Clinical Decision Support and Implications for the Clinician Burnout Crisis

Ivana Jankovic¹, Jonathan H. Chen²

¹ Division of Endocrinology, Stanford University School of Medicine, Stanford, CA, USA
² Center for Biomedical Informatics Research and Division of Hospital Medicine, Stanford University School of Medicine, Stanford, CA, USA

Summary

Objectives: This survey aimed to review aspects of clinical decision support (CDS) that contribute to burnout and identify key themes for improving the acceptability of CDS to clinicians, with the goal of decreasing said burnout.

Methods: We performed a survey of relevant articles from 2018-2019 addressing CDS and aspects of clinician burnout from PubMed and Web of Science¹. Themes were manually extracted from publications that met inclusion criteria.

Results: Eighty-nine articles met inclusion criteria, including 12 review articles. Review articles were either prescriptive, describing how CDS should work, or analytic, describing how current CDS tools are deployed. The non-review articles largely demonstrated poor relevance and acceptability of current tools, and few studies showed benefits in terms of efficiency or patient outcomes from implemented CDS. Encouragingly, multiple studies highlighted steps that succeeded in improving both acceptability and relevance of CDS.

Conclusions: CDS can contribute to clinician frustration and burnout. Using the techniques of improving relevance, soliciting feedback, customization, measurement of outcomes and metrics, and iteration, the effects of CDS on burnout can be ameliorated.

Keywords
Decision Support Systems, Clinical; burnout, professional; Electronic Health Records

Introduction

The Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009 resulted in the rapid incorporation of electronic health records (EHRs) into hospitals and clinics, with the promise of improving healthcare quality and efficiency [1]. Unintentionally, the EHR has created frustrating workflow interruptions, hours of documentation at home, clerical data entry for reimbursement and regulations, cumbersome security requirements, and an interrupted clinician-patient relationship [2]. In one study, 70% of physician EHR users reported stress related to healthcare information technology, almost two-thirds felt that EHRs added to their daily frustrations, and almost half disagreed with the sentiment that EHRs improved patient care [3]. Although superior in many ways to the paper charts they replaced, EHRs’ F-grade usability [1] has become a significant driver of burnout in clinicians [1, 2, 4, 5].

Burnout is a serious issue affecting healthcare professionals. It is characterized as a syndrome of emotional exhaustion, depersonalization, and a low sense of personal accomplishment. The causes of burnout are multifactorial, but include workload pressures, inefficiencies, and moral distress from ethically undesirable situations [2]. Currently, it is estimated that 35-60% of clinicians experience symptoms of burnout [2], though one large study of medical interns in the United States found rates of up to 79% in subscales of the Maslach Burnout Inventory [6]. These rates are higher than in other professions [7], suggesting factors intrinsic to medicine drive burnout. Importantly, burnout is a consequence of employment and not a mental health disorder [2].

Burnout affects not only the health of clinicians but also that of patients and the healthcare system overall. As studied in physicians, burnout is associated with increased risks of depression, substance abuse, occupational injury, and suicide [2, 7]. Patients also suffer when clinicians are burned out. Increases in medical errors, recovery times, and patient mortality, as well as decreased patient satisfaction, have been associated with healthcare team burnout [2, 7]. From the point of view of the healthcare system, burnout is costly. Physician burnout is associated with decreased productivity, absenteeism, intent to leave one’s practice or medicine altogether, and increased malpractice claims [2, 7]. Physician burnout is estimated to cost over $4 billion annually [2].

There is no single cause of burnout in healthcare providers. The National Academy of Medicine in 2019 released a 333-page report on clinician burnout [2] that identified contributors such as long work hours; low nurse-to-patient ratios; patient messaging portals; ethical tensions between serving individual patients, the patient population, and the healthcare system; excessive documentation; threats of harm; and the tendency to always place patient needs above one’s own. The report also identified administrative burdens, such as those related to EHRs, as “the most prominent current complaint by clinicians about their workplaces,” and specifically highlighted how poor usability of healthcare technology contributes to clinician burnout.

Clinical decision support (CDS) is an important aspect of modern EHRs, with the potential to limit patient care errors and improve adherence to evidence-based medicine. For example, CDS alerts can identify
if a patient is allergic to a prescription or if they are due for a flu shot. However, although CDS systems can reduce some errors, they can also be the cause of others, such as treatment delays [8]. CDS has had several impacts on clinicians, including changes in workflow, more or new kinds of work, alert/alarm fatigue, spillover of work into personal time, loss of autonomy, and anxiety about the medico-legal ramifications of CDS suggestions [2,7–9]. Concerns about these changes are not new [10], and many were highlighted in the Ten Commandments for Effective Clinical Decision Support: Making the Practice of Evidence-based Medicine a Reality, published in 2003 [11]. Nonetheless, it remains unclear whether factors that affect clinician burnout are being considered and used to inform CDS design and deployment in vivo. For this reason, we undertook a survey of two years of publications on CDS and clinician burnout, in order to identify current trends and synthesize evidence-based standards for consideration in any CDS implementation project.

2 Methods

Two databases, PubMed and Web of Science™ (WoS), were queried for relevant studies in English published between January 1st, 2018 and December 31st, 2019, yielding 189 total results, which was narrowed to 150 articles when duplicates were excluded (Figure 1). To capture CDS, terms including “cognitive aid,” “user interface,” “expert system,” “decision support systems,” “cds,” “clinical reminder,” “best practice advisory,” “best practice alert,” “decision support,” “Decision Support Systems, Clinical” [MeSH], and “Decision Support Systems, Management” [MeSH] were included in the search terms. To capture burnout concepts, the terms “burnout,” “Stress, Psychological”[MeSH], “pain point,” “workaround,” “psychology,” “stress,” “stressful,” “fatigue,” “alarm fatigue,” “alert fatigue,” “provider resistance,” and “provider satisfaction” were used. The general terms “alarm,” “alert,” and “human factors” were not included as the search results were too broad.

Articles were excluded if their relevance could not be determined (e.g., abstract was not available), if the CDS was not related to technology (e.g., a paper worksheet), if the CDS was not aimed at clinicians, or if the tool was used exclusively for teaching. After applying the exclusion criteria, the remaining articles were evaluated for their themes and study design, whether the evaluation of factors that could contribute to burnout was actively incorporated into the design of the CDS tool, and which metrics, if any, were used to evaluate CDS tools. Study design was categorized from each paper’s methods section and adapted to research hierarchies [12–14]. Metrics and themes were considered relevant if they addressed aspects of usability, user preferences or frustrations, burden, user efficiency at tasks, or other concepts related to burnout and wellness as described in section 3.2, even if the term “burnout” was not used explicitly. If these metrics could not be ascertained from the abstract, the whole paper was reviewed, which was the case for the majority of papers.

3 Results

Of the 89 articles that met inclusion criteria, 12 were identified as review articles, of which seven were systematic reviews. Two were editorials. Five described randomized controlled trials (RCTs). Seventeen studies included a comparator group, such as pre-post or cohort designs. Five publications were pre-trial descriptions of intended studies. The remaining 48 publications were descriptive, including cross-sectional and case study designs (Table 1). Themes and features of review articles and non-review articles were considered separately.

3.1 The Big Picture — Review Articles

The 12 review articles that met inclusion criteria in this literature survey can be broadly categorized as either prescriptive or analytic. The prescriptive reviews offered a perspective on how CDS should work. For example, Marcial et al. proposed that CDS
should support individual patients as well as their clinical care teams and emphasized the previously-established “CDS Five Rights” [15]: the right information at the right time, given to the right people, in the right format, via the right channel [16]. Other articles argued that nurse practitioners, as stakeholders in CDS, should be involved in design, implementation, and optimization [17, 18]. Wilbanks and McMullen reviewed cognitive workload associated with EHRs and generally found that it is hard to measure with validity. Specifically regarding CDS, they proposed that the related concepts of alarm fatigue and desensitization were the most pressing for patient safety and they suggested minimizing within-patient identical alerts to ameliorate the negative effects on cognitive load [19]. Finally, Tolley et al. systematically reviewed CDS literature from 2007-2014 to identify areas of improvement. They suggested that sensitivity and specificity of alerts should always be considered, with an emphasis on high specificity to decrease alert fatigue. They also proposed tiering of alerts to improve acceptance and consideration of human factors in CDS design [20].

The analytic reviews analyzed the features of previously implemented CDS tools. Several reviews examined alert use, generally finding acceptance rates to be low [21] and override rates to be high [22–24]. Hussain et al. delved into why alerts are so commonly overridden and found that although interruptive alerts are the most common design, they are also the least accepted. The authors also found that role-tailoring alerts to pharmacists instead of physicians had improved acceptance [21]. Carli et al. undertook a systematic review of positive predictive values (PPVs) of CDS alerts as a proxy for clinical relevance, finding massive PPV variations from 0% to 97%, with most in the 20%-40% range. PPVs were higher when the contextual information of the individual patient was considered in CDS, so the authors concluded that incorporating as much specific patient data as possible to improve PPV is more important than using a large database of knowledge that may generate many false positives and thus promote alert fatigue [24]. In a systematic review to understand CDS for drug allergies, Légat et al. found that significant inaccuracies in the EHR due to the difficulty of structured documentation of allergies, the lack of standard allergy terminology, and irregular updating of allergy databases resulted in poor specificity and overall performance of CDS allergy alerts, along with override rates of up to 90%. The authors emphasized that clinicians and programmers should review CDS at regular intervals post-implementation to ensure that the allergy rules are up-to-date and to minimize alert fatigue [23].

Finally, Powers et al. evaluated 32 articles on “hard stop” CDS in which the clinician could not proceed without third-party override, if at all, and the effects on healthcare delivery. The majority of studies showed improvement in process outcomes and four of eight that reported patient outcomes found improvements. Only two studies pre-specified patient health outcomes, with one finding no adverse effect and the other being stopped early due to treatment delays from the hard stop alert. In evaluating the user experience, the authors found that CDS for which the end users were involved in design and iteration were the most acceptable. Hard stops for documentation purposes were poorly received [25].

### 3.2 How and How Much — Methods and Metrics

The 75 non-review, non-opinion articles used a variety of metrics to evaluate the usability and acceptability of their CDS tools, with many using more than one metric. The most common evaluation was via interviews (n=29) [26–54], followed by surveys or questionnaires (n=16) [27, 29, 50, 53, 55–66]. Other qualitative feedback responses were obtained by focus groups or workshops (n=5) [43, 57, 61, 67, 68]. The alert firing rate (n=12) [43, 55, 62, 69–77] and alert acceptance rate (n=15) [27, 29, 43, 58, 61, 62, 69, 74, 78–84] were also common metrics used to understand alarm fatigue and acceptability. Although metrics of relevance such as sensitivity, specificity, and positive predictive value (PPV) were referenced several times [77, 84–88], only one study evaluated these metrics as pre-specified outcomes [49]. Tools for capturing clinician burnout, stress, or satisfaction after using a CDS tool were used in nine studies [26, 44, 78, 80, 82, 89–92]. Psychometric measures of the user interface, such as mouse clicks, eye movement tracking, time spent on the page, errors using the tool, or direct observation were captured in seven studies [29, 49, 54, 59, 82, 89, 93]. Efficiency metrics of task completion time or clinic appointment time were obtained in seven studies [29, 44, 60, 61, 89, 91, 94]. How often a non-mandatory CDS tool was used, a proxy of whether clinicians found the tool useful and acceptable, was studied in five publications [27, 53, 56, 93, 95]. Whether or not a CDS tool made a difference in patient outcomes, such as cardiovascular events [48], A1c [46], or mortality [80], was evaluated in 11 studies [28, 43, 47, 48, 77, 80–82, 90, 96, 97]. Only one study considered the economic cost of a CDS system [48].

### Table 1 Categorization of the design of each publication from the results of the literature survey

<table>
<thead>
<tr>
<th>Study type</th>
<th>Manuscripts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systematic Review or Meta-Analysis</td>
<td>18, 20, 21, 23–25, 108</td>
</tr>
<tr>
<td>Review</td>
<td>16, 17, 19, 22, 109</td>
</tr>
<tr>
<td>Randomized Controlled Trial</td>
<td>78, 80, 82, 92, 97</td>
</tr>
<tr>
<td>Controlled Studies</td>
<td>27, 46, 49, 51, 55, 56, 69, 72, 73, 76, 79, 81, 85, 89, 91, 93, 94</td>
</tr>
<tr>
<td>Descriptive Studies</td>
<td>26, 28–41, 43–45, 47, 50, 52, 53, 57–68, 70, 71, 74, 75, 77, 83, 84, 87, 90, 95, 96, 98, 99, 104</td>
</tr>
<tr>
<td>Opinion or Editorial</td>
<td>102, 110</td>
</tr>
<tr>
<td>Pre-Trial Study Proposal</td>
<td>42, 48, 54, 86, 88</td>
</tr>
</tbody>
</table>
3.3 Personalized Medicine (Alerts) – The Relevance of CDS Prompts

Similar to the review articles, the relevance of alerts to patients featured prominently in studies of CDS systems. As one article put it: “high PPV is critical for successful deployment of clinical decision support interventions” [49]. Irrelevant alerts fail to improve patient care, contribute to alarm fatigue, and interrupt workflow, all of which have negative effects on clinician wellbeing and that of patients. For example, one study found that their sepsis alert had a PPV of only 14.6% and was associated with a 66.6% override rate, which the authors felt negatively impacted the ability of the CDS to improve patient outcomes [49]. In another case, CDS for a precision-medicine genotype test to detect patients at risk of a drug toxicity yielded 71% normal status and no case of complete deficiency in tested patients. Dosing was adjusted in only three patients after 500 alerts [96]. Kizzier-Carnahan et al. point out, “the average clinician must deal with both the inappropriate presence as well as the inappropriate absence of alerts,” [98] and targeting alerts to those who stand most to improve would improve clinician trust in the CDS system [97].

Relevance could also be captured indirectly by measuring alert acceptance rates, with the caution that some studies found alerts were accepted just to silence an annoying alert [49], or that alarm fatigue could cause even appropriate alerts to be dismissed [69, 72]. Nonetheless, a cross-sectional study of medication CDS alerts examined the appropriateness of overrides, finding that overall 60% of overrides were appropriate and, in subgroup analysis, >95% of duplicate drug and patient allergy alerts were appropriately overridden [57]. In an example of successful CDS, Wasylewicz et al. describe how relevance was improved in patients with hypokalemia when pharmacists were only alerted if there was no potassium repletion order. Doctors were no longer called about lab results they had already acted upon and their acceptance rate of this targeted alert was 88% [81].

To better capture relevance of CDS to clinicians, one study parlayed “number needed to treat” into “number needed to remind,” a measure of the number of CDS alerts the clinician sees before the desired intervention is undertaken. For the intervention in question, the “number needed to remind” was approximately 83 [45]. Another way of understanding CDS relevance was by directly soliciting feedback from users, for example using a Likert scale. A study of an antibiotic susceptibility CDS found that clinicians using a Likert scale felt that fewer than 30% of alerts were relevant [84]. Improved CDS relevance to individual patients was also high on the wish-lists of clinicians giving feedback on CDS alerts [29, 97].

3.4 Drowning in Information – Alarm Fatigue

In addition to relevance, one of the major themes across studies was the burden of CDS tools. Even with perfect PPV, a thousand daily alerts would soon become overwhelming. There is currently no consensus on how to use CDS without causing alert fatigue [64]. Worries about alert fatigue and the burden of CDS were common in qualitative studies on CDS tools [29, 31, 34, 83, 93, 99], which is perhaps unsurprising given that one study estimated the average ICU clinician is subjected to over 900 active and passive alerts per day [98]. Alarm fatigue can lead to important alerts being overlooked [29]. When one hospital switched from a legacy EHR to a commercial one, the alert burden increased six-fold but the alert acceptance rate plummeted nearly twelve-fold, from 100% to 8.4% for high-severity alerts [69]. This drop suggests that poor relevance or simply alarm fatigue can have a disproportionate effect on CDS burden. One RCT determined that their CDS generated approximately 14,400 alerts during the study period, almost 95% of which were dismissed [80]. Another study described alert rates 51 per 100 orders at one hospital, indicating that an alert was generated for every other order. Over 90% of these alerts were dismissed [72].

3.5 Make it Easy – Workflow, Efficiency, and Integration

The burden of CDS tools is not only an excessive number of alerts, but also how the clinician interacts with CDS. Usability of the tool itself is one example. Encouragingly, several studies emphasized developing tools that incorporated user-centered design, such as clean, concise, and intuitive interfaces [26, 30, 50, 67]. However, several studies cited the need for an improved CDS user interface, such as limiting mouse-clicks or not having the pop-up window block access to the chart [35, 93, 99]. Usability was measured directly in several studies, including by observation or tracking of eye movements, click counts, and use errors. Interestingly, users sometimes had differing opinions on design elements [63, 64].

The ability to use CDS tools efficiently can limit excessive documentation burden and improve efficiency. Relatedly, how CDS incorporated into clinician workflow was commonly addressed. For example, one study found that a large number of antibiotic susceptibility alerts were already addressed by standard workflow (e.g., the clinician reviewing new culture results and adjusting empiric antibiotics appropriately), so early alerts just contributed to alert burden [84]. Another study noted the pharmacists using a tool for antibiotic stewardship were frustrated by automatic log-offs [33]. Additionally, a CDS tool for screening for pregnancy complications in Ghana was criticized by clinicians for adding up to 30 minutes to an antenatal visit [38]. Positively, some studies were specifically targeted at improving workflow, such as early work using natural language processing to pre-populate imaging orders based on clinicians’ notes [87].

One major theme in the usability of CDS addressed in the recent papers was integration. Prior studies have shown that CDS needs to be minimally- interruptive to workflow [100], and literature in this survey emphasizes that this goal is better accomplished by integrating it into the existing EHR. Needing to reference an external database is a deterrent to CDS use [27]. The inability of CDS to understand records from outside the hospital system, such as scanned lab results, can lead to misfiring
of CDS alerts [96]. The lack of integration of CDS into EHRs leads to needing to “do work twice” – once to satisfy the CDS alert and then again, for example, to actually place the order in the EHR [99]. Better integration of CDS into the EHR was identified as a key change that would improve acceptance of decision support systems [34, 78, 83, 85]. For example, automatic documentation of patient vital signs that feed into CDS systems decreased both clinician workload and frustration [91].

3.6 Ends Justify the Means – Measuring Outcomes

Given the potential negative effects on workflow and the risks of alarm fatigue, with the associated burnout-related sequelae, it is important to establish that CDS is accomplishing its patient-centered goals. Prior meta-analyses evaluating RCTs of CDS found only weak evidence, if any, of improved patient outcomes [100, 101]. Only a minority of surveyed studies here explicitly evaluated the effectiveness of the CDS tools. Some studies identified the importance but did not undertake evaluation of effectiveness [36]. One study of a CDS tool for deprescribing found that only 1.2% of alerts resulted in medication discontinuation, which was overall branded ineffective [45]. Another study compared interruptive vs. non-interruptive alerts for ACE inhibitor prescription, finding that the non-interruptive alerts were less annoying but less seen [51]. Unintended consequences should be assessed in addition to expected outcomes [59]. In one case, CDS did not change pre-specified outcomes relating to clinician confidence but was preferred by the users anyway [92].

Studies evaluating the effects of CDS on clinically relevant patient outcomes were uncommon. A CDS tool to point clinicians away from prescribing fluoroquinolone antibiotics showed decreased fluoroquinolone prescription rates but did not examine if patients had fewer adverse events or if their infections resolved appropriately [76]. A hard-stop best practice advisory to promote intensification of blood pressure control in patients with diabetes was successful in increasing treatment, but control of hypertension in the population was not assessed [79]. An evaluation of patient outcomes based on whether the physician accepted or declined CDS found no difference in A1c but statistically significant improvements in blood pressure and low-density lipoprotein (LDL) [46]. As mentioned earlier, a sepsis alert with poor PPV did not change length of stay or mortality [49]. One RCT of CDS found no difference in length of stay or mortality, though a modest increase in alert resolution [80]. Another RCT of a CDS system to assist in the management of patients with cardiovascular risk showed a 2.24% improvement in the surrogate outcome of annual change in predicted 10-year cardiovascular risk [82]. Of the 11 studies that examined CDS effects on end-users as well as patient health outcomes, only two [46, 82] showed benefit in some domains.

3.7 I’m a Believer – Clinician Buy-in

Multiple authors posited that for a CDS tool to be accepted and used by healthcare professionals, they must be convinced that the tool solves a meaningful clinical problem [31, 47, 92, 102], building on previous work on the importance of clinician buy-in for CDS uptake [103]. An investigation into why clinician adoption of a CDS tool was only 4-14% identified lack of buy-in as one of the factors deterring use [93]. Some studies noted concerns that CDS threatened clinician autonomy [67] or that CDS use would be monitored for punitive reasons [33]. For example, users were offended by the CDS prompt to use basal-bolus insulin when only sliding scale insulin had been ordered [67], demonstrating that even when CDS is supporting the evidence-based regimen the prompt needs to be palatable to clinicians. Accordingly, another study proposed increasing buy-in by framing CDS positively instead of negatively and making users feel recognized for their efforts [28]. From the surveyed literature, for there to be optimal clinician buy-in, a CDS system must foster the clinician’s belief that a worthwhile problem is being addressed, that CDS can solve the problem (patient outcomes), that it is targeting the correct patients (sensitivity and specificity, trust), that the alerts are relevant (PPV and alert burden), and that workflow interruptions are worth the benefits, all while using language that lifts up instead of denigrating the clinician. The need for clinician buy-in for CDS success is so significant that Kawamanto et al. state that there should be “a requirement that proposed CDS is actually desired by intended recipients” [75].

3.8 Fix It – Improving CDS

The surveyed studies on CDS and clinician burnout explored how CDS tools could become more acceptable to users. As a participant in one study put it: “[Not having an] alert is better than a poorly designed alert” [65]. The most important theme in improving CDS was iteration, meaning evaluating a tool, revising it, and then evaluating the revised tool. Multiple studies used iterative designs to fine-tune their CDS [26, 34, 42, 68, 74]. Mann et al. provided a description of the development of a CDS tool for antibiotic prescriptions: (i) Pre-deployment, the CDS team collected qualitative feedback during “think aloud” sessions on the draft versions of the tool, (ii) Post-deployment, they solicited feedback in group interviews and measured utilization rates as well as antibiotic prescription rates. Their CDS tool was based on a previous version developed via user-centered iteration that had nearly two-thirds utilization; however, one year after implementation, the utilization of the new tool was 4-14% across sites. Their inability to replicate a previous tool’s utilization success demonstrates that even intensive feedback and iteration may not be able to overcome factors that dissuade clinicians from CDS use, such as the fact that alert fatigue for clinicians had increased since the earlier version [93].

In contrast to the intensive qualitative evaluations of user perceptions undertaken by Mann et al., Yoshida et al. described an automated approach of CDS monitoring. Using this approach, the CDS team monitored for changes in firing rates or patterns indefinitely post-implementation. In addition to automated monitoring, the team undertook targeted monitoring of silent alert firing patterns for two weeks prior to roll-out, including chart review to capture
false positives and negatives, post-implementation monitoring and chart review, and ad hoc monitoring whenever users reported problems. Their automatic monitoring system identified 128 issues with the CDS system over two years, of which 24 were false positives [60].

In a story of successful CDS improvement, Kawamanto et al. described the establishment of a CDS governance that involved a CDS committee, CDS for desired outcomes only, data analytics to monitor alert frequency, a push to switch alerts to other areas of the EHR, and experimental designs to improve effectiveness, resulting in a greater than 50% reduction in alerts [75]. In another example of improvement of CDS, Bubp et al. described the multi-phased development of a drug-disease alert system. In the first phase, pharmacists and physicians scored alerts from a database based on clinical utility and the scores were validated by a survey. During the second phase, alerts fired silently, and alert burden was evaluated. Finally, the alert system was rolled out, and post-implementation analysis showed decreased alert burden and the second-highest acceptance rate of any alert type at nearly 22%, behind dose alerts [62].

Similar to Bubp et al., expert panels were used in several studies to improve the relevance of CDS [62, 71, 88]. Expert panels can be advantageous over stock alerts. For example, one study found only 18% concordance on high severity alerts amongst four drug interaction compendia, whereas an expert panel was able to reach consensus on the relevance of 12/13 alerts [104]. A downside of expert panels is that they do not scale well given time limitations [85, 104].

Customizability was another factor that could improve user experience with CDS. Customizability can be at the scale of the hospital or clinic [69, 73] or by the end-user altering, for example, which alerts are seen [44]. Additionally, role-tailoring, in which CDS is targeted to users by role such as nurse, pharmacist, physician, etc., was also seen as a way to reduce burnout and increase acceptability in CDS [21, 51]. In one study, non-physician care-coordinators were responsible for data entry to support CDS [46]. In another, nurses rooming patients triggered the CDS [82]. Alerts tailored by role could also be used in cases where the ownership of an alert would otherwise be ambiguous, such as results of a genomic test [37]. Separate from role-tailoring, CDS that “nudges” [26] could be used to combat concerns about loss of autonomy [67]. Relevance could also be improved by tiering of alerts [20, 22]. For example, when one hospital disabled the least severe drug interaction alerts, alert burden fell by over 50% [69]. Another hospital was able to decrease “major” alerts by 62% via customization of the stock alerts [73].

Finally, multiple publications discussed the importance of end-user feedback throughout the design, implementation, and modification process. Several publications noted that nurses and nurse practitioners are also stakeholders who should be involved in CDS design [17, 18, 27]. Feedback before iteration was emphasized in multiple studies [42, 47, 68, 74, 93] and “the importance of listening to user experience” was explicitly highlighted [36]. Also important was the idea that implementation does not end with roll-out, but that CDS is an ongoing process of monitoring, feedback, and iteration [52, 60, 62, 65, 72].

4 Discussion
The rapid and widespread implementation of EHRs has contributed to a crisis of clinician dissatisfaction [105]. Clinical decision support is an important aspect of EHRs that is “not merely the use of technology; it is using technology to find meaningful information to make clinical decisions and provide the best possible patient care” [18]. As described above, how CDS is designed and implemented can have significant impacts on clinician users. Alarm fatigue from both a high volume of alerts and alerts of poor relevance, such as low PPV, were commonly described problems, leading to high rates of alert overrides or avoidance of CDS. Although the ideal PPVs, number of alerts, or alert override rates have not been established, this review identifies that generally poor performance of many CDS tools has been reported. Overrides, avoidance, and workarounds decrease effectiveness of CDS [95] in implementing behavioral change, though few studies directly evaluated the effectiveness of CDS on changing health-related patient outcomes. Studies also identified that CDS tools need to have clinician buy-in, which is partly achieved by fostering trust that the CDS is relevant to their patients but also by utilizing user-centered interfaces with integration into the workflow and EHR.

There were many suggestions on how to improve CDS tools in the surveyed literature, however it is not clear that they are being consistently implemented, perhaps because of a lack of randomized trials to guide design [106, 107]. Aspects of CDS that were found to be most helpful and harmful to clinician burnout in the recent RCTs and systematic reviews are highlighted in Table 2. As the meta-analysis by West et al. demonstrated, structural and organizational changes can result in meaningful reductions in burnout [107], so organizations designing and implementing CDS could potentially play a significant role in the wellness of front-line clinicians.

Based on our survey of the recent literature, we have consolidated recurrent ideas into key factors that should be considered when designing and implementing CDS in order to minimize the effects on clinician burnout:

1. **Be relevant.** CDS should solve problems that clinicians feel need to be solved. CDS alerts should incorporate as much patient-specific information as possible to maximize PPV and minimize the number needed to remind.

2. **Solicit feedback.** End-users should be involved in all aspects of design, pre-testing, and implementation.

3. **Customize.** Whether allowing expert panels to tier alerts or clinicians to choose how and when to see CDS tools, customization can minimize alert burden and improve relevance as well as clinician satisfaction.

4. **Measure outcomes.** The effects on alert burden, override rates, workflow, efficiency, burnout, satisfaction, patient outcomes, etc. must be evaluated. Tools should either improve efficiency, patient outcomes, or both. Tools that do neither should be abandoned, especially if they add to alert burden or burnout.

5. **Iterate.** CDS requires ongoing maintenance based on feedback and outcomes, as well as updates to clinical practice standards.
controlled trials in this literature survey.

Factors that ameliorate (helpful) or contribute to (harmful) clinician burnout as extracted from the systematic reviews and randomized controlled trials, or themes may have been misclassified or overlooked.

### Key Factors of CDS and Burnout from Systematic Reviews and Randomized Controlled Trials

<table>
<thead>
<tr>
<th>Helpful</th>
<th>Harmful</th>
</tr>
</thead>
<tbody>
<tr>
<td>High sensitivity, specificity, and positive predictive value [20, 24, 97]</td>
<td>Interruptive or hard stop alerts [21, 25, 80]</td>
</tr>
<tr>
<td>Role tailoring of alerts [21, 82]</td>
<td>Lack of “right information” [18]</td>
</tr>
<tr>
<td>Tiering of alerts [29, 80]</td>
<td>Lack of relevance [80, 97, 108]</td>
</tr>
<tr>
<td>Iterative design [25]</td>
<td>Lack of integration [78]</td>
</tr>
</tbody>
</table>

Using these principles, future CDS tools can minimize their impacts on the multifactorial problem that is clinician burnout.

A strength of this study is that it included a wide array of healthcare professionals’ experiences including physicians [37, 62], advanced practice providers [17, 18, 63], nurses [90, 92], pharmacists [29, 33], midwives [38], and paramedics [59]. We also surveyed a heterogenous mix of publication types including systematic reviews, randomized trials, quality improvement initiatives, and opinion pieces. A limitation is that manual review of articles was undertaken by one individual and some relevant articles, study designs, or themes may have been misclassified or overlooked.

### 5 Conclusion

Clinical decision support tools can contribute to clinician frustration and burnout. Using the techniques of improving relevance, soliciting feedback, customization, measurement of outcomes and metrics, and iteration, the effects of CDS on burnout can be ameliorated.

### Acknowledgements

We would like to thank Stanford Information Services Librarian Christopher Stave for his assistance in formulating the database queries. This research was supported in part by the NIH Big Data 2 Knowledge initiative via the National Institute of Environmental Health Sciences under Award Number K01ES026837 as well as by the Diabetes, Endocrinology, and Metabolism Training Grant 5T32DK007217-44. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

### References


