


# BMJ Open User interface approaches implemented with automated patient deterioration surveillance tools: protocol for a scoping review

Yik-Ki Jacob Wan <sup>1</sup>, Guilherme Del Fiol <sup>1</sup>, Mary M McFarland,<sup>2</sup> Melanie C Wright<sup>3</sup>

**To cite:** Wan Y-KJ, Del Fiol G, McFarland MM, *et al*. User interface approaches implemented with automated patient deterioration surveillance tools: protocol for a scoping review. *BMJ Open* 2022;**12**:e055525. doi:10.1136/bmjopen-2021-055525

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2021-055525>).

Received 14 July 2021

Accepted 04 January 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>Department of Biomedical Informatics, University of Utah, Salt Lake City, Utah, USA

<sup>2</sup>Eccles Health Sciences Library, University of Utah, Salt Lake City, Utah, USA

<sup>3</sup>College of Pharmacy, Idaho State University, Pocatello, Idaho, USA

## Correspondence to

Melanie C Wright;  
melaniewright2@isu.edu

## ABSTRACT

**Introduction** Early identification of patients who may suffer from unexpected adverse events (eg, sepsis, sudden cardiac arrest) gives bedside staff valuable lead time to care for these patients appropriately. Consequently, many machine learning algorithms have been developed to predict adverse events. However, little research focuses on how these systems are implemented and how system design impacts clinicians' decisions or patient outcomes. This protocol outlines the steps to review the designs of these tools.

**Methods and analysis** We will use scoping review methods to explore how tools that leverage machine learning algorithms in predicting adverse events are designed to integrate into clinical practice. We will explore the types of user interfaces deployed, what information is displayed, and how clinical workflows are supported. Electronic sources include Medline, Embase, CINAHL Complete, Cochrane Library (including CENTRAL), and IEEE Xplore from 1 January 2009 to present. We will only review primary research articles that report findings from the implementation of patient deterioration surveillance tools for hospital clinicians. The articles must also include a description of the tool's user interface. Since our primary focus is on how the user interacts with automated tools driven by machine learning algorithms, electronic tools that do not extract data from clinical data documentation or recording systems such as an EHR or patient monitor, or otherwise require manual entry, will be excluded. Similarly, tools that do not synthesise information from more than one data variable will also be excluded. This review will be limited to English-language articles. Two reviewers will review the articles and extract the data. Findings from both researchers will be compared with minimise bias. The results will be quantified, synthesised and presented using appropriate formats.

**Ethics and dissemination** Ethics review is not required for this scoping review. Findings will be disseminated through peer-reviewed publications.

## INTRODUCTION

If recognised and treated early, patients who experience deterioration conditions have a lower risk of developing adverse events, such as sepsis and acute and kidney injury.<sup>1</sup> To

## Strengths and limitations of this study

- This protocol is novel in investigating how deterioration information is presented and integrated into clinical workflows.
- We will review studies across broad definitions of patient deterioration.
- Descriptions or evaluations for some commercial deterioration prediction tools may not be available in academic publications.
- Only English-language articles will be included.

ensure these patients receive interventions early, escalation protocols that include evaluation criteria for the patients are commonly established in hospitals.<sup>1 2</sup> However, by some accounts, bedside staff only follow protocols in 8% of all hospital adverse events<sup>2-4</sup>

Several scoring systems, such as Early Warning Score (EWS) and Modified Early Warning Score, have been developed and adopted widely to help clinicians identify patients whose conditions may deteriorate in the hours to come. However, clinical outcomes from the use of EWS have been mixed.<sup>5-7</sup> Bedside warnings reported by these scores are often not acknowledged or acted on because bedside staff encounter high false positives and low actionable values. The perception of warnings that are not actionable could be due to the timing of the warning.<sup>8</sup> To achieve better predictive performance, researchers have turned to artificial intelligence (AI) and machine learning (ML) algorithms to predict adverse events. Little is known about how the user interface (UI) design of these systems impacts clinician workload and clinical outcomes.

Since the 1990s, the development and continuous refinement of scoring systems to predict patient deterioration have garnered many reviews of their effectiveness. Lagadec *et*

found anecdotal evidence that various EWSs are beneficial to clinical staff when implemented.<sup>9</sup> However, clinical outcomes also depend on factors other than the EWSs' predictive performance and incorporated escalation protocols. McNeill *et al* reviewed studies that included early detection tools in the activation of rapid response teams.<sup>1</sup> They concluded that the lack of appropriate integration into clinical workflows and UI design shortcomings might have curtailed these systems' performance.<sup>1</sup> In studying how nurses activate rapid response teams, Wood *et al*'s review found that mistrust, over-reliance, miscalculation and the lack of understanding of the EWSs contribute to the failure of escalation. In some cases, such failures may place patients at risk.<sup>10</sup> With broader adoption of AI and ML algorithms, Muralitharan *et al* found that, generally, ML algorithms have greater accuracy in predicting clinical deterioration when developed and evaluated retrospectively. However, few studies assess the clinical benefits of these algorithms in the real world.<sup>11</sup>

There are scoping reviews covering issues surrounding the development and implementation of ML decision support tools in general. However, those reviews have objectives that are different than this protocol. Schwartz *et al* reviewed the level of clinicians' involvement in developing and implementing any decision support tools used in the hospitals.<sup>12</sup> Their inclusion criteria were broad, and their analysis did not include design features of the decision support tools. Similarly, Lee *et al* focus their review on implementation issues of decision support tools without feature analysis.<sup>13</sup> According to our knowledge, this is the first scoping review that focuses on design features of the UI tools specifically for the early prediction of patient deterioration and adverse events.

## OBJECTIVES

The objective is to identify design principles, human factors methods, human–computer interactions and sociotechnical factors practised in developing surveillance tools that predict patient adverse events during their hospital stay. We will classify different approaches to UI designs and evaluate the impact of different approaches on usability, clinical decision-making, and patient outcomes. We will chart the types of UI designs, the information provided, the effectiveness of these surveillance tools, and metrics used to evaluate their effectiveness.

## METHODS AND ANALYSIS

We will conduct our scoping review under the guidance of the latest version of the *JBIM Manual for Evidence Synthesis* and organise the protocol on the framework of five stages proposed by Arksey and O'Malley: (1) identifying the research question, (2) identifying relevant studies, (3) study selection, (4) extracting the collected data and (5) reporting the results.<sup>14–16</sup> For transparency and reproducibility, we will adhere to the reporting guidelines defined in the Preferred Reporting Items for Systematic Reviews

and Meta-Analyses extension for scoping reviews.<sup>17</sup> Details regarding electronic sources of data, date ranges, and inclusion and exclusion criteria are outlined in the 'Stage 2 Identifying relevant studies' section.

We will use Covidence (Veritas Health Innovation) an online systematic reviewing platform, to screen and select studies. Citation management and duplicate detection and removal will be accomplished with EndNote (Clarivate Analytics.) We will use a spreadsheet programme to extract and chart our data.

A search for existing reviews was conducted in PubMed (pubmed.gov), Epistemonikos (www.epistemonikos.org), PROSPERO (www.crd.york.ac.uk/PROSPERO) and Open Science Framework (osf.io). None were identified as focusing on the UI of the surveillance tool directed to the clinician.

## Stage 1: identifying research questions

We seek to address the following *research question* constructed with JBI's 'PCC' mnemonic: what approaches, at what frequency, have designers and developers used to present patient deterioration risk information to clinicians?<sup>15</sup> *Participants* in the studies include clinicians who use or represent intended users of automated surveillance tools that supply computed deterioration risk information in clinical decision-making. The key *concept* we are exploring is evaluations of automated surveillance tools that support the prediction of patient deterioration by measuring user experience, human–system clinical performance, workflow processes or clinical outcomes. The relevant *contexts* include automated patient surveillance tools in hospital settings in any country.

## Stage 2: identifying relevant studies

The second stage of Arksey and O'Malleys' framework is identifying relevant studies. While many studies evaluate algorithms that provide predictions of patient deterioration, this scoping review focuses on only studies that operationalise these algorithms into usable tools with relevant clinician UIs. Settings should be live or simulated clinical settings that incorporate realistic patient data.

An information specialist (MMM) will develop the search string for our primary database (Medline) and translate it to the other preselected databases by database subject terms and keywords. Library colleagues will peer review the strategy using PRESS guidelines.<sup>18</sup> An example of the search string is included as an online supplemental appendix.

Before the incentives under the EHR Meaningful Use program in 2009, EHR adoption was low.<sup>19</sup> Tools that predict patient deteriorations became technically feasible for design and development only after clinical data were made available electronically. While there may have been decision support tools using automated surveillance before 2009, the potential for the implementation of such tools was limited. Accordingly, we will search for articles from 1 January 2009 to the present.

**Table 1** Search strategy concepts and example search terms used in the search strategy

Concept	Examples
Patient deterioration events	sepsis, acute cardiac, kidney, liver or nervous system failure, respiratory failure, bleeding
Clinicians	Physicians, medical or nursing staff, rapid response teams
Hospital	hospital, intensive care, critical care
User design	Clinical performance or workflow, usability, ergonomics
Machine learning	artificial intelligence, deep learning, algorithms
Surveillance tools	Decision support, electronic health record, track and trigger, clinical alarms, point-of-care tools, rapid response team

Electronic sources will include Medline (Ovid), Embase (embase.com), CINAHL Complete (Ebscohost), Cochrane Library (wiley.com), CENTRAL (wiley.com) and IEEE Xplore (IEEE.org). No methodological nor language filters will be applied.

We will check references of included studies for relevant studies. No grey literature will be selected to search.

### Search terms

The queries will include the following general concepts. [Table 1](#) shows an example of the concepts and example search terms used in the query strings. Medline is our primary database, and our search is highly sensitive to our research question. Search strategies for the other four databases will include more precision and not be as sensitive. The exact preliminary search strategy for Medline is included in online supplemental appendix 1.

### Eligibility criteria

We will include studies that engage clinicians who use or represent intended users of surveillance tools that supply computed deterioration risk information in clinical decision-making as *participants*. As a minimum criterion, studies must include participants recruited from outside of the investigating team.

Studies will be included that address evaluating the UI or user experience of automated surveillance tools that support the prediction, classification or identification of patient deterioration by measuring user experience, human–system clinical performance, workflow processes or clinical outcomes.

Automated surveillance tools are defined as tools that: (1) leverage and aggregate multiple data types that are already being collected within standard care practices, (2) analyse these data dynamically, and (3) provide information to support patient monitoring or clinician decision-making. We limit our review to tools that leverage some form of computational, algorithmic, AI, or ML approach to predict or to classify the risk of patient deterioration in

advance of a relevant, clearly defined clinical outcome. Relevant outcomes may include the following: cardiac arrest, stroke, sepsis, acute kidney injury, acute lung injury, haemorrhage, ventilator-associated pneumonia, thrombosis, seizures, syncope, loss of consciousness, or death. Prediction or risk assessment of surrogate outcomes for clinical deterioration will also be included. Examples include transfer to a higher level of care, activation of rapid response, or code team. Emergent treatments such as mechanical ventilation or rescue medication delivery also are relevant outcomes for inclusion.

For the user experience and subsequent outcomes of automated surveillance tools, we are limiting our review to evaluations that engage clinicians in evaluating any part of the system, including:

1. the UI: the device used for conveying the information such as a phone, pager, or monitor; details of the interface such as display design, message content, risk scoring approach; and integration of information into existing clinical systems such as an EHR or patient monitor.
2. clinical workflow processes: to whom the information is provided and in what clinical situations.

We will include all English-language articles. Non-English studies appearing to meet inclusion criteria via English abstract will be noted as non-English in our data charting form (and no further data abstracted). Funding for translation services has not been allocated.

We will include evaluations in the *context* of automated patient surveillance tools in hospital settings in any country.

Any study that engages users in an evaluation of the relevant tool will be included. For example, original studies including observational, cohort, case control, clinical trial, usability tests, qualitative evaluations will be included.

In sum, the following inclusion criteria will be applied:

- ▶ Original research.
- ▶ Must include descriptions of tools that are used for the surveillance, prediction and detection of patient deterioration events.
- ▶ Algorithms must automatically synthesise multiple types of information.
- ▶ The articles must contain formal evaluation involving human subjects.
- ▶ Intended end-users must be hospital clinicians.

Naturally, any articles that do not meet the inclusion will be excluded. However, to ensure consistency and agreement among evaluators, the exclusion criteria are outlined as follows:

- ▶ Studies that only include analysis of algorithm performance without a clinical use.
- ▶ Studies that describe the UI or architecture designs without an evaluation.
- ▶ Simple monitors that only trigger on preset thresholds for a single parameter.
- ▶ Systems that are only intended for epidemiology studies.

- Calculators that require manual entry.

### Step 3: study selection

Pairs of evaluators will screen the title and abstract from the first 20 randomised entries of the queries' result set for inclusion based on defined criteria. Discrepancies will be resolved through discussions. After resolution, the following 20 studies will be evaluated. This cycle will be repeated until an acceptable kappa agreement of 0.8 is achieved between the reviewers. All titles and abstracts will then be reviewed to identify studies to include for full-text review.

The subset included for full-text review will be evaluated by two reviewers for inclusion. Discrepancies will be resolved through discussions. If discussions fail to resolve differences, a third reviewer will adjudicate.

As in common for scoping review methodology, we do not plan to conduct a quality assessment of included studies. Our goal is to map the literature rapidly to understand the scope of approaches that have been implemented and evaluated.

### Stage 4: data extraction

Electronic spreadsheets will be used in the data extraction process. Three researchers will develop an initial data extraction form and present it to a panel of experts for review and revision. Using the revised form, two researchers will independently perform data extraction on a small sample of articles to evaluate the form's reliability and clarity by calculating interrater agreements. Discrepancies of the extracted data will be resolved by discussion. If new categories are found during the review, they will be added to the extraction form. Redundant categories will be removed, and ambiguous categories will be clarified. The abstraction form will be fine tuned iteratively until good agreement of the extracted data is reached. Core data elements of the data extraction have been submitted as a supplemental in this protocol.

Pairs of researchers will review the included articles and extract data using the extraction process during the extraction process. Differences will be resolved by discussions. A third researcher will adjudicate any unsolved differences.

The following data should be collected:

1. Definition of patient deterioration.
2. The clinical workflow and the targeted patient population.
3. Demographics of the targeted end-users and their professional roles.
4. The users that are included in the evaluation process, along with their demographics and professional roles.
5. The design process/method that was used in developing the tool.
6. Display data: what and how data are displayed in the tool.
7. Contextual data supporting the prediction or risk assessment.

8. Evaluation metrics being used to measure the effects of the tool.
9. The subject focus of the journals.

The extracted data will be classified into categories such as design approach, problem predicted and definitions used to define relevant outcomes. Once classified, the frequency of each of the categories will be counted. We will use descriptive statistics to analyse their frequencies. If available, descriptive statistics will be applied to the sample sizes of the included manuscript. Correlations among related categories will also be analysed.

### Stage 5: data reporting

Along with a narrative description of results, frequency counts of each category identified will be reported in tabular formats. Categories, such as defined patient deterioration outcomes, methods of users' interaction with the systems, and types of information displayed, will be displayed as bar charts or other figure formats for comparison. For example, the types of information displayed in the UI and correlation with definitions of patient deterioration may be displayed as bubble charts.

Change(s) in scoping protocol methodology will be acknowledged and defined in the manuscript.

### Current status

The queries for other databases are under development, and an initial version extraction form has been drafted. We have begun title and abstract screening for articles retrieved with the Medline search. Depending on the size of the result set, the entire project is expected to be completed by April 2022.

### Patient and public involvement

Due to the limited scope of our research support, patient and public involvement has not been included as part of the protocol.

### Ethics and dissemination

Ethics review is not required for this scoping review. Findings will be dissemination through peer-reviewed publications.

**Acknowledgements** We wish to acknowledge the following experts for their input to the search terms and inclusion criteria (in alphabetical order of their surnames): Samir Abdelrahman, Deniz Dishman, Xiaoqian Jiang, Kensaku (Ken) Kawamoto, Kendall Lemmons, Brekk MacPherson, Karl Madaras-Kelly, Jonathan Mark, Mary Nies, Mihai Podgoreanu, Thomas Reese, and Noa Segall.

**Contributors** Y-KJW drafted the manuscript. MCW and GDF reviewed and edited the manuscript. MMM provided feedback and structure of the manuscript. MMM also provided the sample query of the search.

**Funding** This work is supported by the National Institute of General Medical Sciences, National Institute of Health grant number R01GM137083. Research reported in this publication was also supported by the University of Utah Systematic Review Core, with funding in part from the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant UL1TR002538.

**Disclaimer** The content is solely the authors' responsibility and does not necessarily represent the official views of the National Institutes of Health.

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Yik-Ki Jacob Wan <http://orcid.org/0000-0001-9767-2046>

Guilherme Del Fiol <http://orcid.org/0000-0001-9954-6799>

## REFERENCES

- McNeill G, Bryden D. Do either early warning systems or emergency response teams improve hospital patient survival? A systematic review. *Resuscitation* 2013;84:1652–67.
- O'Neill SM, Clyne B, Bell M, *et al*. Why do healthcare professionals fail to escalate as per the early warning system (EWS) protocol? A qualitative evidence synthesis of the barriers and facilitators of escalation. *BMC Emerg Med* 2021;21:15.
- van Galen LS, Struik PW, Driesen BEJM, *et al*. Delayed recognition of deterioration of patients in general wards is mostly caused by human related monitoring failures: a root cause analysis of unplanned ICU admissions. *PLoS One* 2016;11:e0161393.
- Massey D, Chaboyer W, Aitken L. Nurses' perceptions of accessing a medical emergency team: a qualitative study. *Aust Crit Care* 2014;27:133–8.
- Smith MEB, Chiovaro JC, O'Neil M, *et al*. Early warning system scores for clinical deterioration in hospitalized patients: a systematic review. *Ann Am Thorac Soc* 2014;11:1454–65.
- Gerry S, Bonnici T, Birks J, *et al*. Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. *BMJ* 2020;369:m1501.
- Kumar A, Ghabra H, Winterbottom F, *et al*. The modified early warning score as a predictive tool during unplanned surgical intensive care unit admission. *Ochsner J* 2020;20:176–81.
- Shearer B, Marshall S, Buist MD, *et al*. What stops hospital clinical staff from following protocols? An analysis of the incidence and factors behind the failure of bedside clinical staff to activate the rapid response system in a multi-campus Australian metropolitan healthcare service. *BMJ Qual Saf* 2012;21:569–75.
- Le Lagadec MD, Dwyer T. Scoping review: the use of early warning systems for the identification of in-hospital patients at risk of deterioration. *Aust Crit Care* 2017;30:211–8.
- Wood C, Chaboyer W, Carr P. How do nurses use early warning scoring systems to detect and act on patient deterioration to ensure patient safety? A scoping review. *Int J Nurs Stud* 2019;94:166–78.
- Muralitharan S, Nelson W, Di S, *et al*. Machine learning-based early warning systems for clinical deterioration: systematic scoping review. *J Med Internet Res* 2021;23:e25187.
- Schwartz JM, Moy AJ, Rossetti SC, *et al*. Clinician involvement in research on machine learning-based predictive clinical decision support for the hospital setting: a scoping review. *J Am Med Inform Assoc* 2021;28:653–63.
- Lee TC, Shah NU, Haack A, *et al*. Clinical implementation of predictive models embedded within electronic health record systems: a systematic review. *Informatics* 2020;7:25.
- Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol* 2005;8:19–32.
- Peters M, Godfrey C, McInerney P, *et al*. Chapter 11: Scoping reviews. In: Aromataris E, Munn Z, eds. *JBI manual for evidence synthesis*, 2020.
- Peters MDJ, Godfrey CM, Khalil H, *et al*. Guidance for conducting systematic scoping reviews. *Int J Evid Based Healthc* 2015;13:141–6.
- Tricco AC, Lillie E, Zarin W, *et al*. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018;169:467–73.
- McGowan J, Sampson M, Salzwedel DM, *et al*. PRESS peer review of electronic search strategies: 2015 guideline statement. *J Clin Epidemiol* 2016;75:40–6.
- Services USDH. *HITECH act enforcement interim final rule*, 2009.