Journal of the American Medical Informatics Association, 2026, 1–9
https://doi.org/10.1093/jamia/ocag064
Review

Alert fatigue measurement in clinical decision support: a systematic review

Cara E. Ray, PhD, MA*[,1], Geneva M. Wilson, PhD, MPH[1,2], Ashley M. Hughes, PhD, MS[1,3], Cassie Cunningham Goedken, MPH[4], Eric Ping-Fei Liu, MPH[2,5], Margaret A. Fitzpatrick, MD, MS[6,7], Katie J. Suda, Pharmd, MS, FCCP[8,9], Satya Manasa Kota, MBBS, MS[10], Chinonyerem Nwankpa[11], Charlesnika T. Evans, PhD, MPH[1,2]

[1]Center of Innovation for Complex Chronic Healthcare, Edward Hines Jr. VA Medical Center, United States Department of Veterans Affairs, Hines, Illinois, United States
[2]Department of Preventive Medicine, Center for Health Services and Outcomes Research, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States
[3]Department of Medicine at MetroHealth, Case Western Reserve University School of Medicine, Cleveland, Ohio, United States
[4]Center for Access & Delivery Research and Evaluation, Iowa City VA Health Care System, United States Department of Veterans Affairs, Iowa City, Iowa, United States
[5]ID Care, Hillsborough, New Jersey, United States
[6]Seattle-Denver Center of Innovation for Veteran-Centered and Value-Driven Care, VA Eastern Colorado Healthcare System, United States Department of Veterans Affairs, Aurora, Colorado, United States
[7]Division of Infectious Diseases, Department of Medicine, University of Colorado School of Medicine, Aurora, Colorado, United States
[8]Center for Healthcare Evaluation, Research, and Promotion, VA Pittsburgh Healthcare System, United States Department of Veterans Affairs, Pittsburgh, Pennsylvania, United States
[9]Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, United States
[10]Department of Epidemiology and Biostatistics, University of Illinois Chicago School of Public Health, Chicago, Illinois, United States
[11]Department of Biology, North Carolina A&T State University College of Science & Technology, Greensboro, North Carolina, United States
*Corresponding author: Cara E. Ray, PhD, MA, Center of Innovation for Complex Chronic Healthcare (CINCCH), Edward Hines, Jr. VA Hospital (151-H), 5000 S. 5th Avenue, Hines, IL 60141-3030, United States (Cara.Ray@va.gov)

Abstract

Background: **Alert fatigue is defined as alert dismissals due to excessive or irrelevant alerts and is frequently cited as a barrier to clinical decision support system use and impact. However, the criteria for determining the presence or absence of alert fatigue are poorly defined. The objective of this systematic review of systematic reviews was to identify operationalized definitions and measures of alert fatigue or alert-related metrics.**

Methods: **Systematic reviews reporting at least one alert-related metric or measure/operationalization of alert fatigue for physician-directed electronic alerts were included. The Cochrane Library, Embase, and PubMed were searched from database start to 2024. The Revised Assessment of Multiple Systematic Reviews was used to assess study quality and risk of bias. Data were synthesized narratively and with descriptive statistics.**

Results: **A total of 22 studies were included in the review. Studies reported between 1 and 11 alert metrics. Studies were most often of medium quality. Reporting of primary study characteristics was frequently judged to be insufficient. Only one article reported an operational definition of alert fatigue. The most common alert metrics were quantity, override rate, and acceptance rate.**

Discussion: **Alert fatigue measurement methods are not clearly or consistently defined in systematic reviews related to alert fatigue in clinical decision support. Reporting of other primary study characteristics is often limited. We recommend that future efforts use a significant, sustained decrease in appropriate alert response rates from an established baseline as a measure of alert fatigue.**

Key words: alert fatigue, health personnel; decision support systems, clinical; medical informatics; health care evaluation mechanisms; sensitivity and specificity.

Received: November 3, 2025. **Accepted:** April 4, 2026
Published by Oxford University Press on behalf of the American Medical Informatics Association 2026.

Background and significance

Clinical decision support systems (CDSS) are intended to provide information about individual patients' health and clinical knowledge relevant to care decisions.[1,2] They are commonly used to optimize care quality and safety.[3] While they show promise for improving provider behaviors such as compliance with guidelines for prescriptions and patient health outcomes such as risk of adverse drug events,[4–6] CDSSs can generate large alert volumes,[7,8] thereby increasing users' cognitive load.[9] Inappropriately timed or irrelevant alerts can impose distraction and disruption to clinical work.[10,11] CDSSs are noted to produce alert fatigue, which is often conceptualized as when excessive and/or irrelevant alerts lead recipients to ignore, overlook, or override alerts regardless of their relevance for appropriate clinical care given the surrounding context.[3,7,12–17] Inappropriately overridden alerts can lead to adverse health consequences.[16,18] For example, overriding an alert regarding an order for a drug with high potential for a severe drug-drug interaction could cause an adverse drug event. Therefore, to maximize the benefits of CDSSs for patient care, it is important to prevent and combat alert fatigue.

Past studies have explored factors associated with appropriateness of alert overrides per evaluators' criteria[8,15] and described interventions intended to reduce alert volume and override rates.[12,13,19] These alert-related metrics often appear to be treated as proxies for alert fatigue. However, no literature to-date has evaluated whether these metrics actually constitute alert fatigue and whether alert fatigue was successfully prevented. More specifically, prior studies fail to provide operational definitions of alert fatigue; that is, explicit criteria for determining alert fatigue's presence or onset in the study's context.

It is unclear whether or how to intervene on alert fatigue without a validated measure of it. Interventions directed toward proxies of alert fatigue may be ineffective if alert fatigue is not actually present. In one example study, order check alert override rates varied significantly by alert type.[8] Of the alert types with available override data, drug-drug interaction (DDI) alerts had the highest inappropriate override rates, but not the highest overall override rates. Alert fatigue would be more strongly implicated if override rates were similar regardless of alert type. The most relevant type of alert would then have the most inappropriate overrides. Instead, recipients in this study may have selectively underestimated the relevance of DDI alerts compared to other alerts due to misjudgment or poor design of DDI alert features. Alert fatigue countermeasures, such as making DDI alerts more specific, may be less effective for reducing inappropriate overrides in such instances than more clearly emphasizing DDI severity. They could also result in dangerous false negatives. Conversely, if override rates were, in fact, similarly high regardless of alert type, this could suggest alert fatigue and an opportunity to improve the impact of CDSSs on patient care by making alerts more specific. Thus, operational definitions using validated measures of alert fatigue are important for determining whether and how deploying strategies to combat alert fatigue may help prevent patient harm from CDSSs.

Objectives

The initial purpose of this study was to identify common measures or operational definitions of alert fatigue in the CDSS literature and evaluate their implications for care. We focused on provider-targeted systems in any care setting. We operationally defined "measures of alert fatigue" as criteria by which alert fatigue was determined to be present or absent, or something reported to be a measure or operational definition of alert fatigue by the primary review authors.

We sought to conduct a review of systematic reviews. We assumed that systematic reviews examining alert fatigue would have broadly applicable, standardized operational definitions of alert fatigue reported in their inclusion and criteria per publication standards.[20,21] However, we discovered only one review that reported an operational definition or measure of alert fatigue.[7] Therefore, our team expanded the focus of this review of systematic reviews to identify any metrics related to alert quantity, alert appropriateness or performance, alert responses, or appropriateness of responses—hereafter called "alert metrics" for simplicity—regardless of whether they were explicitly intended as measures or operational definitions of alert fatigue. The goal of identifying such metrics was to assess their relation to alert fatigue and how they could inform or be incorporated into a measure of alert fatigue as we conceptualized it: *when excessive and/or irrelevant alerts lead recipients to insufficiently engage with alerts for appropriate patient care given the surrounding context*.

Methods

We report on our systematic review of systematic reviews. Our methods and reporting were guided by the Revised Assessment of Multiple Systematic Reviews (R-AMSTAR) and the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) 2020 checklist.[20,21]

Inclusion criteria

Reviews needed to be systematic to be included. We used the criteria outlined by the Cochrane Collaboration for determining whether a review was systematic.[22,23] In brief, a review needed to explicitly report pre-defined objectives, reproducible methods, and systematic presentation of findings to be counted as a systematic review.[22] No exclusions based on language or peer review status were applied. Reviews needed to discuss provider-facing alerts. Providers needed to be free to determine how to respond to alerts.

Articles were included if they reported any alert metric, including, but not limited to operational definitions or measures of alert fatigue. Variables were considered eligible alert metrics if they pertained to alert quantity, response, or appropriateness. Alert metrics of alert quantity could include the rate of alerts per person, for example, patient or provider, or per time period. Metrics related to alert response could have included the number or rate of overrides or acceptances. Overrides could have included nonresponse, limited response, or dismissal. These would include a failure to engage in the target clinical practice following alerts that can essentially be dismissed by clicking "acknowledge" or for alerts that do not need to be dismissed to continue work.[12] The number of alerts or alert-triggering events followed by action would have needed to be clearly described.

Appropriateness or inappropriateness of overrides were also eligible. Metrics related to alert appropriateness could include system performance metrics such as sensitivity, specificity, or PPV.

Articles were considered eligible if they presented at least one of the aforementioned types of alert metrics. Only metrics of alerts that were actually presented to target recipients, and therefore capable of causing alert fatigue in recipients, were eligible. For instance, reviewers could not confirm whether performance metrics from some articles pertained to systems that sent alerts to end-users. Rather, they may have only come from silent trials. Such metrics were considered ineligible. Alert type was considered a factor that may influence alert metrics rather than an alert metric in and of itself; therefore, articles that only reported alert type and no other alert metric were excluded. An alert metric only had to be reported once in the review to be included, whether in the methods, results, or both. Thus, even metrics that review authors set out to extract a priori were included even if they were not discovered in the primary articles. Metrics that were only mentioned in the results for a single primary article were included, even if it was not clear that the authors extracted information consistently for all included primary articles.

Search methods

The Cochrane Library, Embase, and PubMed were searched. All databases were searched from the database start time to 2024. A set of search terms related to CDSS and alert systems in healthcare was combined with a set of terms related to alert fatigue with using an "AND" operator. The first set contained 3 subsets related to the following topics: (1) decision support tools without mentioning healthcare in the phrase, such as "computerized decision support*"; (2) healthcare, such as "medical informatic*" and (3) CDSSs, such as "clinical decision support system.*" The second set comprised 3 subsets related to 1. alerts, 2. acceptances or overrides, and 3. alert fatigue. In each set, the first 2 subsets were combined with each other using "AND" and then combined with the third subset using "OR." Full search term sets and dates are in Appendix S1.

The lead author (C.E.R.) sought and screened additional articles from Department of Veterans Affairs (VA) research report registries, contacting experts, and manually searching reference lists of articles identified via database search. VA research report registries included VA Health Systems Research (HSR) Reports in Progress, Evidence Synthesis Program Reports with the "systematic review" filter applied, and HSR Publications using the terms "alert fatigue" OR "clinical reminder."

Screening

All identified articles were screened for inclusion/exclusion by C. N. and C.E.R. Both researchers first screened titles and abstracts for relevance. Reports with relevant titles and abstracts underwent full article eligibility screening. Any discrepancies were resolved by discussion between the 2 reviewers or, if consensus could not be reached, with input from A.M.H. and/or G.W.

Data extraction

Two reviewers independently extracted data from each article: C. E.R. and either S.M.K. or E.L. Differences between data extractors were resolved using the same procedures as for screening. They used a standardized data extraction tool (in Appendix S2) to ensure each reviewer completed the same data extraction elements. The following data elements were extracted from each review article as available: alert metrics, measures or operational definitions of alert fatigue used or proposed in the review, time frame, analysis/synthesis methods, alert targets such as drug-drug interaction, and quality per R-AMSTAR. Data extraction items besides those in the R-AMSTAR were developed by C.E.R. to meet the research objective, namely, to identify operational definitions of alert fatigue and alert metrics. Analysis/synthesis methods were extracted in case they yielded additional insights regarding how alerts were quantified or evaluated. Article quotes regarding additional study characteristics, such as potential predictors and outcomes of alert fatigue, were obtained for potential future extraction and synthesis in case new research questions emerged once alert fatigue measurement was better understood.

We combined alert metrics that we deemed as similar or the same into a single metric. For instance, variations of alert quantity such as mean alarm count per day[24] were subsumed under alert quantity. We considered something to be a measure or operational definition of alert fatigue if the authors called it a proxy for, substitute for, operational definition of, indicator of, measure of, or metric of alert fatigue. The terms "alarm," "reminder," or "prompt" instead of "alert" and the terms "overload" or "burden" instead of "fatigue" were acceptable. Measures, metrics, and indicators needed to be reportedly *of* alert fatigue in order to be counted as a measure or operational definition of alert fatigue rather than solely an alert metric. They could not be solely *related to* alert fatigue unless the authors indicated what values or ranges of these metrics were considered unacceptable and to result from cognitive load or desensitization.[7]

Analysis/synthesis

The R-AMSTAR contains an item asking whether reviews reported characteristics of primary studies. We required articles to report the number of patients, the number of physicians, *and* the setting of primary studies to receive credit for this item. We synthesized findings in narrative and tabular format, including frequencies and percentages of each reported alert metric and R-AMSTAR criterion met. The alert metrics, measures or operational definitions of alert fatigue, analysis/synthesis methods, time frame, and R-AMSTAR scores of each study were also tabulated. We divided the range of possible R-AMSTAR scores into thirds to get low (scores 11-21), medium (22-32), and high (33-44) study quality rankings.

Additional methods and results details can be found in Appendix S3.

Results

The results of the search and selection process are depicted in the inclusion flow diagram Figure 1. Articles excluded upon full

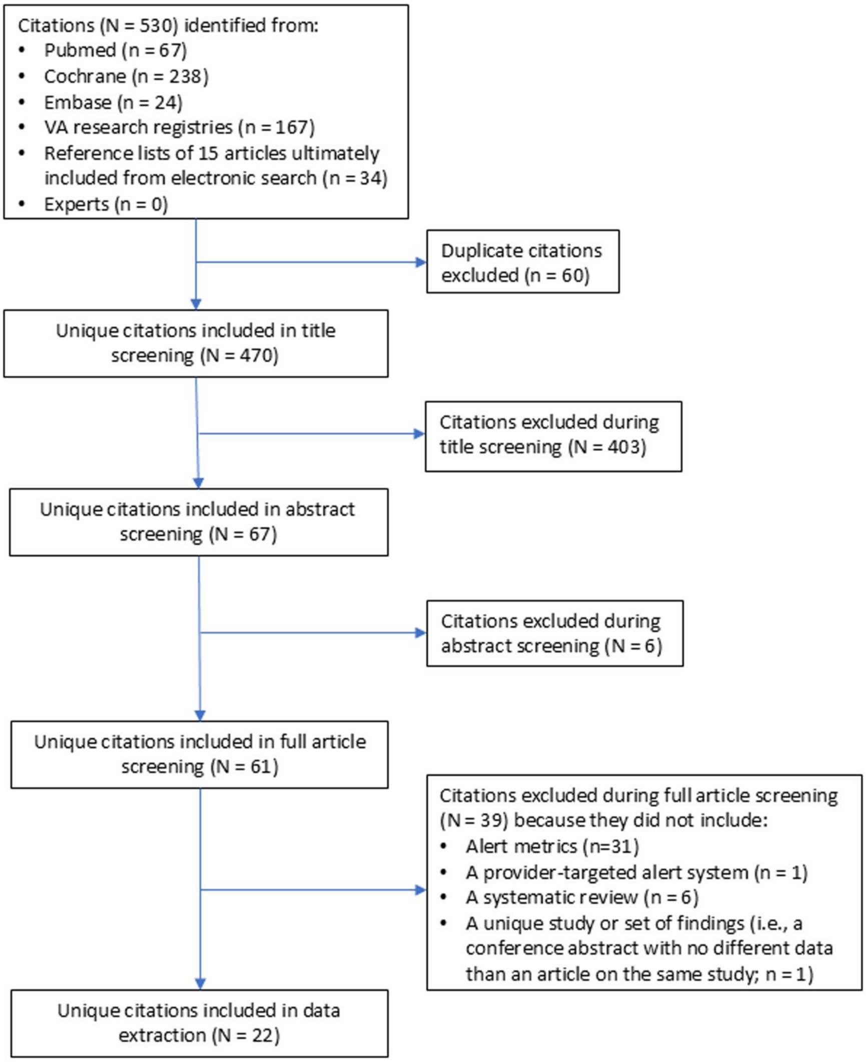

Figure 1. Inclusion flow diagram. Alt text: *Description of articles identified, included, and excluded.*

article screening are cited along with reasons for exclusions in Appendix S4. Twenty-two articles were systematic reviews reporting at least one alert metric and were thus included in our review. Study characteristics, including alert metrics and R-AMSTAR score(s), are in Table 1.

Operational definitions or measures of alert fatigue

Only one included article reported an operational definition or measure of alert fatigue. Specifically, they used alert quantity as a proxy for fatigue.[7] In their discussion, the authors proposed measuring alert fatigue as the rate of inappropriate alert overrides per alert accompanied by high cognitive load or desensitization. They noted that cognitive load could be measured with pre-existing workload assessment tools.

Alert metrics

Alert metrics reported by each study and their definitions are in Table 2. The most common alert metrics were alert quantity, such as mean alarm count per day;[24,37] override rate, and acceptance rates. The largest number of unique alert metrics reported by a single article was 6.[10,26] Of note, while Kane-Gill et al. recommended inappropriate override rates as a component of an operational definition of alert fatigue,[7] only 3 studies (13.6%) reported this metric.[5,18,35] Appropriateness was determined in the primary studies either by clinician evaluators using criteria based on previous research and guidelines,[40,41] or via unreported methods.[5,18,35]

Study quality

Overall, R-AMSTAR study quality scores ranged from 20 (low quality) to 34 (high quality; Carli et al.[26] and Dunn et al.,[27] respectively) with a mean of 26.6 (medium quality; SD = 3.9). Three studies were of low quality (13.6%), one was high quality (4.6%),[27] and the rest were medium quality (81.8%). The percentages of studies meeting each criterion are in Appendix S5. All reported inclusion criteria, a list of included studies *not* just in the reviews' reference list, and electronic search parameters. All but one study reported an a priori design.[38] All but one study reported using 2 electronic sources.[26] No reviews reported

Table 1. Characteristics of included studies.

Author, Year	R-AMSTAR score	Review Time Frame, including databases if start date depended on database inception	# of primary articles	Alert metrics
Bennett, 2003[25]	26	January 1966 - December 31, 2001	26	acceptance rate
Carli, 2018[26]	20	February 2009 - March 2015	17	alert quantity | sensitivity | specificity | PPV | clinical relevance of alerts | alert response time
Dunn, 2021[27]	34	Ovid Medline, Embase, CINAHL, Web of Science, Scopus, and Cochrane Central Register of Controlled Trials from inception—May 13, 2018	11	override rate | alert quantity
Eslami, 2007[28]	28	Ovid Medline & Medline in-process and Embase from inception to—August 2006	67	acceptance rate | override rate
Handler, 2007[29]	32	1985 - 2006	12	PPV
Hemens, 2011[30]	32	2004 - 2010	65	alert quantity | PPV | False alarm or inappropriate alert rate | time to process alerts
Hussain, 2019[12]	21	2007 and 2017	39	acceptance rate | override rate
Hwang, 2020[31]	25	PubMed MEDLINE, Embase, the Cochrane Library, and CINAHL from inception—August 2, 2018	10	sensitivity | specificity | PPV | False alarm or inappropriate alert rate
Kane-Gill, 2017[7]	27	1966-February 1, 2017	26	alert quantity | acceptance rate | override rate | alert value
Lee, 2020[32]	27	1 January 2010-31 July 2019	44	alert quantity | override rate | sensitivity
Légat, 2018[16]	25	CINAHL, Cochrane Library, Embase, Ovid, and PubMed from inception—February 2016.	69	alert quantity | acceptance rate | alert prevalence and acceptance rates
Luri, 2022[18]	27	Same as Légat et al. (2018) plus Medline & Cochrane Library from March 2016-March 2020	28	alert quantity | acceptance rate | override rate | rate of appropriate or inappropriate overrides
Makam, 2015[33]	29	PubMed MEDLINE, Embase, the Cochrane Library, and CINAHL from inception—June 27, 2014	8	sensitivity | specificity | PPV | negative predictive value | likelihood ratio
Marasinghe, 2015[5]	24	Medline, Embase, & Cochrane Library from -1st week of February 2014.	7	alert quantity | acceptance rate | rate of appropriate or inappropriate overrides | PPV
Olakotan, 2021[10]	24	1997 to 2018	76	alert quantity | override rate | sensitivity | specificity | PPV | False alarm or inappropriate alert rate
Page, 2017[34]	21	January 2000-February 2016	23	acceptance rate | override rate
Poly, 2020[35]	30	January 1, 2000-March 31, 2019	23	alert quantity | override rate | rate of appropriate or inappropriate overrides
Shiffman, 1999[36]	22	1992-January 1998	25	alert response time
Van der Vegt, 2024[24]	27	January 1, 2010-April 1, 2023	37	alert quantity | sensitivity | specificity | alert hours before event | risk of event for patients meeting alert criteria versus those who do not| acceptance rate |
Van der Vegt, 2023[37]	30	January 1, 2012-June 23, 2022	30	acceptance rate
Van Dort, 2021[38]	22	January 2010-April 2020	8	alert quantity | acceptance rate | override rate | false alarm or inappropriate alert rate
Warttig, 2018[39]	32	1900-September 18, 2017	3	false negative rate

PPV = positive predictive value; R-AMSTAR = Revised Assessment of Multiple Systematic Reviews.
[*]Database abbreviations and inception dates: Cumulative Index to Nursing and Allied Health Literature (CINAHL; 1982), Cochrane Library (1996), Embase (1980), Ovid and Ovid Medline (1984), and PubMed (1996) Medline (1950), Medline in-process (1966), Web of Science (1964), and Scopus (1966). R-AMSTAR scores are out of 27. The following study quality rankings from possible R-AMSTAR: low (scores 11-21), medium (22-32), and high (33-44).

Table 2. Alert metrics commonly extracted from included articles ($N = 22$)

Common alert metrics	*N* (%)
Alert quantity (how many alerts fired)	12 (54.5)
Override rate (of the total alert quantity, how many were overridden)	10 (45.5)
Alert acceptance (of the total alert quantity, how many were accepted)	10 (45.5)
Alert appropriateness/Positive predictive value (of the total number of alerts, how many were supposed to fire [true positives])	7 (31.8)
Sensitivity (out of the total number of alert-eligible events, how many had an alert)	6 (27.3)
Specificity (out of the total number of non-alert-eligible events, how many did *not* have an alert)	4 (18.2)
Inappropriate alert rate or false alarm rate (false positives out of the total number of alerts or actual negative events, respectively)	4 (18.2)
Rate of appropriate or inappropriate overrides (out of the total number of overrides, how many were appropriate or inappropriate)	3 (13.6)
Alert response time	2 (9.1)
False negative rate (false negatives out of actual positive events [true positives + false negatives])	1 (4.6)
Negative predictive value	1 (4.6)
Risk of event for patients meeting alert criteria versus those who did not	1 (4.6)
Alert hours before event (how many hours before the event did the alert fire)	1 (4.6)
Time to process alerts	1 (4.6)
Likelihood ratio (the true positive rate divided by the false positive rate)	1 (4.6)
Alert "prevalence and acceptance rates,"	1 (4.6)
Clinical relevance of alerts	1 (4.6)
Alert value as measured by number of accepted alerts, decrease in override rates, or judged as clinically significant	1 (4.6)

translating articles in languages other than English, having readers sufficiently trained in said language[s], or graphical/statistical assessments of publication bias. Only one reported tests of homogeneity.[29] Furthermore, only 3 (13.6%) reviews reported the minimum primary study characteristics we determined were necessary to receive credit for the R-AMSTAR item on that topic.[7,25,28]

Discussion

The overarching purpose of our review was to identify measures or operationalizations of alert fatigue. Measuring, identifying, and operationally defining alert fatigue is important for finding ways in which CDSSs may be harming patient care that might otherwise be overlooked in broader evaluations. Only one review included in our results explicitly reported an operational definition or measure of alert fatigue.[7] Specifically, this study operationalized alert fatigue by using alert quantity as a proxy.[7] Twenty-one other studies reported metrics related to alert quantity, appropriateness, or responses—most often, quantity—without explicitly describing them as measures or operational definitions of alert fatigue. Kane-Gill et al. proposed that future studies operationalize alert fatigue as declining appropriate response rates accompanied by cognitive load and/or desensitization as measured by tools external to the alert metrics themselves.[7] However, override appropriateness was infrequently reported. Reporting of key primary study characteristics within each review besides alert metrics was poor (13.6%). Only one included review was categorized as high quality.[27]

By our interpretation, existing alert metrics and operational definitions vary in how they reflect alert fatigue when it is conceptualized as excessive and/or irrelevant alerts leading recipients to insufficiently engage with alerts for appropriate patient care given the surrounding context.[3,7,12–17] Some metrics reflect relevance and quantity. For instance, the positive predictive value, sensitivity, specificity, false negative rate, false positive rate, rate of appropriate or inappropriate alerts, risk of event for patients meeting alert criteria versus those who do not, and alert value relate to the relevance of alerts given the clinical context. Alert hours before event could determine alerts' timeliness,[24] and therefore clinical relevance. Metrics such as sensitivity and specificity also determine alert quantity for a given event prevalence. Alerts could become excessive at some quantities. In contrast, negative predictive value does not directly indicate excessiveness, relevance, or engagement.

Appropriate response rates could measure sufficient engagement when investigators categorize alert responses—acceptances, overrides, and response times—by their contextual appropriateness for patient care, ideally using explicit, evidence-based criteria. Declining appropriate response rates could then reflect decreasing engagement. Cognitive overload, desensitization, and distrust have been proposed to drive the role of excessive or irrelevant alerts in alert fatigue.[7,9,26,31,42–44] Therefore, declining appropriate response rates in the presence of such mechanisms may comprise alert fatigue.[7]

We speculate that alert fatigue could manifest as a substantial and sustained decline in appropriate response rates from a previously established baseline, assuming for now that it takes some time for alerts to accumulate enough to cause cognitive overload, desensitization, or distrust.[7,9,26,31,42–44] Appropriate response rates may also be influenced by factors besides alert fatigue. These factors could include external sources of cognitive load that are high during the baseline period or that vary seasonally. Sustained declines in appropriate response rates following a prior baseline period may be less attributable to these other factors. If these propositions are correct, they align with findings of inconsistent and nonlinear relationships between alert quantity and acceptance rates in prior publications[9,45] and in prior descriptions of alert fatigue as the point of decreasing alert impact[42] or the "Goldilocks" point.[46] However, these propositions are based on our interpretation of the current conceptualization of alert fatigue rather than evidence from our review.

If, after controlling for alternative determinants of engagement, a statistically significant decrease in the relationship between appropriate response rates and time relative to baseline persisted, it could indicate the presence of alert fatigue. The magnitude of this change could measure the extent of fatigue. These propositions are hypothetical and require future empirical testing, including further exploration of other determinants of engagement. Validation could involve examining whether alert quantity and relevance metrics, and psychological mechanisms such as cognitive overload, distrust, and desensitization[7,9,26,31,42,43,47–49] are associated with the onset or severity of such declines. Doing so may ultimately help identify levels of these underlying metrics that are "ideal… to avoid fatigue" (p. 1482).[7]

Limitations and future research

This review has a few limitations. We did not take specific steps to ensure meta-analyses were included in our search results. Meta-analyses may have been more likely to report alert metrics, given their quantitative focus. Future efforts could expand search criteria to identify whether any meta-analyses include alert metrics or alert fatigue operational definitions other than the ones listed here. Also, there was subjectivity in what we considered "alert metrics," which could have artificially narrowed the scope of metrics we identified.

Another limitation is that we did not systematically extract the type of CDSS from each included review. Over- or under-representation of certain CDSS types may have thus biased the alert metrics reported. Finally, we did not perform coverage analysis to examine overlap in reviews' included studies. Thus, alert metrics reported by primary studies included in more than one review may be over-counted. Nevertheless, we do not believe that differences in the exact frequencies of individual alert metrics or additional performance metrics would change our proposed measure of alert fatigue.

Our review also has many strengths. Reviewing systematic reviews rather than individual primary articles offers a high-level overview of common metrics. This approach allowed us to address our research question efficiently by leveraging past systematic reviews evaluating their included studies within the broader context of each other. All included articles underwent dual independent screening and data extraction. We used a previously validated systematic review quality assessment tool. We had minimal eligibility restrictions outside of those needed to address our research questions. We also successfully pivoted our research objective in response to the lack of reported measures of alert fatigue to gain a broader understanding of commonly used alert metrics and propose guidance for future efforts to measure and understand alert fatigue.

Conclusion

In conclusion, operational definitions or measures of alert fatigue are rare. CDSS evaluations would benefit from standard measures of alert fatigue. Such measures can identify instances where CDSSs harm care that might otherwise be overlooked and allow interventions to be targeted appropriately. We propose measuring alert fatigue via a statistically significant, sustained decrease in appropriate response rates over time relative to a previously established baseline after controlling for other determinants of engagement. Future work can validate this measure by examining its relation to potential causes and mechanisms of alert fatigue.

Registration and protocol

The review protocol was not registered. It is available along with a log of amendments by request.

Author Contributions

Cara E. Ray (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing—original draft, Writing—review & editing), Geneva M. Wilson (Conceptualization, Data curation, Methodology, Supervision, Validation, Writing—review & editing), Ashley M. Hughes (Methodology, Validation, Visualization, Writing—review & editing), Cassie Cunningham Goedken (Writing—review & editing), Eric Ping-Fei Liu (Data curation, Writing—review & editing), Margaret A. Fitzpatrick (Writing—review & editing), Katie J. Suda(Conceptualization, Writing—review & editing), Satya Manasa Kota (Data curation, Formal analysis, Investigation, Writing—review & editing) Chinonyerem Nwankpa (Data curation, Writing—review & editing), and Charlesnika T. Evans (Funding acquisition, Supervision, Writing—review & editing)

Disclaimer

The views expressed in this presentation are those of the authors and do not necessarily reflect the decision, position, or policy of the United States Department of Veterans Affairs, or the United States government.

Funding statement

This study received funding support from the Department of Veterans Affairs Health Services Research & Development (HSR&D) Small Award Initiative for Impact program, Center of Innovation for Complex Chronic Healthcare award # SWC 24-004 (Principal Investigator: C.E.R.); HSR&D Senior Research Career Scientist award # MRA0-001-24W (Principal Investigator: C.T.E); and Quality Enhancement Research Initiative program # QUE 20-016 (Principal Investigator: C.T.E.).

Conflict of interests statement

None declared..

Data availability

See Appendix S2 for extraction forms. Data and analytic materials (i.e., Microsoft Excel formulas) are available upon request.

References

1. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ*. 2005;330:765.
2. Osheroff JA, Teich J, Levick D, et al. *Improving Outcomes with Clinical Decision Support: An Implementer's Guide*. Himss Publishing; 2012.
3. Joseph AL, Borycki EM, Kushniruk AW. Alert fatigue and errors caused by technology: A scoping review and introduction to the flow of cognitive processing model. *Knowledge Management & E-Learning: An International Journal*. 2021; 2021:500-521.
4. Shojania KG, Jennings A, Mayhew A, Ramsay CR, Eccles MP, Grimshaw J. The effects of on-screen, point of care computer reminders on processes and outcomes of care. *Cochrane Database Syst Rev*. 2009;2009:CD001096.
5. Marasinghe KM. Computerised clinical decision support systems to improve medication safety in long-term care homes: a systematic review. *BMJ Open*. 2015;5:e006539.
6. Murphy EV. Clinical decision support: effectiveness in improving quality processes and clinical outcomes and factors that may influence success. *Yale J Biol Med*. 2014;87:187-197.
7. Kane-Gill SL, O'Connor MF, Rothschild JM, et al. Technologic Ddstractions (Part 1): summary of approaches to manage alert quantity with intent to reduce alert fatigue and suggestions for alert fatigue metrics. *Crit Care Med*. 2017; 45:1481-1488.
8. Nanji KC, Slight SP, Seger DL, et al. Overrides of medication-related clinical decision support alerts in outpatients. *J Am Med Inform Assoc*. 2014;21:487-491.
9. Ancker JS, Edwards A, Nosal S, et al. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. *BMC Med Inform Decis Mak*. 2017;17:1-9.
10. Olakotan OO, Mohd Yusof M. The appropriateness of clinical decision support systems alerts in supporting clinical workflows: A systematic review. *Health Informatics J*. 2021;27: 14604582211007536.
11. Osheroff JA, editor. *Improving Medication Use and Outcomes with Clinical Decision Support:: A Step by Step Guide 2009*. HIMSS; 2009.
12. Hussain MI, Reynolds TL, Zheng K. Medication safety alert fatigue may be reduced via interaction design and clinical role tailoring: a systematic review. *J Am Med Inform Assoc*. 2019; 26:1141-1149.
13. Chaparro JD, Hussain C, Lee JA, Hehmeyer J, Nguyen M, Hoffman J. Reducing interruptive alert burden using quality improvement methodology. *Appl Clin Inform*. 2020;11:46-58.
14. Wan PK, Satybaldy A, Huang L, Holtskog H, Nowostawski M. Reducing alert fatigue by sharing low-level alerts with patients and enhancing collaborative decision making using blockchain technology: scoping review and proposed framework (MedAlert). *J Med Internet Res*. 2020;22:e22013.
15. van der Sijs H, van Gelder T, Vulto A, Berg M, Aarts J. Understanding handling of drug safety alerts: a simulation study. *Int J Med Inform*. 2010;79:361-369.
16. Légat L, Van Laere S, Nyssen M, Steurbaut S, Dupont AG, Cornu P. Clinical decision support systems for drug allergy checking: systematic review. *J Med Internet Res*. 2018; 20:e258.
17. McCoy AB, Waitman LR, Lewis JB, et al. A framework for evaluating the appropriateness of clinical decision support alerts and responses. *J Am Med Inform Assoc*. 2012;19:346-352.
18. Luri M, Leache L, Gastaminza G, Idoate A, Ortega A. A systematic review of drug allergy alert systems. *Int J Med Inform*. 2022;159:104673.
19. Phansalkar S, Desai A, Choksi A, et al. Criteria for assessing high-priority drug-drug interactions for clinical decision support in electronic health records. *BMC Med Inform Decis Mak*. 2013;13:65.
20. Kung J, Chiappelli F, Cajulis OO, et al. From systematic reviews to clinical recommendations for evidence-based health care: validation of Revised Assessment of Multiple Systematic Reviews (R-AMSTAR) for Grading of Clinical Relevance. *Open Dent J*. 2010;4:84-91.
21. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.
22. Kitsiou S, Pare G, Jaana M. Systematic reviews and meta-analyses of home telemonitoring interventions for patients with chronic diseases: a critical assessment of their methodological quality. *J Med Internet Res*. 2013;15:e150.
23. Green S, Higgins JP, Alderson P, Clarke M, Mulrow CD, Oxman AD. Introduction. Cochrane handbook for systematic reviews of interventions. 2008. p. 1–9.
24. van der Vegt AH, Campbell V, Mitchell I, et al. Systematic review and longitudinal analysis of implementing Artificial Intelligence to predict clinical deterioration in adult hospitals: what is known and what remains uncertain. *J Am Med Inform Assoc*. 2024;31:509-524.
25. Bennett JW, Glasziou PP. Computerised reminders and feedback in medication management: a systematic review of randomised controlled trials. *Med J Aust*. 2003;178:217-222.
26. Carli D, Fahrni G, Bonnabry P, Lovis C. Quality of decision support in computerized provider order entry: Systematic literature review. *JMIR Med Inform*. 2018;6:e3.
27. Dunn AN, Radakovich N, Ancker JS, Donskey CJ, Deshpande A. The impact of clinical decision support alerts on clostridioides difficile testing: a systematic review. *Clin Infect Dis*. 2021;72:987-994.
28. Eslami S, de Keizer NF, Abu-Hanna A. The impact of computerized physician medication order entry in hospitalized patientsa systematic review. *Int J Med Inform*. 2008; 77:365-376.
29. Handler SM, Altman RL, Perera S, et al. A systematic review of the performance characteristics of clinical event monitor signals used to detect adverse drug events in the hospital setting. *J Am Med Inform Assoc*. 2007;14:451-458.
30. Hemens BJ, Holbrook A, Tonkin M, et al. Computerized clinical decision support systems for drug prescribing and

management: a decision-maker-researcher partnership systematic review. *Implement Sci*. 2011;6:89.

31. Hwang MI, Bond WF, Powell ES. Sepsis alerts in emergency departments: a systematic review of accuracy and quality measure impact. *West J Emerg Med*. 2020;21:1201-1210.
32. Lee TC, Shah NU, Haack A, Baxter SL. Clinical implementation of predictive models embedded within electronic health record systems: a systematic review. *Informatics (MDPI)*. 2020;7:25.
33. Makam AN, Nguyen OK, Auerbach AD. Diagnostic accuracy and effectiveness of automated electronic sepsis alert systems: A systematic review. *J Hosp Med*. 2015;10:396-402.
34. Page N, Baysari MT, Westbrook JI. A systematic review of the effectiveness of interruptive medication prescribing alerts in hospital CPOE systems to change prescriber behavior and improve patient safety. *Int J Med Inform*. 2017;105:22-30.
35. Poly TN, Islam MM, Yang HC, Li YJ. Appropriateness of overridden alerts in computerized physician order entry: systematic review. *JMIR Med Inform*. 2020;8:e15653.
36. Shiffman RN, Liaw Y, Brandt CA, Corb GJ. Computer-based guideline implementation systems: a systematic review of functionality and effectiveness. *J Am Med Inform Assoc*. 1999; 6:104-114.
37. van der Vegt AH, Scott IA, Dermawan K, Schnetler RJ, Kalke VR, Lane PJ. Deployment of machine learning algorithms to predict sepsis: systematic review and application of the SALIENT clinical AI implementation framework. *J Am Med Inform Assoc*. 2023;30:1349-1361.
38. Van Dort BA, Zheng WY, Sundar V, Baysari MT. Optimizing clinical decision support alerts in electronic medical records: a systematic review of reported strategies adopted by hospitals. *J Am Med Inform Assoc*. 2021; 28:177-183.
39. Warttig S, Alderson P, Evans DJ, Lewis SR, Kourbeti IS, Smith AF. Automated monitoring compared to standard care for the early detection of sepsis in critically ill patients. *Cochrane Database Syst Rev*. 2018;6:CD012404.
40. Wong A, Amato MG, Seger DL, et al. Prospective evaluation of medication-related clinical decision support over-rides in the intensive care unit. *BMJ Qual Saf*. 2018;27:718-724.
41. Judge J, Field TS, DeFlorio M, et al. Prescribers' responses to alerts during medication ordering in the long term care setting. *J Am Med Inform Assoc*. 2006;13:385-390.
42. Cash JJ. Alert fatigue. *Am J Health Syst Pharm*. 2009; 66:2098-2101.
43. Cvach M. Monitor alarm fatigue: an integrative review. *Biomed Instrum Technol*. 2012;46:268-277.
44. Reddy E. Crying 'Crying Wolf': How Misfires and Mexican engineering expertise are made meaningful. *Ethnos*. 2020; 85:335-350.
45. Murad DA, Tsugawa Y, Elashoff DA, Baldwin KM, Bell DS. Distinct components of alert fatigue in physicians' responses to a noninterruptive clinical decision support alert. *J Am Med Inform Assoc*. 2022;30:64-72.
46. Glassman PA, Good CB, Cunningham FE. Drug alerts and the Goldilocks principle: striving for "just right". *Am J Health Syst Pharm*. 2016;73:517.
47. Fitousi D, Wenger MJ. Processing capacity under perceptual and cognitive load: a closer look at load theory. *J Exp Psychol Hum Percept Perform*. 2011;37:781-798.
48. Lavie N, Hirst A, de Fockert JW, Viding E. Load theory of selective attention and cognitive control. *J Exp Psychol Gen*. 2004;133:339-354.
49. Singh H, Spitzmueller C, Petersen NJ, Sawhney MK, Sittig DF. Information overload and missed test results in electronic health record-based settings. *JAMA Intern Med*. 2013; 173:702-704.

Published by Oxford University Press on behalf of the American Medical Informatics Association 2026.

Journal of the American Medical Informatics Association, 2026, 00, 1–9
https://doi.org/10.1093/jamia/ocag064
Review