patient in order of where patients are located in the ICU: *“Unless I’m told something specific about a patient, I’m starting at bed one and I’m working around room to room.”* (RT5) This individual approach taken by different RTs makes it impossible for the nurses to infer when the RT might arrive at their patient and to use this inference to choose when to start an SAT meant to coordinate with the RT.

5.2.3 Uncertainty after performing SAT and SBT. How long is ‘good enough’ before deciding if a patient should be liberated?

Another point of uncertainty was around when to end the protocol to inform patient liberation decisions. While the protocol suggests that one hour or less is sufficient to evaluate whether patients can breathe on their own, there did not seem to be a standard duration in clinical practice. RTs pointed out that the duration of SBTs rather depended on physicians: *“Some physicians will be like, they pass an SBT. It’s been an hour. They’re doing well. Let’s go ahead and extubate. Love that. Other physicians, it’s been an hour. Okay. Let’s give it another hour and another one. ... it’s just kind of undefined and it feels like you’re kicking the can down the road.”* (RT4) Lack of agreement on protocol duration, combined with the delayed timing of care meant that nurses and RTs had to chase physicians to get them to make a decision: *“You need to go bother the doctors: Hey,*

this guy’s been on a wean for so long. Can we decide one way or the other?” (Nurse 4)

5.2.4 Uncertainty in collaboration and coordination. When should SAT and SBT be performed?

Our interviews revealed that there was no standard timing of care for performing the protocol. While all participants shared that the protocol is supposed to be performed in the morning before rounds, they noted that it might be delayed until the afternoon if the unit is busy. Some nurses performed SAT first thing in the morning, whereas others preferred getting patient care tasks out of the way before cutting sedation: *“Usually I’ll get my other patient squared away and then I’ll go in [the mechanically ventilated patient’s] room. And the first thing I do is stop their sedation, then I go on to perform my assessment, give them their meds so that I’m in the room when they wake up.”* (Nurse 4) However, mornings seemed to be the busiest time for nurses, and depending on the patients: *“I don’t feel like 9 am is a great time when you’re trying to pass meds.”* (Nurse 7) These statements suggest that the timing is driven more on convenience rather than intentional decision making.

How to coordinate with the nurse/RT? All participants emphasized that collaboration between the nurse and RT was essential for the protocol. The coordination for SAT/SBT was often informal: *“Just the coordination, like, ‘Hey, I’m gonna cut my sedation’. [The RT will] come back when they see me in the room.”* (Nurse 2) This seemed to work well for participants with more experience who have been working together for a long while: *“Having a crew that works together on the regular, things happen quicker, patients recover faster. But when they’re all brand new travel nurses and brand new people, everything is hard. We don’t know how to communicate with one another.”* (RT5) All nurses we interviewed seemed confident working with RTs, except for one nurse: *“Treatment wise, some patients get breathing treatments and they do like ventilator care and stuff. I’m not too sure what [RT’s] schedule is.”* (Nurse 5)

Interestingly, several RTs reported that they do not wait for an SAT to perform SBT due to time constraints and challenges in coordination: *“I need to get through those 16 beds, I want to get all my treatments done in a timely fashion. So that’s the quickest way for me to get it done. I don’t want to negate doing an SBT on a patient just because I can’t work with the nurse at that point in time.”* (G1-RT2) They shared that they will start an SBT during their first pass, and they will circle back for a second or third pass for SAT: *“If their sedation is higher, and they’re not able to SBT, then I can let the nurse know. That gives them time to finish their work, pass their meds and start the sedation vacation. So that I can start a second attempt.”* (G1-RT2)

Some RTs expressed concerns with this practice: *“We never skip the wean but [due to heavy workload] we might not do the start over multiple times or it might happen late in the day.”* (RT1) Some nurses felt frustrated about the lack of coordination: *“Sometimes [RTs] won’t even tell the nurse, they’ll just put ‘em on an SBT and I’m like, well, they’re still on sedation, you know?”* (Nurse 7) Moreover, nurses did not want to cut sedation without knowing when the RT would come for an SBT for their patient: *“Sometimes [nurses] don’t want to cut patients’ sedation because they just can’t be there to watch the patient. They have to take their other patient to MRI and they’re gonna be gone for two hours.”* (Nurse 4) Uncertainty around

each other's schedule seemed to be a major barrier to accurately performing SAT and SBT.

5.3 Interactions with EHR

In this section, we detail how nurses and RTs interacted with the EHR and other IT systems with respect to WUB. We capture their pain points as well as their needs for support.

Two IT systems are not better than one. Details on WUB live on a different system than the EHR. Clinicians access both on the same computer, but it is not possible to see both at the same time. None of the computers we observed had two screens, which would make it possible to see guidelines and information about a specific patient at the same time.

The EHR system prompted nurses for SAT (i.e. *"Have you done an SAT for this patient?"*). However, participants shared that the prompt is buried in an obscure location in the EHR rather than being integrated into current workflows: *"It's under the iView, under like patient care, there's a section saying 'started' sedation vacation ... But I don't know if everyone knows about it."* (Nurse 3) Some nurses shared that they rarely used the iView interface: *"I feel like sometimes it's not there and I've never really dug deep enough to see why it's not. But yeah, I have seen it fire a task for sedation interruption."* (Nurse 4) A nurse leader recalled a case where the hospital management received a complaint for noncompliance: *"I think the manager got a complaint or whatever, saying nobody was documenting we did the sedation vacation. I don't think many people are doing it."* (Nurse 1)

EHR does not assess which patients are eligible. When asked how they assess patient eligibility for the protocol, all participants shared that they use the information passed during handoff along with their current knowledge of the patient, rather than using the EHR to identify eligible patients. While the eligibility criteria was often captured (e.g., whether the patient is paralyzed, whether the patient has an open chest, etc.), the system did not provide any support for assessing patient eligibility. Several participants pointed out that the EHR could do a better job in indicating contraindications to help assess which patients are eligible: *"[If there was] a daily SBT SAT thing that fires, but then it fires if it's appropriate. It's smart enough in the background to say there's a contraindication present, seek out physician or something like that, that would be really cool."* (G2-RT3) This information would serve to reduce uncertainty by minimizing the cognitive burden necessary to assess the patient, so long as the data were transparent and interpretable.

The system only prompts for a task, but does not provide guidance on protocol. While the system prompted for SAT and SBT in the form of an alert, inputting data into EHR did not seem to follow the structure of the pathway. Instead, clinicians relied on their memory to go through contraindications and carry out the steps. Nearly all participants brought up the need for an EHR that can support and reinforce the WUB clinical pathway (a flow chart that documents the specific implementation steps for SAT and SBT): *"[When you click on the order set for the respiratory therapist] it brings up an archaic 2009 image of the pathway that hasn't been updated. [The system] is prompting a task for us, but having the pathway visible and easily accessible as [RTs are] getting ready to perform SBT would be extremely helpful."* (G2-RT3) Some RTs noted that training new personnel has been a challenge, as there has been

a high turnover rate since the COVID-19 pandemic. Several nurses (Nurse 1, Nurse 2, Nurse 4, Nurse 7) also indicated that the SAT pathway was not clear, and suggested having step-by-step guidance in the EHR: *"[An integrated module in the EHR] would help so that nurses could know the series of steps that should happen. If the person is agitated, then you go get the physician to assess the patient rather than just restarting their sedation."* (Nurse 4)

SAT/SBT are separate "checkbox" tasks, EHR reinforces individual accountability rather than coordination. Participants' reflections revealed that they prioritized their individual accountability in performing and documenting SAT/SBT over coordination. As a nurse put it: *"When you have people that do their SAT, it's kind of like I'm just coming in and doing my assessment. They're not really concerned with coupling it with the SBT portion of it."* (G3-Nurse 4) Specifically the way these two tasks are designed in the EHR seemed to unintentionally separate nurses and RTs instead of helping them coordinate: *"We're just chasing metrics, you know what I mean? Like, did I do it? And then once it's checked, okay, I'm done. Because everybody is busy. It's not that they're trying to get around it, it's just, that's their work."* (G3-RT4)

Clinicians are penalized for cases with uncertainty, the system does not give them credit for postponed SAT/SBT. The system interface for documenting SAT and SBT prompted clinicians to either record their execution of the protocol or document why the protocol was not performed (e.g., contraindication present, patient not eligible). The task was in the form of an alert that would prompt clinicians in the morning around 9am, and would require documentation within two hours. Missing the task window would be automatically captured as noncompliance. Both nurses and RTs shared that for patients with uncertainty, the protocol was often postponed for later in the day with the documentation option *"MD order, contraindication"* – which led the system to disregard the succeeding clinician efforts: *"It doesn't matter if I go back anytime throughout the day. Since I already documented that contraindication, I don't get credit for any more of an SBT that we would do that day [for that patient]."* (G2-RT3)

Moreover, the EHR entry led to inaccurate data capture: *"It's more of a data [limitation]. It would show that the RT didn't do an SBT on that patient that day. I would just make the assumption as her leader that that patient must not have been eligible, even though they were."* (G2-RT3) These statements suggest that the official records indicating that WUB was not adhered to may not be accurate. A nurse further elaborated that the system would not accept 'continuous sedation interruption' (a patient who receives no sedation). The EHR would force them to document SAT as a discrete action: *"You only get credit if you start and stop [sedation interruption]. You're not technically stopping [to increase sedation back] but the trial ended, patients are advanced through the trial."* (G2-Nurse 3) The system behavior made the unit look noncompliant, even though the protocol was performed accurately: *"We are realizing we weren't getting credit for not hitting 'stopped'. So our [compliance] score was very low. So we're trying to re-educate our staff."* (G2-Nurse 3) Credit for compliance often meant serving the needs of an inflexible EHR more than providing care patients needed.

No unit-level view of mechanically ventilated patients to monitor protocol status and extubation readiness. RTs shared that they have access to a patient-level SBT dashboard showing the

patient's SBT history, yet this was not designed as a shared view; none of the nurses had seen it before (except a nurse leader). Our group interviews revealed the need for a unit-level shared view: *"I think a dashboard would be helpful, like, everybody's everywhere, all at once."* (G3-Nurse 4) Some participants suggested that having a shared view of SAT/SBT status of all patients would save time when getting a hold of physicians, as patient liberation decisions were mostly physician-driven: *"I wish we had a clinical pathway where it was: we put them on the wean, now we get the gas, now the physician comes to see them now, and then we extubate. I feel like a lot of our time is spent going and telling people: Hey this is happening, what do you wanna do?"* (Nurse 4)

6 DISCUSSION

Despite providing best practices for improved patient outcomes, the implementation and uptake of clinical protocols and guidelines in clinical practice remains a persistent challenge [34, 99]. Care providers face challenges in operationalizing the protocols from highly controlled clinical trials in messy, complex, and resource constrained day-to-day practice. Interestingly, the healthcare literature approaches this challenge largely as a problem with human behavior, often overlooking the opportunity for technology to change how people work. Recently, a few researchers suggested taking a more systemic approach to investigate not only clinicians' interactions with each other, but also their interactions with computing systems [40].

We approached this challenge using an HCI lens to investigate the lack of protocol adherence in the ICU, specifically in the context of WUB. Our field study revealed many barriers in how nurses and RTs interact with the EHR that negatively impacted protocol adherence. Current systems do not automate the assessment of patient eligibility, scaffold clinicians in making that assessment, nor support planning how a team of clinicians might efficiently assess all relevant patients within an ICU. We identified several opportunities for technology to better support the execution of WUB. At a higher level, these insights led us to reflect on what 'protocol adherence' means, and on what the role and behavior of technology might be.

Below, we first discuss how our findings point to new design opportunities for technology to help clinicians effectively execute WUB. We then reframe 'protocol non-adherence' and discuss opportunities for HCI research to explore better forms and behaviors of ICU technology. Finally, we reflect on open research questions for future investigation.

6.1 Implications for Designing Protocol-Based Care

A core goal of our study was to explore how healthcare computing systems might better support clinicians in executing WUB within the messiness of real world practice. Our study revealed that current EHR systems provide little support. The protocol seems to have been encoded in the EHR in an inflexible way that did not capture clinicians' mental models. It forces them to readjust their actions and go out of their way to receive electronic credit when they have complied. No information on collaboration gets captured nor do

clinicians seem to be rewarded for working together across a set of patients.

Reflecting on these insights, we draw three implications to inform and inspire new healthcare computing systems in addressing adoption barriers to WUB in the ICU: automating/predicting a patient's eligibility for the protocol, guiding clinician actions through the steps of the protocol, and coordinating protocol actions across clinicians and the many patients they must support. We suspect most if not all of these design considerations will generalize to other clinical guidelines and protocols employed in the ICU, and that they will likely be valuable to clinical practice beyond the ICU.

6.1.1 Support assessing patient eligibility for protocol. Recent literature suggests that indicating which patients are eligible to receive SAT and SBT increases the adherence to the protocol [4]. Our findings echo this. Both nurses and RTs expressed challenges in assessing a patient's eligibility. They desired better support from technology. Current SAT and SBT tasks in EHR are not intelligent; the system does not take into account when a patient has contraindications, even though these are often captured in the patient record. A better system might pre-qualify patients to indicate if they are eligible to receive a protocol. It might make delaying SAT a more explicit task for nurses and RTs, so that they seek physician input to address their uncertainty. Notably, our study raises issues around data trustworthiness; patients with contraindications might have been simply documented as ineligible due to delayed decision making. Future work should investigate better data collection and data fusion (e.g., patient records, data from mechanical ventilators, etc.) to assess the feasibility of pre-qualifying patients for WUB.

We also see opportunities for AI-based systems to support clinicians in assessing patient eligibility for clinical guidelines. Recent research demonstrates the use of large language models (LLMs) to extract patient information from team discussions in ICU rounds in order to suggest evidence-based practices and guidelines [57]. Future work can explore summarizing patient trajectory with regards to specific guidelines (e.g., what was discussed about patient extubation in the past week during rounds?). Additionally, as foundation models open up new avenues in extracting patient variables from EHR [1, 73], information from multiple sources (rounds, EHR, unstructured notes from nursing and RT documentation) can be combined to present patient-specific, contextual clinical guidelines. Indicating patient eligibility for a protocol is likely to increase the uptake of guidelines while reducing clinicians' uncertainty and anxiety around whether they will receive credit for their work.

6.1.2 Support clinicians in adapting clinical guidelines. Prior work highlighted that clinical guidelines should account for the complexity of the ICU environment for increased adoption in clinical practice [34]. Our work raises additional design considerations around the granularity and adaptability of clinical guidelines. Our findings point out that guidelines can remain too high-level, leading to ambiguities in successful execution. Building on the large body of literature on medical checklists [59, 68], we see opportunities creating team-based, shared WUB plans that allow clinicians to set goals as a team – a critical aspect for improved care coordination [2]. Future research should explore how clinical guidelines can be designed in a way that allows clinicians to customize, adapt, and add more granularity onto existing guidelines to operationalize

them in their specific clinical context (e.g., offering clarity on what sedatives suppress the respiratory system, highlighting the ones a specific patient is receiving, and nudging nurses to cut sedation completely rather than reducing it). Central to this point is that computer systems could motivate clinicians to effectively perform the protocol, rather than penalizing them.

6.1.3 Support unit-level coordination of protocol actions between clinicians. HCI and CSCW literature has highlighted the importance of *common information spaces* and *mutual awareness* to help clinicians coordinate care in complex healthcare settings. Our study shows that role-based views can lead to information silos when they are not presented as team-based, shared plans. A better system could support unit-level coordination of SAT and SBT actions between nurses and RTs, rather than separating these as individual responsibilities. We see an opportunity for HCI research to explore interaction designs that present the work for multiple care providers across multiple patients. For instance, a dedicated SAT/SBT dashboard could present a shared view of all ventilated patients to nurses, RTs, and physicians to indicate patient status (e.g., SAT/SBT: ready, in process, completed, not eligible) as well as patient trajectory (e.g., days on the ventilator, prior SAT/SBT performance, prior unsuccessful extubations, etc). Anderson et al.'s study presents a great example in this direction [4].

We also see opportunities for data and AI-based approaches for prioritizing patients in busy wards – a major pain point where clinicians think AI can help [32]. Prior research has mostly focused on prediction tasks around patient extubation for clinical decision support (e.g., whether a patient should be extubated based on the risk of extubation failure [105, 124]). This is a high-risk decision that brings up likely challenges around clinician acceptance [114]. Instead, researchers can focus on predictions around WUB to support care coordination and prioritization. For instance, predictions of the likelihood of a patient to receive SAT/SBT (e.g., highly likely versus less likely) can help nurses and RTs coordinate WUB efforts for high priority patients (e.g., highlighting top three patients that should receive the protocol). Additionally, identifying patients with high uncertainty (e.g., patients with dynamic conditions who might become eligible later in the day) can prompt nurses and RTs to resolve uncertainties for patients who might otherwise not receive the protocol. Exploring whether and how this information can be best incorporated into clinical workflows marks a clear direction for future research.

6.2 Reframing Non-Adherence

A large body of healthcare research has investigated the barriers to the adoption of clinical guidelines in practice [15, 31, 34, 99]. The majority of this literature approaches this challenge using the lens of *compliance*: clinicians are 'non-compliant' when they do not adhere to guidelines. This view has an inherent assumption that clinicians are either negligent as they are not aware of the existence of a protocol or they fail to adequately perform them, or they are non-compliant as they ignore or actively choose not to follow guidelines.

Our findings suggest that in the case for WUB, this characterization of clinician negligence does not capture the real story. Current

literature suggests clinicians' lack of adherence stems from their reluctance or resistance to the protocol. However, what we observed and heard during the interviews indicates something closer to *procrastination*: clinicians are overwhelmed with way more tasks than they can complete. When they experience uncertainty around a patient's eligibility for a particularly unpleasant task – such as cutting a patient's sedation or breathing support as the patient anxiously struggles to remove their tube – and uncertainty around the care team's priorities and schedule, they default to prioritizing tasks with high certainty. They do not explicitly choose not to perform a required task. Instead, they choose to prioritize tasks where they have certainty, and this unintentionally causes the uncertain tasks to remain undone. Similarly, clinicians' reflections on how they prioritize patients for WUB indicate that patients with high certainty for extubation seem most likely to receive the protocol. They prioritize and adhere to the cases where they think the protocol is most important. Their behavior feels a lot like triage, the allocation of resources where they will have the greatest impact, a normal part of clinical practice.

These insights provide the opportunity to reframe what 'protocol adherence' means in the clinical setting. A new way to see this is as supporting how several clinicians from different roles work across different patients to execute an interrelated set of actions. There's an opportunity for ICU level coordination. Current EHR systems seem to unintentionally hinder this work by separating individual clinician's actions. Most healthcare IT systems privilege a patient-level treatment view for a single clinician performing care at a single time and place [114]. A better healthcare system could be designed with a focus on the collective delivery of care, where the patient-level EHR is a component of a larger system that has a better knowledge of activities across all patients and clinicians [38]. For example, many shipping companies provide a customer-centric view, showing people where their package is, as well as a logistics view for executing an efficient delivery of packages across many packages, customers, and delivery drivers. We suspect this alternative, service-oriented view [89, 115] will inspire new possibilities for computational systems that can support a coordinated set of activities, more efficiently and fairly allocating clinical efforts and skills across many patients to maximize outcomes for all. Recent work (e.g., [19, 44, 72, 122]) exploring how to design datasets with domain stakeholders provide great starting points for this strand of research. We see an opportunity for HCI research to envision new forms and behaviors for technical systems and datasets that help clinicians effectively and collectively perform their day-to-day work.

6.3 Open Research Questions

Below we detail three research questions our study raised that merit further investigation.

6.3.1 How to design an AI system that coordinates interactions between care team members for increased use of protocols? While our study reveals opportunities for designing AI systems that support coordination between nurses and RTs, there are open questions from an interaction design standpoint. For instance, a ranked list of mechanically ventilated patients (e.g., based on eligibility and predicted WUB performance) can help RTs follow a specific patient

order, so that nurses might better be able to predict RTs' actions. However, questions remain regarding whether RTs and nurses will accept and use this type of AI system. We anticipate that the human-AI interaction design will be critical for communicating the criteria for patient prioritization in order to establish trust without hindering agency [54, 114] (e.g., high priority patient due to successful SAT/SBT in the past few days). On the other hand, prior HCI work on algorithmic management systems note that such systems can easily become tyrannical tools (e.g., worker surveillance issues raised by the use of an algorithmic room assignment system in the hospitality industry [96]). Future research should follow participatory approaches to co-design future technology with clinicians in ways that support them in coordinating while at the same time protecting their autonomy.

6.3.2 How to situate protocol coordination systems in clinical workflows? In addition to interaction design questions around presenting patient eligibility and priority, there are open research questions around situating this type of systems in critical care workflows. When, where, and in what form should WUB information be presented to nurses, RTs, and physicians? For example, an SAT/SBT dashboard with draft patient priority can be reviewed by the night shift physician fellow for them to mark high priority patients before the day shift handover. Moreover, such a dashboard can be placed in patient rooms or can be situated in the ICU corridor as a mobile, interactive display for use during rounds. Future work should explore specific interaction designs and how clinicians might interact with these in-situ.

6.3.3 Does using AI to recommend or draft coordination plans offer an opportunity for lower-risk AI in healthcare? Would deploying these types of systems make clinicians more open to collaborating with AI in clinical practice? Recent research exploring AI innovation in healthcare proposed patient triage, workload management, documentation, and resource planning as lower-risk application areas that might be better starting places for integrating AI in healthcare compared to high-risk, critical applications such as clinical decision support [22, 116]. This provides a counterintuitive perspective, given that AI systems are increasingly more capable with the rise of multimodal foundation models [62, 73]. Our findings bring up a similar point. From a responsible AI perspective [104, 123], supporting clinicians in the execution of clinical guidelines seems to be less risky, both in terms of clinician acceptance and regulatory challenges. We encourage HCI and AI researchers to further investigate how computational systems support the translation of clinical guidelines into practice.

7 LIMITATIONS

Our study has three limitations. First, our fieldwork focused on nurses and RTs as the key clinician roles involved in the execution of WUB. However, there are many other roles involved in the broader context of patient liberation from mechanical ventilation, including physicians, patients, and patient's family members whose perspectives are not covered in this work. Future work should recruit a broader set of participants to understand their involvement. Second, our qualitative approach mostly focused on capturing the

current state with little insight into clinicians' desired futures. Future research should take participatory, generative approaches to explore how clinicians envision computational systems to better support the incorporation of protocols in day-to-day practice. Finally, our study had a North American focus with its involvement of RT as a distinct clinical role. We acknowledge that ICU settings and practices may differ based on hospitals and geographical locations. Future research should investigate other settings to provide insights into how these findings may differ or generalize to other ICUs.

8 CONCLUSION

In this paper, we have presented a field study to understand why clinicians deviate from the clinical practice guidelines for liberating patients from mechanical ventilation in the ICU. We expanded prior literature by capturing the clinical workflows of nurses and RTs for mechanically ventilated patients, and identifying opportunities where technology might support clinicians in executing the patient liberation protocol. These findings suggest an alternative perspective for clinical protocol adherence, proposing that computational systems should work in service of clinicians to make it easier to consider and perform protocols, rather than characterizing clinicians as noncompliant. We strongly encourage HCI researchers to create socio-technical systems that can support the adoption and execution of clinical practice guidelines in complex, real world healthcare settings within and beyond the ICU.

ACKNOWLEDGMENTS

We thank the participants in this work for their time and valuable input. This material is based upon work supported by the National Science Foundation under Grant No. (2007501) and work supported by the National Institutes of Health (R35HL144804). The first author was also supported by the Center for Machine Learning and Health (CMLH) Translational Fellowships in Digital Health. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the National Institutes of Health.

REFERENCES

- [1] Blythe Adamson, Michael Waskom, Auriane Blarre, Jonathan Kelly, Konstantin Krismer, Sheila Nemeth, James Gippet, John Ritten, Katherine Harrison, George Ho, et al. 2023. Approach to machine learning for extraction of real-world data variables from electronic health records. *Frontiers in Pharmacology* 14 (2023).
- [2] Ofra Amir, Barbara J Grosz, Krzysztof Z Gajos, Sonja M Swenson, and Lee M Sanders. 2015. From care plans to care coordination: opportunities for computer support of teamwork in complex healthcare. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 1419–1428.
- [3] Tariq Osman Andersen, Francisco Nunes, Lauren Wilcox, Enrico Coiera, and Yvonne Rogers. 2023. Introduction to the Special Issue on Human-Centred AI in Healthcare: Challenges Appearing in the Wild. , 11 pages.
- [4] Brian J Anderson, David Do, Corey Chivers, Katherine Choi, Yevgeniy Gitelman, Shivan J Mehta, Venkat Panchandam, Steve Gudowski, Margie Pierce, Maurizio Cereda, et al. 2019. Clinical impact of an electronic dashboard and alert system for sedation minimization and ventilator liberation: a before-after study. *Critical care explorations* 1, 10 (2019).
- [5] Michele C Balas, Judith Tate, Alai Tan, Brennon Pinion, and Matthew Exline. 2021. Evaluation of the perceived barriers and facilitators to timely extubation of critically ill adults: an interprofessional survey. *Worldviews on Evidence-Based Nursing* 18, 3 (2021), 201–209.
- [6] Jakob E Bardram. 2000. Temporal coordination—on time and coordination of collaborative activities at a surgical department. *Computer Supported Cooperative*

- Work (CSCW) 9 (2000), 157–187.
- [7] Jakob E Bardram and Thomas Riisgaard Hansen. 2010. Why the plan doesn't hold: a study of situated planning, articulation and coordination work in a surgical ward. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 331–340.
 - [8] Jakob E Bardram, Thomas R Hansen, and Mads Soegaard. 2006. AwareMedia: a shared interactive display supporting social, temporal, and spatial awareness in surgery. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*. 109–118.
 - [9] Jakob E Bardram and Steven Houben. 2018. Collaborative affordances of medical records. *Computer Supported Cooperative Work (CSCW)* 27 (2018), 1–36.
 - [10] David W Bates. 2010. Getting in step: electronic health records and their role in care coordination. , 174–176 pages.
 - [11] Melissa Beauchemin, Meghan T Murray, Lillian Sung, Dawn L Hershman, Chunhua Weng, and Rebecca Schnall. 2019. Clinical decision support for therapeutic decision-making in cancer: A systematic review. *International journal of medical informatics* 130 (2019), 103940.
 - [12] Pavan K Bhatraju, Bijan J Ghassemieh, Michelle Nichols, Richard Kim, Keith R Jerome, Arun K Nalla, Alexander L Greninger, Sudhakar Pipavath, Mark M Wurfel, Laura Evans, et al. 2020. Covid-19 in critically ill patients in the Seattle region—case series. *New England Journal of Medicine* 382, 21 (2020), 2012–2022.
 - [13] Mary Jo Bitner, Amy L Ostrom, and Felicia N Morgan. 2008. Service blueprinting: a practical technique for service innovation. *California management review* 50, 3 (2008), 66–94.
 - [14] Pernille Bjørn and Morten Hertzum. 2011. Artefactual multiplicity: A study of emergency-department whiteboards. *Computer Supported Cooperative Work (CSCW)* 20 (2011), 93–121.
 - [15] Bronagh Blackwood, Karen EA Burns, Chris R Cardwell, and Peter O'Halloran. 2014. Protocolized versus non-protocolized weaning for reducing the duration of mechanical ventilation in critically ill adult patients. *Cochrane database of systematic reviews* 11 (2014).
 - [16] Andra L Blomkalns, Matthew T Roe, Eric D Peterson, E Magnus Ohman, Elizabeth S Fraulo, and W Brian Gibler. 2007. Guideline implementation research: exploring the gap between evidence and practice in the CRUSADE Quality Improvement Initiative. *Academic Emergency Medicine* 14, 11 (2007), 949–954.
 - [17] M Bodi, A Rodriguez, J Solé-Violán, MC Gilavert, J Garnacho, J Blanquer, J Jimenez, MV De La Torre, JM Sirvent, J Almirall, et al. 2005. Antibiotic prescription for community-acquired pneumonia in the intensive care unit: impact of adherence to Infectious Diseases Society of America guidelines on survival. *Clinical Infectious Diseases* 41, 12 (2005), 1709–1716.
 - [18] Leanne M Boehm, Mary S Dietrich, Eduard E Vasilevskis, Nancy Wells, Pratik Pandharipande, E Wesley Ely, and Lorraine C Mion. 2017. Perceptions of workload burden and adherence to ABCDE bundle among intensive care providers. *American Journal of Critical Care* 26, 4 (2017), e38–e47.
 - [19] Sander Bogers, Janne Van Kollenburg, Eva Deckers, Joep Frens, and Caroline Hummels. 2018. A situated exploration of designing for personal health ecosystems through data-enabled design. In *Proceedings of the 2018 Designing Interactive Systems Conference*. 109–120.
 - [20] Emily M Boltey, Theodore J Iwashyna, Robert C Hyzy, Sam R Watson, Corine Ross, and Deena Kelly Costa. 2019. Ability to predict team members' behaviors in ICU teams is associated with routine ABCDE implementation. *Journal of critical care* 51 (2019), 192–197.
 - [21] Claus Bossen. 2002. The parameters of common information spaces: The heterogeneity of cooperative work at a hospital ward. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*. 176–185.
 - [22] Claus Bossen and Kathleen H Pine. 2023. Batman and Robin in Healthcare Knowledge Work: Human-AI Collaboration by Clinical Documentation Integrity Specialists. *ACM Transactions on Computer-Human Interaction* 30, 2 (2023), 1–29.
 - [23] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
 - [24] Suzanne M Burns. 2012. Adherence to sedation withdrawal protocols and guidelines in ventilated patients. *Clinical Nurse Specialist* 26, 1 (2012), 22–28.
 - [25] Naomi E Cahill, Jeanette Suurdt, Hélène Ouellette-Kuntz, and Daren K Heyland. 2010. Understanding adherence to guidelines in the intensive care unit: development of a comprehensive framework. *Journal of parenteral and Enteral Nutrition* 34, 6 (2010), 616–624.
 - [26] Carrie J Cai, Samantha Winter, David Steiner, Lauren Wilcox, and Michael Terry. 2019. "Hello AI": uncovering the onboarding needs of medical practitioners for human-AI collaborative decision-making. *Proceedings of the ACM on Human-computer Interaction* 3, CSCW (2019), 1–24.
 - [27] Carrie J Cai, Samantha Winter, David Steiner, Lauren Wilcox, and Michael Terry. 2021. Onboarding Materials as Cross-functional Boundary Objects for Developing AI Assistants. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–7.
 - [28] Chia-An Chao. 2016. The impact of electronic health records on collaborative work routines: A narrative network analysis. *International journal of medical informatics* 94 (2016), 100–111.
 - [29] Sarah A Collins, Matthew Fred, Lauren Wilcox, and David K Vawdrey. 2012. Workarounds used by nurses to overcome design constraints of electronic health records. In *NI 2012: 11th International Congress on Nursing Informatics*, June 23–27, 2012, Montreal, Canada., Vol. 2012. American Medical Informatics Association.
 - [30] Deena Kelly Costa, Thomas S Valley, Melissa A Miller, Milisa Manojlovich, Sam R Watson, Phyllis McLellan, Corine Pope, Robert C Hyzy, and Theodore J Iwashyna. 2018. ICU team composition and its association with ABCDE implementation in a quality collaborative. *Journal of critical care* 44 (2018), 1–6.
 - [31] Deena Kelly Costa, Matthew R White, Emily Ginier, Milisa Manojlovich, Sushant Govindan, Theodore J Iwashyna, and Anne E Sales. 2017. Identifying barriers to delivering the awakening and breathing coordination, delirium, and early exercise/mobility bundle to minimize adverse outcomes for mechanically ventilated patients: a systematic review. *Chest* 152, 2 (2017), 304–311.
 - [32] Bar Eini-Porat, Ofra Amir, Danny Eytan, and Uri Shalit. 2022. Tell me something interesting: Clinical utility of machine learning prediction models in the ICU. *Journal of Biomedical Informatics* 132 (2022), 104107.
 - [33] Geraldine Fitzpatrick and Gunnar Ellingsen. 2013. A review of 25 years of CSCW research in healthcare: contributions, challenges and future agendas. *Computer Supported Cooperative Work (CSCW)* 22 (2013), 609–665.
 - [34] Phillip J Gary, Yue Dong, and Alexander S Niven. 2021. Addressing the Elephant in the Intensive Care Unit: The Role of Interprofessional Education. , 293–296 pages.
 - [35] Luciano Gattinoni, Michael Quintel, and John J Marini. 2020. "Less is More" in mechanical ventilation. *Intensive care medicine* 46 (2020), 780–782.
 - [36] Pratik Ghosh, Karen L Posner, Stephanie L Hyland, Wil Van Cleve, Melissa Bristow, Dustin R Long, Konstantina Palla, Bala Nair, Christine Fong, Ronald Pauldine, et al. 2023. Framing Machine Learning Opportunities for Hypotension Prediction in Perioperative Care: A Socio-Technical Perspective. *ACM Transactions on Computer-Human Interaction* (2023).
 - [37] Timothy D Girard, John P Kress, Barry D Fuchs, Jason WW Thomason, William D Schweickert, Brenda T Pun, Darren B Taichman, Jan G Dunn, Anne S Pohlman, Paul A Kinniry, et al. 2008. Efficacy and safety of a paired sedation and ventilator weaning protocol for mechanically ventilated patients in intensive care (Awakening and Breathing Controlled trial): a randomised controlled trial. *The Lancet* 371, 9607 (2008), 126–134.
 - [38] John Glaser. 2020. It's time for a new kind of electronic health record. *Harvard Business Review* (2020).
 - [39] Giacomo Grasselli, Alberto Zangrillo, Alberto Zanella, Massimo Antonelli, Luca Cabrinini, Antonio Castelli, Danilo Cereda, Antonio Coluccello, Giuseppe Foti, Roberto Fumagalli, et al. 2020. Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy Region, Italy. *Jama* 323, 16 (2020), 1574–1581.
 - [40] AP Gurses, KL Seidl, V Vaidya, G Boichicchio, AD Harris, J Hebden, and Y Xiao. 2008. Systems ambiguity and guideline compliance: a qualitative study of how intensive care units follow evidence-based guidelines to reduce healthcare-associated infections. *BMJ Quality & Safety* 17, 5 (2008), 351–359.
 - [41] Katharine E Henry, Rachel Kornfield, Anirudh Sridharan, Robert C Linton, Catherine Groh, Tony Wang, Albert Wu, Bilge Mutlu, and Suchi Saria. 2022. Human-machine teaming is key to AI adoption: clinicians' experiences with a deployed machine learning system. *NPJ digital medicine* 5, 1 (2022), 97.
 - [42] Sean M Hickey and Al O Giwa. 2020. Mechanical ventilation. *Treasure Island (FL): StatPearls Publishing* (2020).
 - [43] Michael H Hooper, Lisa Weavind, Arthur P Wheeler, Jason B Martin, Supriya Srinivasa Gowda, Matthew W Semler, Rachel M Hayes, Daniel W Albert, Norment B Deane, Hui Nian, et al. 2012. Randomized trial of automated, electronic monitoring to facilitate early detection of sepsis in the intensive care unit. *Critical care medicine* 40, 7 (2012), 2096.
 - [44] Aspen Hopkins, Fred Hohman, Luca Zappella, Xavier Suau Quadros, and Dominik Moritz. 2023. Designing data: Proactive data collection and iteration for machine learning. *arXiv preprint arXiv:2301.10319* (2023).
 - [45] John A Hughes, Wolfgang Prinz, Tom Rodden, Kjeld Schmidt, Liam Bannon, and Susanne Bødker. 1997. Constructing common information spaces. In *Proceedings of the Fifth European Conference on Computer Supported Cooperative Work*. Springer, 81–96.
 - [46] James C Hurley. 2023. Trends in ICU mortality and underlying risk over three decades among mechanically ventilated patients. A group level analysis of cohorts from infection prevention studies. *Annals of Intensive Care* 13, 1 (2023), 62.
 - [47] Maia Jacobs, Jeffrey He, Melanie F Pradier, Barbara Lam, Andrew C Ahn, Thomas H McCoy, Roy H Perlis, Finale Doshi-Velez, and Krzysztof Z Gajos. 2021. Designing AI for trust and collaboration in time-constrained medical decisions: A sociotechnical lens. In *Proceedings of the 2021 chi conference on human factors in computing systems*. 1–14.
 - [48] Alistair Johnson, Lucas Bulgarelli, Tom Pollard, Steven Horng, Leo Anthony Celi, and Roger Mark. 2020. Mimic-iv. *PhysioNet*. Available online at: <https://physionet.org/content/mimiciv1.0/> (accessed August 23, 2021) (2020).
 - [49] Barbara E Jones, Dave S Collingridge, Caroline G Vines, Herman Post, John Holmen, Todd L Allen, Peter Haug, Charlene R Weir, and Nathan C Dean.

2019. CDS in a learning health care system: identifying physicians' reasons for rejection of best-practice recommendations in pneumonia through computerized clinical decision support. *Applied Clinical Informatics* 10, 01 (2019), 001–009.
- [50] Annika Kaltenhauser, Verena Rheinstädter, Andreas Butz, and Dieter P Wallach. 2020. "You Have to Piece the Puzzle Together" Implications for Designing Decision Support in Intensive Care. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 1509–1522.
- [51] Arvind Karunakaran, Young Hee-Nam, and Madhu Reddy. 2012. Investigating barriers to electronic medical record use during collaborative information seeking activities. In *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*. 743–748.
- [52] Barbara Katzenberg, Fred Pickard, and John McDermott. 1996. Computer support for clinical practice: embedding and evolving protocols of care. In *Proceedings of the 1996 ACM conference on Computer supported cooperative work*. 364–369.
- [53] Brian P Kavanagh and Michael Nurok. 2016. Standardized intensive care. Protocol misalignment and impact misattribution. *American Journal of Respiratory and Critical Care Medicine* 193, 1 (2016), 17–22.
- [54] Saif Khairat, David Marc, William Crosby, Ali Al Sanousi, et al. 2018. Reasons for physicians not adopting clinical decision support systems: critical analysis. *JMIR medical informatics* 6, 2 (2018), e8912.
- [55] Babar A Khan, William F Fadel, Jason L Tricker, W Graham Carlos, Mark O Farber, Siu L Hui, Noll L Campbell, E Wesley Ely, and Malaz A Boustani. 2014. Effectiveness of Implementing a Wake up and Breathe Program on Sedation and Delirium in the Intensive Care Unit. *Critical care medicine* 42, 12 (2014), e791.
- [56] Sucharita Kher, Russel J Roberts, Erik Garpestad, Chris Kunkel, William Howard, Dorothy Didominico, Anne Fergusson, and John W Devlin. 2013. Development, implementation, and evaluation of an institutional daily awakening and spontaneous breathing trial protocol: a quality improvement project. *Journal of intensive care medicine* 28, 3 (2013), 189–197.
- [57] Andrew J King, Derek C Angus, Gregory F Cooper, Danielle L Mowery, Jennifer B Seaman, Kelly M Potter, Leigh A Bukowski, Ali Al-Khafaji, Scott R Gunn, and Jeremy M Kahn. 2023. A voice-based digital assistant for intelligent prompting of evidence-based practices during ICU rounds. *Journal of Biomedical Informatics* 146 (2023), 104483.
- [58] Michael Klompas and Chanu Rhee. 2020. Current sepsis mandates are overly prescriptive, and some aspects may be harmful. *Critical care medicine* 48, 6 (2020), 890–893.
- [59] Leah Kulp, Aleksandra Sarcevic, Megan Cheng, Yanan Zheng, and Randall S Burd. 2019. Comparing the effects of paper and digital checklists on team performance in time-critical work. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [60] Leah Kulp, Aleksandra Sarcevic, Richard Farneth, Omar Ahmed, Dung Mai, Ivan Marsic, and Randall S Burd. 2017. Exploring design opportunities for a context-adaptive medical checklist through technology probe approach. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. 57–68.
- [61] Diana S Kusunoki and Aleksandra Sarcevic. 2015. Designing for temporal awareness: The role of temporality in time-critical medical teamwork. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 1465–1476.
- [62] Peter Lee, Sebastian Bubeck, and Joseph Petro. 2023. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. *New England Journal of Medicine* 388, 13 (2023), 1233–1239.
- [63] Soyoung Lee, Charlotte Tang, Sun Young Park, and Yunan Chen. 2012. Loosely formed patient care teams: communication challenges and technology design. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*. 867–876.
- [64] Xiaoli Liu, Tongbo Liu, Zhengbo Zhang, Po-Chih Kuo, Haoran Xu, Zhicheng Yang, Ke Lan, Peiyao Li, Zhenchao Ouyang, Yeuk Lam Ng, et al. 2021. TOP-Net prediction model using bidirectional long short-term memory and medical-grade wearable multisensor system for tachycardia onset: algorithm development study. *JMIR Medical Informatics* 9, 4 (2021), e18803.
- [65] Aaron R Lyon, Sean A Munson, Brenna N Renn, David C Atkins, Michael D Pullmann, Emily Friedman, and Patricia A Areán. 2019. Use of human-centered design to improve implementation of evidence-based psychotherapies in low-resource communities: protocol for studies applying a framework to assess usability. *JMIR research protocols* 8, 10 (2019), e14990.
- [66] Bella Martin, Bruce Hanington, and Bruce M Hanington. 2012. Universal methods of design: 100 ways to research complex problems. *Develop Innovative Ideas, and Design Effective Solutions* (2012), 12–13.
- [67] Angela Mastrianni, Aleksandra Sarcevic, Lauren Chung, Issa Zakeri, Emily Alberto, Zachary Milestone, Ivan Marsic, and Randall S Burd. 2021. Designing Interactive Alerts to Improve Recognition of Critical Events in Medical Emergencies. In *Designing Interactive Systems Conference 2021*. 864–878.
- [68] Angela Mastrianni, Aleksandra Sarcevic, Hua Cui, Megan A Krentsa, Travis M Sullivan, Issa Zakeri, Ivan Marsic, and Randall S Burd. 2023. Supporting Awareness of Dynamic Data: Approaches to Designing and Capturing Data within Interactive Clinical Checklists. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 1293–1308.
- [69] Melissa A Miller, Sarah L Krein, Christine T George, Sam R Watson, Robert C Hyzy, and Theodore J Iwashyna. 2013. Diverse attitudes to and understandings of Spontaneous Awakening Trials (SAT): results from a statewide quality improvement collaborative. *Critical care medicine* 41, 8 (2013), 1976.
- [70] Alessandro Morandi, Nathan E Brummel, and E Wesley Ely. 2011. Sedation, delirium and mechanical ventilation: the 'ABCDE' approach. *Current opinion in critical care* 17, 1 (2011), 43–49.
- [71] Shamim Nemati, Andre Holder, Fereshteh Razmi, Matthew D Stanley, Gari D Clifford, and Timothy G Buchman. 2018. An interpretable machine learning model for accurate prediction of sepsis in the ICU. *Critical care medicine* 46, 4 (2018), 547.
- [72] Renee Noortman, Peter Lovei, Mathias Funk, Eva Deckers, Stephan Wensveen, and Berry Eggen. 2022. Breaking up data-enabled design: expanding and scaling up for the clinical context. *AI EDAM* 36 (2022).
- [73] H Nori, N King, SM McKinney, D Carignan, and E Horvitz. 2023. Capabilities of GPT-4 on medical challenge problems. arXiv. Preprint posted online March 20 (2023).
- [74] Tariq Osman Andersen, Francisco Nunes, Lauren Wilcox, Elizabeth Kaziunas, Stina Matthiesen, and Farah Magrabi. 2021. Realizing AI in healthcare: challenges appearing in the wild. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–5.
- [75] Ann S O'Malley, Joy M Grossman, Genna R Cohen, Nicole M Kemper, and Hoangmai H Pham. 2010. Are electronic medical records helpful for care coordination? Experiences of physician practices. *Journal of general internal medicine* 25 (2010), 177–185.
- [76] Sun Young Park, Pei-Yi Kuo, Andrea Barbarin, Elizabeth Kaziunas, Astrid Chow, Karandeep Singh, Lauren Wilcox, and Walter S Lasecki. 2019. Identifying challenges and opportunities in human-AI collaboration in healthcare. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*. 506–510.
- [77] Sharoda A Paul and Madhu C Reddy. 2010. Understanding together: sensemaking in collaborative information seeking. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 321–330.
- [78] Arne Peine, Ahmed Hallawa, Johannes Bickenbach, Guido Dartmann, Lejla Begic Fazlic, Anke Schmeink, Gerd Ascheid, Christoph Thiernemann, Andreas Schuppert, Ryan Kindle, et al. 2021. Development and validation of a reinforcement learning algorithm to dynamically optimize mechanical ventilation in critical care. *NPJ digital medicine* 4, 1 (2021), 32.
- [79] Kathleen H Pine and Melissa Mazmanian. 2014. Institutional logics of the EMR and the problem of 'perfect' but inaccurate accounts. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 283–294.
- [80] Juliette Rambourg, Hélène Gaspard-Boulinc, Stéphane Conversy, and Marc Garbey. 2018. Welcome OnBoard: An Interactive Large Surface Designed for Teamwork and Flexibility in Surgical Flow Management. In *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*. 5–17.
- [81] Madhu Reddy and Paul Dourish. 2002. A finger on the pulse: temporal rhythms and information seeking in medical work. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*. 344–353.
- [82] Madhu Reddy, Wanda Pratt, Paul Dourish, and MM Shabot. 2003. Sociotechnical requirements analysis for clinical systems. *Methods of information in medicine* 42, 04 (2003), 437–444.
- [83] Madhu C Reddy, Paul Dourish, and Wanda Pratt. 2001. Coordinating heterogeneous work: Information and representation in medical care. In *ECSCW 2001: Proceedings of the Seventh European Conference on Computer Supported Cooperative Work 16–20 September 2001, Bonn, Germany*. Springer, 239–258.
- [84] Madhu C Reddy, Paul Dourish, and Wanda Pratt. 2006. Temporality in medical work: Time also matters. *Computer Supported cooperative work* 15, 1 (2006), 29–53.
- [85] Safiya Richardson, Jamie S Hirsch, Mangala Narasimhan, James M Crawford, Thomas McGinn, Karina W Davidson, Douglas P Barnaby, Lance B Becker, John D Chelico, Stuart L Cohen, et al. 2020. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. *Jama* 323, 20 (2020), 2052–2059.
- [86] Markus Rittenbruch and Gregor McEwan. 2009. An historical reflection of awareness in collaboration. *Awareness systems: Advances in theory, methodology and design* (2009), 3–48.
- [87] Samantha Robertson and Niloufar Salehi. 2020. What If I Don't Like Any Of The Choices? The Limits of Preference Elicitation for Participatory Algorithm Design. arXiv preprint arXiv:2007.06718 (2020).
- [88] Laurent Robriquet, Hugues Georges, Olivier Leroy, Patrick Devos, Thibaut D'escrivain, and Benoit Guery. 2006. Predictors of extubation failure in patients with chronic obstructive pulmonary disease. *Journal of critical care* 21, 2 (2006), 185–190.
- [89] Virpi Roto, Jung-Joo Lee, Effie Lai-Chong Law, and John Zimmerman. 2021. The overlaps and boundaries between service design and user experience design. In

- Designing Interactive Systems Conference 2021*. 1915–1926.
- [90] Aleksandra Sarcevic, Leysia A Palen, and Randall S Burd. 2011. Coordinating time-critical work with role-tagging. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. 465–474.
 - [91] Ian Scott. 2007. The evolving science of translating research evidence into clinical practice. *BMJ Evidence-Based Medicine* 12, 1 (2007), 4–7.
 - [92] Peter G Scupelli, Yan Xiao, Susan R Fussell, Sara Kiesler, and Mark D Gross. 2010. Supporting coordination in surgical suites: Physical aspects of common information spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1777–1786.
 - [93] Noa Segall, Alberto S Bonifacio, Atilio Barbeito, Rebecca A Schroeder, Sharon R Perfect, Melanie C Wright, James D Emery, B Zane Atkins, Jeffrey M Taekman, Jonathan B Mark, et al. 2016. Operating room-to-ICU patient handovers: a multidisciplinary human-centered design approach. *The Joint Commission Journal on Quality and Patient Safety* 42, 9 (2016), 400–AP5.
 - [94] Julia Simkus. 2023. Snowball sampling method: Definition, Techniques & Examples. <https://www.simplypsychology.org/snowball-sampling.html>
 - [95] Venkatesh Sivaraman, Leigh A Bukowski, Joel Levin, Jeremy M Kahn, and Adam Perer. 2023. Ignore, trust, or negotiate: understanding clinician acceptance of AI-based treatment recommendations in health care. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18.
 - [96] Franchesca Spektor, Sarah E Fox, Ezra Awumey, Christine A Riordan, Hye Jin Rho, Chinmay Kulkarni, Marlen Martinez-Lopez, Betsy Stringam, Ben Begleiter, and Jodi Forlizzi. 2023. Designing for Wellbeing: Worker-Generated Ideas on Adapting Algorithmic Management in the Hospitality Industry. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 623–637.
 - [97] Preethi Srinivas, Anthony Faiola, and Gloria Mark. 2016. Designing guidelines for mobile health technology: Managing notification interruptions in the ICU. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 4502–4508.
 - [98] Jean E Stevenson, Gunilla C Nilsson, Göran I Petersson, and Pauline E Johansson. 2010. Nurses' experience of using electronic patient records in everyday practice in acute/inpatient ward settings: A literature review. *Health Informatics Journal* 16, 1 (2010), 63–72.
 - [99] Lisa Stoneking, Kurt Denninghoff, Lawrence DeLuca, Samuel M Keim, and Benson Munger. 2011. Sepsis bundles and compliance with clinical guidelines. *Journal of intensive care medicine* 26, 3 (2011), 172–182.
 - [100] Harini Suresh, Nathan Hunt, Alistair Johnson, Leo Anthony Celi, Peter Szolovits, and Marzyeh Ghassemi. 2017. Clinical intervention prediction and understanding using deep networks. *arXiv preprint arXiv:1705.08498* (2017).
 - [101] Charlotte Tang and Sheelagh Carpendale. 2007. An observational study on information flow during nurses' shift change. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 219–228.
 - [102] Anja Thieme, Danielle Belgrave, and Gavin Doherty. 2020. Machine learning in mental health: A systematic review of the HCI literature to support the development of effective and implementable ML systems. *ACM Transactions on Computer-Human Interaction (TOCHI)* 27, 5 (2020), 1–53.
 - [103] Anja Thieme, Maryann Hanratty, Maria Lyons, Jorge Palacios, Rita Faia Marques, Cecily Morrison, and Gavin Doherty. 2023. Designing human-centered AI for mental health: Developing clinically relevant applications for online CBT treatment. *ACM Transactions on Computer-Human Interaction* 30, 2 (2023), 1–50.
 - [104] Anja Thieme, Aditya Nori, Marzyeh Ghassemi, Rishi Bommasani, Tariq Osman Andersen, and Ewa Luger. 2023. Foundation Models in Healthcare: Opportunities, Risks & Strategies Forward. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–4.
 - [105] Flavia Torrini, Ségolène Gendreau, Johanna Morel, Guillaume Carteaux, Arnaud W Thille, Massimo Antonelli, and Armand Mekontso Dessap. 2021. Prediction of extubation outcome in critically ill patients: a systematic review and meta-analysis. *Critical Care* 25, 1 (2021), 1–12.
 - [106] Stephanie Tulk Jesso, Aisling Kelliher, Harsh Sanghavi, Thomas Martin, and Sarah Henrickson Parker. 2022. Inclusion of clinicians in the development and evaluation of clinical artificial intelligence tools: a systematic literature review. *Frontiers in Psychology* 13 (2022), 830345.
 - [107] Marion van der Kolk, Mark van den Boogaard, Corine ter Brugge-Speelman, Jeroen Hol, Luc Noyez, Kees van Laarhoven, Hans van der Hoeven, and Peter Pickkers. 2017. Development and implementation of a clinical pathway for cardiac surgery in the intensive care unit: effects on protocol adherence. *Journal of evaluation in clinical practice* 23, 6 (2017), 1289–1298.
 - [108] Lauren Wilcox, Jie Lu, Jennifer Lai, Steven Feiner, and Desmond Jordan. 2010. Physician-driven management of patient progress notes in an intensive care unit. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1879–1888.
 - [109] Heather Willden. [n. d.]. Best Practices: Unplanned Extubations. ([n. d.]).
 - [110] F Perry Wilson, Michael Shashaty, Jeffrey Testani, Iram Aqeel, Yuliya Borovskiy, Susan S Ellenberg, Harold I Feldman, Hilda Fernandez, Yevgeniy Gitelman, Jennie Lin, et al. 2015. Automated, electronic alerts for acute kidney injury: a single-blind, parallel-group, randomised controlled trial. *The Lancet* 385, 9981 (2015), 1966–1974.
 - [111] Hannah Wunsch, Walter T Linde-Zwirble, Derek C Angus, Mary E Hartman, Eric B Milbrandt, and Jeremy M Kahn. 2010. The epidemiology of mechanical ventilation use in the United States. *Critical care medicine* 38, 10 (2010), 1947–1953.
 - [112] Yan Xiao, Stephen Schenkel, Samer Faraj, Colin F Mackenzie, and Jacqueline Moss. 2007. What whiteboards in a trauma center operating suite can teach us about emergency department communication. *Annals of emergency medicine* 50, 4 (2007), 387–395.
 - [113] Qian Yang, Aaron Steinfeld, and John Zimmerman. 2019. Unremarkable AI: Fitting intelligent decision support into critical, clinical decision-making processes. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–11.
 - [114] Qian Yang, John Zimmerman, Aaron Steinfeld, Lisa Carey, and James F Antaki. 2016. Investigating the heart pump implant decision process: opportunities for decision support tools to help. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 4477–4488.
 - [115] Nur Yildirim, Alex Kass, Teresa Tung, Connor Upton, Donnacha Costello, Robert Giusti, Sinem Lacin, Sara Lovic, James M O'Neill, Rudi O'Reilly Meehan, et al. 2022. How Experienced Designers of Enterprise Applications Engage AI as a Design Material. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–13.
 - [116] Nur Yildirim, Changhoon Oh, Deniz Sayar, Kayla Brand, Supriya Challa, Violet Turri, Nina Crosby Walton, Anna Elise Wong, Jodi Forlizzi, James McCann, et al. 2023. Creating Design Resources to Scaffold the Ideation of AI Concepts. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 2326–2346.
 - [117] Nur Yildirim, Mahima Pushkarna, Nitesh Goyal, Martin Wattenberg, and Fernanda Viégas. 2023. Investigating How Practitioners Use Human-AI Guidelines: A Case Study on the People+ AI Guidebook. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–13.
 - [118] Nur Yildirim, Hannah Richardson, Maria T. Wetscherek, Junaid Bajwa, Joseph Jacob, Mark A. Pinnock, Stephen Harris, Daniel Coelho de Castro, Shruthi Bannur, Stephanie L. Hyland, Pratik Ghosh, Mercy Ranjit, Kenza Bouzid, Anton Schwaighofer, Fernando Pérez-García, Harshita Sharma, Ozan Oktay, Matthew Lungren, Javier Alvarez-Valle, Aditya Nori, and Anja Thieme. 2024. Multimodal Healthcare AI: Identifying and Designing Clinically Relevant Vision-Language Applications for Radiology. *arXiv:2402.14252 [cs.HC]*
 - [119] Nur Yildirim, John Zimmerman, and Sarah Preum. 2021. Technical Feasibility, Financial Viability, and Clinician Acceptance: On the Many Challenges to AI in Clinical Practice. In *HUMAN@ AAAI Fall Symposium*.
 - [120] Nur Yildirim, Susanna Zlotnikov, Deniz Sayar, Jeremy M. Kahn, Leigh A. Bukowski, Sher Shah Amin, Kathryn A. Riman, Billie S. Davis, John S. Minturn, Andrew J. King, Dan Ricketts, Lu Tang, Venkatesh Sivaraman, Adam Perer, Sarah M. Preum, James McCann, and John Zimmerman. 2024. Sketching AI Concepts with Capabilities and Examples: AI Innovation in the Intensive Care Unit. *arXiv:2402.13437 [cs.HC]*
 - [121] Joo Heung Yoon, Vincent Jeanselmie, Artur Dubrawski, Marilyn Hravnak, Michael R Pinsky, and Gilles Clermont. 2020. Prediction of hypotension events with physiologic vital sign signatures in the intensive care unit. *Critical Care* 24, 1 (2020), 1–9.
 - [122] Hubert Dariusz Zajac, Natalia Rozalia Avlona, Finn Kensing, Tariq Osman Andersen, and Irina Shklovski. 2023. Ground Truth Or Dare: Factors Affecting The Creation Of Medical Datasets For Training AI. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*. 351–362.
 - [123] Hubert D Zajac, Dana Li, Xiang Dai, Jonathan F Carlsen, Finn Kensing, and Tariq O Andersen. 2023. Clinician-facing AI in the Wild: Taking Stock of the Sociotechnical Challenges and Opportunities for HCI. *ACM Transactions on Computer-Human Interaction* 30, 2 (2023), 1–39.
 - [124] Qin-Yu Zhao, Huan Wang, Jing-Chao Luo, Ming-Hao Luo, Le-Ping Liu, Shen-Ji Yu, Kai Liu, Yi-Jie Zhang, Peng Sun, Guo-Wei Tu, et al. 2021. Development and validation of a machine-learning model for prediction of extubation failure in intensive care units. *Frontiers in medicine* 8 (2021), 676343.