



OPEN ACCESS



# Computerised clinical decision support systems and absolute improvements in care: meta-analysis of controlled clinical trials

Janice L Kwan,<sup>1,2</sup> Lisha Lo,<sup>3</sup> Jacob Ferguson,<sup>4</sup> Hanna Goldberg,<sup>4</sup> Juan Pablo Diaz-Martinez,<sup>5</sup> George Tomlinson,<sup>5</sup> Jeremy M Grimshaw,<sup>6</sup> Kaveh G Shojania<sup>2,3,7</sup>

<sup>1</sup>Sinai Health System, Department of Medicine, 600 University Avenue, Toronto, ON M5G 1X5, Canada

<sup>2</sup>Department of Medicine, University of Toronto, Toronto, ON, Canada

<sup>3</sup>Centre for Quality Improvement and Patient Safety, University of Toronto, Toronto, ON, Canada

<sup>4</sup>Faculty of Medicine, University of Toronto, Toronto, ON, Canada

<sup>5</sup>Biostatistics Research Unit, University Health Network and Sinai Health System, Toronto, ON, Canada

<sup>6</sup>Clinical Epidemiology Program, Ottawa Hospital Research Institute and Department of Medicine, University of Ottawa, Ottawa, ON, Canada

<sup>7</sup>Department of Medicine, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

Correspondence to: Janice L Kwan [janice.kwan@utoronto.ca](mailto:janice.kwan@utoronto.ca) (or @KwanJanice on Twitter ORCID 0000-0001-9252-0609)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2020;370:m3216 <http://dx.doi.org/10.1136/bmj.m3216>

Accepted: 7 August 2020

## ABSTRACT

### OBJECTIVE

To report the improvements achieved with clinical decision support systems and examine the heterogeneity from pooling effects across diverse clinical settings and intervention targets.

### DESIGN

Systematic review and meta-analysis.

### DATA SOURCES

Medline up to August 2019.

### ELIGIBILITY CRITERIA FOR SELECTING STUDIES AND METHODS

Randomised or quasi-randomised controlled trials reporting absolute improvements in the percentage of patients receiving care recommended by clinical decision support systems. Multilevel meta-analysis accounted for within study clustering. Meta-regression was used to assess the degree to which the features of clinical decision support systems and study characteristics reduced heterogeneity in effect sizes. Where reported, clinical endpoints were also captured.

### RESULTS

In 108 studies (94 randomised, 14 quasi-randomised), reporting 122 trials that provided analysable data from 1 203 053 patients and 10 790 providers, clinical decision support systems increased the proportion of patients receiving

desired care by 5.8% (95% confidence interval 4.0% to 7.6%). This pooled effect exhibited substantial heterogeneity ( $I^2=76\%$ ), with the top quartile of reported improvements ranging from 10% to 62%. In 30 trials reporting clinical endpoints, clinical decision support systems increased the proportion of patients achieving guideline based targets (eg, blood pressure or lipid control) by a median of 0.3% (interquartile range  $-0.7\%$  to  $1.9\%$ ). Two study characteristics (low baseline adherence and paediatric settings) were associated with significantly larger effects. Inclusion of these covariates in the multivariable meta-regression, however, did not reduce heterogeneity.

### CONCLUSIONS

Most interventions with clinical decision support systems appear to achieve small to moderate improvements in targeted processes of care, a finding confirmed by the small changes in clinical endpoints found in studies that reported them. A minority of studies achieved substantial increases in the delivery of recommended care, but predictors of these more meaningful improvements remain undefined.

### Introduction

Although the first electronic health record (EHR) was introduced almost six decades ago,<sup>1</sup> dissemination has occurred surprisingly slowly.<sup>2</sup> In the United States, as recently as 2008, fewer than 10% of hospitals had EHRs, and only 17% had computerised order entry of drugs. By 2015, however, 75% of US hospitals had adopted at least basic EHR systems.<sup>3</sup> This rapid uptake reflects the financial incentives in the Health Information Technology for Economic and Clinical Health (HITECH) Act passed in 2009.<sup>4</sup> Although EHR adoption has been slower in the UK,<sup>5</sup> the NHS long term plan, released in 2019, sets a goal for all trusts and providers to move to digital health records by 2024.<sup>6</sup>

Clinical decision support systems embedded within EHRs—from pop-up alerts about serious drug allergies to more sophisticated tools incorporating clinical prediction rules—prompt clinicians to deliver evidence based processes of care,<sup>7</sup> discourage non-indicated care,<sup>8 9</sup> optimise drug orders,<sup>10 11</sup> and improve documentation.<sup>12 13</sup> Despite optimism over the effects of these support systems,<sup>14-16</sup> a systematic review published by our group in 2010<sup>17</sup> found that clinical decision support systems typically improved the proportion of patients who received target processes of care by less than 5%. The subsequent decade has seen a dramatic rise in the application and evaluation of clinical decision support systems. Systematic reviews of an increasingly large number of

## WHAT IS ALREADY KNOWN ON THIS TOPIC

Clinical decision support systems embedded in electronic health records prompt clinicians to deliver recommended processes of care

Despite enthusiasm over the potential for clinical decision support systems to improve care, a previous systematic review in 2010 found that such systems typically improved the proportion of patients receiving target processes of care by less than 5%

The number of trials published over the ensuing decade has grown considerably, but subsequent systematic reviews have focused only on identifying features of clinical decision support systems associated with positive results, rather than quantifying the actual sizes of improvements achieved

## WHAT THIS STUDY ADDS

Most clinical decision support system interventions achieve small to moderate improvements in the percentage of patients receiving recommended processes of care

The pooled effect size exhibited extreme heterogeneity (that is, variation across trials beyond that expected from chance alone), which did not diminish with a meta-regression model using significant predictors of larger effect sizes

Thus although a minority of studies have shown that clinical decision support systems deliver clinically worthwhile increases in recommended care, the circumstances under which such improvements occur remain undefined

publications, however, have typically looked only at the features of these support systems that predicted improvements in care,<sup>14 18</sup> and reported odds ratios or risk ratios,<sup>19-21</sup> without quantifying the actual sizes of the improvements achieved.

In our systematic review and meta-analysis, we sought firstly, to estimate the typical improvement in processes of care—and thus the potential for clinical effect—conferred by clinical decision support systems delivered at the point of care; and secondly, to identify any study characteristics or features of these support systems consistently associated with larger effects.

### Methods

We followed established methods recommended by Cochrane<sup>22</sup> and report our findings in accordance with PRISMA (preferred reporting items for systematic reviews and meta-analyses).<sup>23</sup>

#### Search strategy and selection criteria

We searched Medline from the earliest available date to August 2019 without language restrictions (supplementary appendix 1) and scanned reference lists from included studies and relevant systematic reviews. We did not search Embase, CINAHL, and Cochrane Central Register of Controlled Trials, as these databases did not increase the yield of eligible studies included in our previous review.<sup>17 24</sup> We did, however, conduct fresh study screening and data abstraction for all articles, even those covered by the previous review, to accommodate new and modified data elements reflecting changes in technology in the intervening decade.<sup>17 24</sup>

Eligible studies evaluated the effects of clinical decision support systems on processes or outcomes of care using a randomised or quasi-randomised controlled design (allocation on the basis of an arbitrary but not truly random process, such as even or odd patient identification numbers). Patients in control arms received “usual care” contemporaneous with care delivered in the intervention arm (that is, we excluded head-to-head comparisons of different clinical decision support systems).

We defined a clinical decision support system as any on-screen tool designed to improve adherence of physicians to a recommended process of care. Eligible studies delivered the support system intervention within a clinical information system routinely used by the provider (not a computer application separate from the EHR) at the time of providing care to the targeted patient (eg, while entering an order or a clinical note). We excluded specialised diagnostic decision support systems (eg, in medical imaging) and systems not directly related to patient care, such as decision support for billing or health record coding. We also excluded studies using simulated patients and studies in which fewer than half of participants were physicians or physician trainees.

We focused on improvements in processes of care (eg, prescribing drugs, immunisations, test ordering, documentation), rather than clinical outcomes,

because we sought to determine the degree to which clinical decision support systems achieve their immediate goal—namely, changing provider behaviour. The extent to which such changes ultimately improve patient outcomes will vary depending on the strength of the relation between targeted processes and clinical outcomes. Nonetheless, we did capture clinical outcomes where reported, including intermediate endpoints, such as haemoglobin A1c and blood pressure. We coded all results so that larger numbers always corresponded to improvements in care. For example, if a study reported the proportion of patients who received inappropriate drugs,<sup>8 25</sup> we recorded the complementary proportion of patients who did not receive inappropriate drugs.

Two investigators independently evaluated the eligibility of all identified studies based on titles and abstracts. Studies not excluded in this first step were independently assessed for inclusion after full text evaluation by two investigators. For articles that met all inclusion criteria, two investigators independently extracted the following information: clinical setting, participants, methodological details, characteristics of the design and content of the clinical decision support system, presence of educational and non-educational co-interventions, and outcomes. We extracted studies with more than one eligible intervention arm as separate trials (that is, comparisons). Two investigators also independently assessed risk of bias for each study using criteria outlined in the Cochrane Handbook for Systematic Reviews of Interventions.<sup>26 27</sup> Specific biases considered for each study included selective enrolment, attrition bias, similarity of baseline characteristics, unit of analysis errors, performance bias (systematic differences between groups with respect to the care provided or exposures other than the interventions of interest), and detection bias (systematic differences between groups in outcome ascertainment). We resolved all discrepancies and disagreements by discussion and consensus among the study team.

#### Data analysis

We used a multilevel meta-analysis model to estimate the absolute improvements (risk differences) in processes of care between intervention and control groups, using the number of patients receiving the target process of care and the total sample size for each reported outcome. This approach allowed us to account for heterogeneity between studies, and for clustering of multiple outcomes reported for the same patients within a given study.

Most trials used clustered designs, assigning intervention status to the provider or provider group rather than to the individual patient, but did not always report cluster adjusted estimates. We accounted for clustering by multiplying the standard error of the risk differences by the square root of the design effect.<sup>22</sup> Intraclass correlation coefficients were abstracted directly from the study when reported. To impute intraclass correlation coefficients in studies

that did not report them, we used a published database of intraclass correlation coefficients<sup>28</sup> stratified by type of setting (eg, hospital versus ambulatory care) and endpoint (eg, process versus outcome). We calculated the median intraclass correlation coefficient for hospital and ambulatory process measures across the 200 studies in this database and applied the relevant value to a given study.

For clinical endpoints, we quantified the median improvement and interquartile range across all studies that reported dichotomous clinical endpoints, such as the percentage of patients who achieved a target blood pressure or the percentage of patients who experienced a clinical event (eg, a critical laboratory result, adverse drug event, or venous thromboembolism). This method<sup>29</sup> for summarising the effects of improvement interventions on disparate clinical endpoints was first developed in a large systematic review of implementation strategies for clinical practice guidelines,<sup>30</sup> and subsequently applied in other systematic reviews.<sup>31 32</sup> We also calculated the median improvement and interquartile range for changes in blood pressure, the most commonly reported continuous clinical endpoint.

We performed univariate meta-regression analyses to explore the extent to which effects varied according to the study characteristics and features of the clinical decision support system. These analyses estimated the difference in absolute improvements reported between studies with and without each feature. Most features are easy to understand (eg, the setting of the intervention as hospital or ambulatory, the presence of co-interventions other than clinical decision support system). Table 1 defines those features with less obvious meanings.

Finally, we fitted a multivariable meta-regression model that included covariates with  $P < 0.1$  in the univariate analyses, and simplified the model by stepwise selection.<sup>33</sup> We used the  $I^2$  statistic to summarise statistical heterogeneity (that is, the degree to which trials exhibited non-random variation in effect sizes). We expected substantial heterogeneity in effect sizes given the range of care settings, targeted clinician

behaviours, and design features of the clinical decision support system. The multivariable meta-regression model was used to identify study and clinical decision support system features that predicted larger effects and to determine if heterogeneity could be reduced.

We performed all statistical analyses using R Software, version 3.5.1 (R Foundation for Statistical Computing, Vienna). *Rma.mv* function from *metafor* library was used to fit all models.

### Patient and public involvement

Patients or the public were not involved in the design, conduct, reporting, or dissemination plans of this study. There were no funds or time allocated for patient and public involvement at the time of the study so we were unable to involve patients.

### Results

Our search identified 5895 citations, of which 5428 were excluded at the initial stage of screening and an additional 352 on full text review, yielding a total of 115 studies that met all inclusion criteria (fig 1 and supplementary appendix 2). Nine studies contained two trials (that is, two comparisons of a clinical decision support system intervention with a control group)<sup>7 34-41</sup> and one study contained six such trials,<sup>42</sup> resulting in 129 included trials.

Of the 129 included trials (table 2 and supplementary table 1), most used a true randomised design, with only 16 (12%) of the 129 trials involving a quasi-random allocation process. Most trials (113/129; 88%) had a clustered design, allocating intervention status to providers or provider groups rather than patients. Most trials (93/129; 72%) occurred in outpatient settings, and 85 of 129 trials (66%) came from US centres. Moreover, 93 (72%) of all 129 trials were published from 2009 onwards, when the US HITECH Act became law. The period since 2009 also included significantly more trials using commercial EHRs (39/93 (42%) v 2/36 (6%),  $P < 0.05$  with Bonferroni correction for multiple comparisons). Epic accounted for 26 (63%) of all 41 interventions on commercial EHR platforms. As shown in supplementary figures 1a-b, the risk of

**Table 1 | Study and clinical decision support system features**

| Study or CDSS feature             | Definition  |
|-----------------------------------|---|
| Acknowledgement and documentation | Requiring the user to register receipt of the information (eg, clicking on a YES or OK button) and to record the rationale for the clinical decision (eg, selecting from a dropdown list of guideline based indications when ordering a broad spectrum antibiotic)  |
| Ambush                            | A CDSS that is likely to appear when not relevant to the immediate clinical task being performed by the user (eg, a pop-up prompting the user to order vaccinations upon opening a patient's record at all visits, not just preventative visits)  |
| Baseline adherence                | The proportion of patients in the control group receiving the desired process of care. In studies that reported pre-intervention data, the mean percentage of patients in the pre-intervention control and intervention groups was calculated   |
| Complex decision support          | A CDSS that incorporates two or more pieces of clinical or demographic information to guide clinical decision making (eg, a tool that recommends deep vein thrombosis prophylaxis for a newly admitted inpatient over the age of 65). This complex decision support is in contrast to simple decision support, which incorporates no pieces, or one piece, of clinical or demographic information                             |
| Consideration of alert fatigue    | Studies that specifically mentioned alert fatigue in designing or delivering the CDSS intervention under study  |
| Interruptive                      | A CDSS that appears on screen and stops the user from performing any further tasks until it is either "acknowledged", or a button is pressed, such as the Escape key  |
| Push                              | A CDSS is fully presented to the user without requiring additional clicks versus "pull", in which the user is required to click a link or icon to receive information contained within the CDSS   |
| Target underuse                   | A CDSS that aims to increase the percentage of patients who receive a recommended process of care (eg, adherence to guideline recommended care). Systems that target underuse are in contrast to those systems that target overuse, whereby improvements correspond to reductions in the percentage of patients receiving inappropriate or unnecessary processes of care (eg, daily bloodwork in a clinically stable patient) |

CDSS=clinical decision support system.

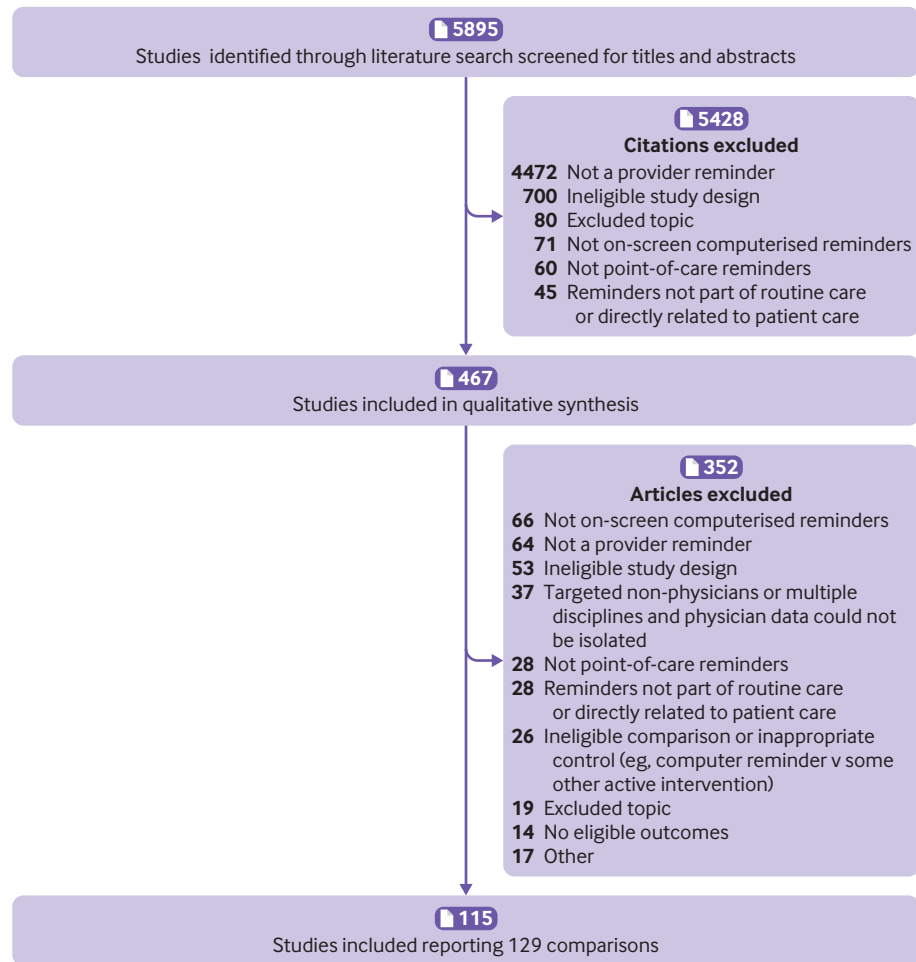


Fig 1 | Flow of studies through the review process

bias was generally low. The one exception involved unit of analysis errors, which occurred in 18 (16%) of 113 clustered studies, and thus possessed a high risk of bias. (We included these studies because our use of intraclass correlation coefficients and a multilevel model avoided replicating unit of analysis errors from the primary studies.)

Across the 122 trials that provided analysable data from 1 203 053 patients and 10 790 providers (fig 2), clinical decision support systems produced an average absolute improvement of 5.8% (95% confidence interval 4.0% to 7.6%) in the percentage of patients receiving the desired process of care. For specific types of physician behaviours, prescribing improved by 4.4% (95% confidence interval 2.6% to 6.2%; 68 trials); test ordering by 6.8% (4.9% to 8.6%; 30 trials); documentation adherence by 7.1% (5.4% to 8.9%; 25 trials); vaccination by 5.9% (4.1% to 7.7%; 12 trials); and other process measures by 6.8% (5.0% to 8.6%; 35 trials). Other process measures included referral for specialty consultations,<sup>43-45</sup> overall guideline concordance,<sup>46-48</sup> and documenting key elements of the diagnostic process.<sup>49-51</sup> Supplementary figures 2a-e present the multilevel models for each category.

In univariate meta-regression analyses (table 3), clinical decision support systems requiring

acknowledgement and documentation of a reason for not adhering to the recommended action achieved improvements 4.8% larger than support systems without this feature (95% confidence interval 0.1% to 9.6%). The ability to execute the desired action through the clinical decision support system was also associated with larger effects than interventions without this feature, with an incremental difference of 4.4% (0.9% to 7.9%). Only 17 (14%) of 122 trials considered alert fatigue in designing or delivering the clinical decision support system, and the associated incremental increase for this feature was small (1.7%) and non-significant (95% confidence interval -3.5% to 6.8%;  $P=0.52$ ).

Improvements seen for clinical decision support systems in paediatric settings exceeded those in other patient populations by 14.7% (95% confidence interval 8.4% to 21.0%; table 4). Additional study features associated with significant incremental effects included small studies (lower than the median patient sample size), with an incremental improvement of 3.7% (0.2% to 7.3%) greater than larger studies, and studies that took place in the US, with an incremental improvement of 3.7% (0% to 7.4%).

As expected, the meta-analytic improvement in processes of care across all trials exhibited substantial

**Table 2 | Summary of characteristics of included trials, according to publication year. Data are number (%) of trials unless stated otherwise**

| Characteristics of trials         | Before 2009 (n=36) | 2009 onwards (n=93) | Total (n=129)   |
|-----------------------------------|--------------------|---------------------|-----------------|
| Targeted process of care:         |                    |                     |                 |
| Prescribing                       | 24 (67)            | 44 (47)             | 68 (53)         |
| Test ordering                     | 12 (33)            | 18 (19)             | 30 (23)         |
| Documentation                     | 3 (8)              | 22 (24)             | 25 (19)         |
| Vaccination                       | 5 (14)             | 7 (8)               | 12 (9)          |
| Other                             | 7 (19)             | 28 (30)             | 35 (27)         |
| Study features:                   |                    |                     |                 |
| Randomised controlled trial       | 30 (83)            | 83 (89)             | 113 (88)        |
| United States                     | 23 (64)            | 62 (67)             | 85 (66)         |
| Outpatient                        | 23 (64)            | 70 (75)             | 93 (72)         |
| No of providers (median (IQR))    | 89 (41-151)        | 79 (36-171)         | 81 (36-171)     |
| No of patients (median (IQR))     | 2254 (452-6275)    | 2204 (450-11 282)   | 2237 (450-8574) |
| Baseline adherence (median (IQR)) | 29 (21-50)         | 45 (17-64)          | 39 (17-62)      |
| Co-interventions (education)      | 12 (33)            | 48 (52)             | 60 (47)         |
| Co-interventions (non-education)  | 8 (22)             | 28 (30)             | 36 (28)         |
| CDSS features:                    |                    |                     |                 |
| Push                              | 22 (61)            | 72 (77)             | 94 (73)         |
| Complex decision support          | 3 (8)              | 40 (43)             | 43 (33)         |
| Acknowledgement and documentation | 2 (6)              | 18 (19)             | 20 (16)         |
| Execute action through CDSS       | 14 (39)            | 46 (49)             | 60 (47)         |
| Ambush                            | 3 (8)              | 30 (32)             | 33 (26)         |
| Interruptive                      | 14 (39)            | 45 (48)             | 59 (46)         |
| Other concurrent CDSS             | 2 (6)              | 12 (13)             | 14 (11)         |
| Consideration of alert fatigue    | 4 (11)             | 15 (16)             | 19 (15)         |
| Target underuse                   | 27 (75)            | 57 (61)             | 84 (65)         |
| Commercial EHR (eg, Epic)*        | 2 (6)              | 39 (42)             | 41 (32)         |

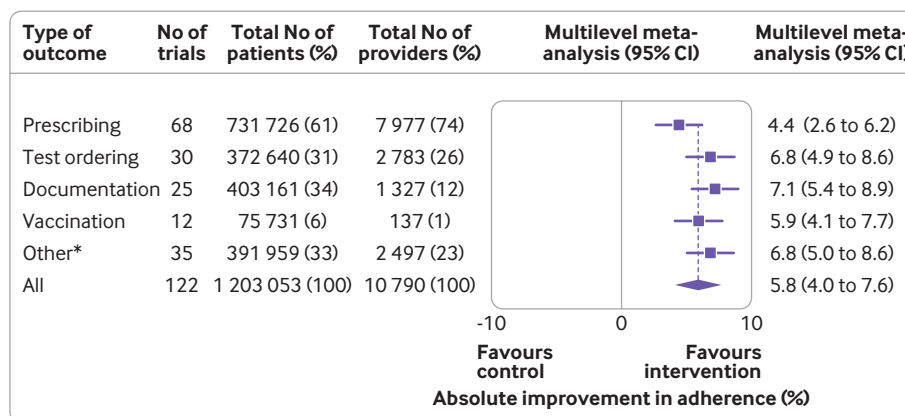
CDSS=clinical decision support system; EHR=electronic health record; IQR=interquartile range.

\*Proportion of studies involving commercial systems increased significantly on or after 2009 compared with earlier (Bonferroni adjusted  $P < 0.05$ ).

heterogeneity ( $I^2=76\%$ ), with a minority of studies reporting much larger improvements than the meta-analytic average. The top quartile of trials reported improvements in process adherence ranging from 10% to 62%. Using multivariable meta-regression (table 5), the final model identified paediatric studies as achieving incremental improvements of 13.6% (95% confidence interval 7.4% to 19.8%), and those trials with low baseline adherence (relative to the median across all studies) reported incrementally greater improvements by 3.2% (0% to 6.4%). Even when these characteristics were incorporated in the meta-

regression model, heterogeneity remained high and essentially unchanged.

Thirty trials reported dichotomous clinical endpoints (supplementary table 1). These endpoints included achieving guideline based targets for blood pressure<sup>52-54</sup> and lipid levels<sup>53-54</sup>; adverse events, such as bleeding<sup>55-56</sup>; in-hospital pulmonary embolism<sup>57-58</sup>; hospital readmission<sup>59-60</sup>; and mortality.<sup>58-60</sup> For these various endpoints, clinical decision support systems increased the proportion of patients achieving guideline based targets by a median of 0.3% (interquartile range -0.7% to 1.9%). Twenty trials reported continuous



**Fig 2 | Absolute improvements in desired care by different categories of clinical care. Results from the multilevel random effects meta-analysis are shown. The diamond shows the summary overall absolute improvement and 95% confidence interval across all types of outcomes; the squares with lines represent estimates and their 95% confidence intervals for different categories of clinical care. \*Other process outcomes included referrals for specialty consultations, overall guideline concordance, and diagnosis**

Table 3 | Absolute incremental improvements in desired care by CDSS feature

| CDSS feature                                    | Category (No of trials)      | Absolute incremental improvement (%; 95% CI)* | P value |
|---|------------------------------|---|---------|
| Acknowledgement and documentation required      | Yes (20) v no (102)          | 4.8 (0.1 to 9.6)                              | 0.05    |
| Execute desired action through CDSS             | Yes (58) v no (64)           | 4.4 (0.9 to 7.9)                              | 0.01    |
| Behaviour targeted, when reported               | Underuse (82) v overuse (37) | 4.0 (0.1 to 7.9)                              | 0.05    |
| EHR platform                                    | Epic (25) v other (97)       | 3.8 (-0.5 to 8.2)                             | 0.09    |
| Interruptive                                    | Yes (55) v no (67)           | 3.2 (-0.4 to 6.7)                             | 0.08    |
| Mode of delivery, when reported                 | Push (90) v pull (13)        | 2.4 (-3.5 to 8.2)                             | 0.43    |
| Other concurrent CDSS                           | Yes (14) v no (108)          | 2.1 (-3.4 to 7.6)                             | 0.46    |
| Considered alert fatigue in design              | Yes (17) v no (105)          | 1.7 (-3.5 to 6.8)                             | 0.52    |
| User workflow specifically considered in design | Yes (35) v no (87)           | 1.2 (-2.7 to 5.2)                             | 0.53    |
| Level of decision support                       | Complex (40) v simple (82)   | 0.9 (-2.9 to 4.7)                             | 0.64    |
| Developed in consultation with users            | Yes (7) v no (115)           | 0.1 (-7.3 to 7.5)                             | 0.98    |
| Makes recommendation for care                   | Yes (98) v no (24)           | 0 (-4.4 to 4.5)                               | 0.98    |
| Appearance differed based on urgency            | Yes (3) v no (119)           | -0.5 (-11.4 to 10.4)                          | 0.93    |
| Inclusion of supporting information on screen   | Yes (53) v no (69)           | -1.1 (-4.7 to 2.5)                            | 0.55    |
| Ambush  | Yes (32) v no (90)           | -1.3 (-5.4 to 2.7)                            | 0.51    |
| Developed by study investigators                | Yes (62) v no (60)           | -1.5 (-5.1 to 2.0)                            | 0.40    |
| Requires input of clinical information          | Yes (9) v no (113)           | -1.9 (-8.8 to 4.9)                            | 0.58    |
| Conveys patient-specific information            | Yes (107) v no (15)          | -4.4 (-9.8 to 1.0)                            | 0.11    |

CDSS=clinical decision support system; EHR=electronic health record.

\*The third column shows the estimated pooled difference (and 95% confidence interval) in the percentage of desired care between the trials classified according to the CDSS features in the first column. Results from 122 trials (108 studies) reporting dichotomous outcomes were each estimated in a univariate meta-regression model.

clinical endpoints (supplementary table 1), including laboratory test values (eg, haemoglobin A1c<sup>54 61</sup>), and questionnaire scores, such as the 36-Item Short Form Survey.<sup>62 63</sup> Blood pressure was the most commonly reported continuous clinical endpoint.<sup>52-54 64 65</sup> Patients in intervention groups had a median reduction in diastolic blood pressure of 1.0 mm Hg (interquartile range 1.0 mm Hg decrease to 0.2 mm Hg increase). Median systolic blood pressure for patients receiving an intervention, however, increased by 1.0 mm Hg (interquartile range 0.3 mm Hg decrease to 1.0 mm Hg increase).

We conducted sensitivity analyses by reanalysing the percentage of patients receiving desired processes of care using the largest improvement from each trial. Instead of each trial contributing data for all eligible processes of care, we used only the largest improvement among the outcomes reported in a given trial for these sensitivity analyses. These “best case

analyses” produced a pooled absolute improvement in patients receiving desired care of 8.5% (95% confidence interval 6.8% to 10.2%). We also repeated our analyses using odds ratios rather than risk differences, since absolute risk differences can produce less stable meta-analytic estimates than relative risks or odds ratios across varying levels of background risk.<sup>66</sup> The pooled odds ratio for improvement in adherence was 1.43 (95% confidence interval 1.30 to 1.58). Overall, when odds ratios were used, our results remained qualitatively the same for important study and clinical decision support system features, types of outcome, and level of heterogeneity.

## Discussion

### Principal findings

Across 122 controlled, mostly randomised, trials involving 1 203 053 patients and 10 790 providers, clinical decision support systems improved the

Table 4 | Absolute incremental improvements in desired care by study feature

| Study feature                               | Category (No of trials)              | Absolute incremental improvement (%; 95% CI)* | P value |
|---|--------------------------------------|---|---------|
| Clinical specialty                          | Paediatrics (9) v other (113)        | 14.7 (8.4 to 21.0)                            | <0.001  |
| Patient sample size†                        | Small (<median) v large (≥median)    | 3.7 (0.2 to 7.3)                              | 0.04    |
| Country                                     | US (82) v other (40)                 | 3.7 (0 to 7.4)                                | 0.05    |
| Baseline adherence‡                         | <Median v ≥median                    | 3.3 (-0.2 to 6.9)                             | 0.06    |
| Co-interventions beyond clinician education | Yes (35) v no (87)                   | 1.3 (-2.6 to 5.1)                             | 0.53    |
| Provider sample size§                       | Small (<median) v large (≥median)    | 0.2 (-5.7 to 6.2)                             | 0.94    |
| Study design                                | Quasi-RCT (16) v RCT (106)           | 0.1 (-5.4 to 5.6)                             | 0.96    |
| Educational co-interventions                | Yes (57) v no (65)                   | -0.8 (-4.4 to 2.8)                            | 0.66    |
| Publication year                            | 2009 onwards (87) v before 2009 (35) | -0.8 (-4.8 to 3.2)                            | 0.70    |
| Care setting                                | Outpatient (90) v other (32)         | -3.7 (-7.8 to 0.5)                            | 0.08    |

RCT=randomised controlled trial.

\*The third column shows the estimated pooled difference (and 95% confidence interval) in the percentage of desired care between the two study types listed in the first column. Results from 122 trials (108 studies) reporting dichotomous outcomes were each estimated using a univariate meta-regression model.

†Defined as small or large relative to the median patient sample size. Across all trials, the median was 2237 (interquartile range 450 to 8574).

‡The proportion of patients in control groups who received the desired process of care. Across all trials, the median was 39.4% (interquartile range 17.3% to 62.3%).

§Defined as small and large relative to the median provider sample size. Across all trials, the median was 81 (interquartile range 36 to 171).

**Table 5 | Multivariable meta-regression model for the absolute incremental improvements in desired care by study and CDSS features**

| Study and CDSS features* | Category (No of trials)           | Absolute incremental improvement (%; 95% CI)† | P value |
|--------------------------|-----------------------------------|---|---------|
| Clinical specialty       | Paediatrics (9) v other (113)     | 13.6 (7.4 to 19.8)                            | <0.001  |
| Baseline adherence‡      | <Median v ≥median                 | 3.2 (0 to 6.4)                                | 0.05    |
| Patient sample size§     | Small (<median) v large (≥median) | 2.4 (-0.8 to 5.7)                             | 0.14    |

CDSS=clinical decision support system.

\*The covariates listed in the table include all those that were included in the final multivariable meta-regression model, fitted through a backwards stepwise procedure that initially included all study and CDSS feature covariates found to have  $P < 0.1$  in the univariate meta-regression models.

†Incremental improvements refer to the additional increases in the percentages of patients receiving the desired care beyond those reported in interventions without this feature.

‡The proportion of patients in control groups who received the desired process of care. Across all trials, the median was 39.4% (interquartile range 17.3% to 62.3%).

§Defined as small and large relative to the median patient sample size. Across all trials, the median was 2237 (interquartile range 450 to 8574).

average percentage of patients receiving the desired element of care by 5.8% (95% confidence interval 4.0% to 7.6%). As expected, these trials exhibited substantial heterogeneity ( $I^2=76\%$ ). Although it is generally not advisable to perform a meta-analysis with this degree of heterogeneity, we have reported these results to show that the current literature, despite its substantial size, provides little guidance for identifying the circumstances under which clinical decision support system interventions produce worthwhile improvements in care.

### Implications

On the one hand, the extreme heterogeneity indicates non-random variation in effect sizes, such that a minority of interventions might have achieved significantly larger effects than the 95% confidence intervals around the meta-analytic average. Indeed, 25% of studies reported absolute improvements greater than 10%, with one as high as 62%.<sup>67</sup> Yet, even with the identification of two significant predictors of larger effects—namely, paediatric studies and those with low baseline adherence—the meta-regression model still showed extreme heterogeneity. Thus even when these characteristics were taken into account, a wide, non-random variation remained in the improvements seen with clinical decision support systems. The reason for this remains largely unknown.

Other systematic reviews have similarly reported extreme heterogeneity, with  $I^2$  as high as 97% in one instance.<sup>18</sup> Moja and colleagues reported more moderate heterogeneity ( $I^2=41-64\%$ ),<sup>21</sup> but their review focused specifically on measures of morbidity and mortality. Other reviews alluded to high heterogeneity without reporting formal analyses,<sup>19</sup> whereas others did not mention it.<sup>14 20</sup>

Previous reviews of clinical decision support systems have typically looked only at predictors of improvements in care,<sup>14 18</sup> or reported odds or risk ratios<sup>19-21</sup> for receiving recommended care. We sought to characterise the typical improvement achieved—namely, a 5.8% increase in the percentage of patients receiving the desired process of care. To put the magnitude of this improvement into perspective, for the control groups in the included trials, a median of 39.4% of patients received care recommended by the clinical decision support system. Thus, in the typical intervention group, about 45% would receive

this recommended process of care. Even if the meta-analytic result of 8.5% from our sensitivity analysis incorporating the largest improvement reported in each trial were used, the typical clinical decision support system intervention would still leave over 50% of patients not receiving recommended care.

We would characterise these absolute increases of 5.8% to 8.5% in the percentage of patients receiving recommended care as a small to moderate effect, but this does not imply that all such improvements are unimportant. For some processes of care, such as vaccinations and evidence based cancer screening, even a small increase in the percentage of patients who receive this care will translate into worthwhile benefits at the population level. For many other processes of care, where the recommendation has a weaker connection with important outcomes, a small increase in adherence may not justify its implementation and subsequent contribution to clinicians' frustrations with EHRs.<sup>68</sup>

Clinical decision support systems embody one of the eagerly awaited applications of artificial intelligence and machine learning for patient care.<sup>69</sup> By leveraging the power of "big data," these technologies promise earlier recognition of sepsis<sup>70 71</sup> and impending clinical deterioration.<sup>72 73</sup> Even with the most effective artificial intelligence algorithms,<sup>74</sup> however, the decision support tools alerting clinicians to their complex outputs will still largely depend on—and therefore be limited by—the small to moderate effect sizes typically obtained in our analysis.

Notably, studies that targeted a paediatric population were associated with the largest absolute improvement in process adherence. Many of these studies occurred at large health centres, affiliated to universities, with mature clinical information systems. In our earlier review,<sup>17</sup> we noted that studies conducted at institutions with longstanding experience in clinical informatics showed significantly larger improvements. In this updated review, we performed similar analyses and found no association between effect size and specific institutions or mature homegrown systems.

Although much of the early work on EHRs and clinical decision support systems took place on inpatients, 93 (72%) of the 129 trials included in this analysis took place in the outpatient setting. Across both settings, most trials (66%; 85/129) came from the US. Nonetheless, we believe our findings are relevant

elsewhere. For instance, primary care EHRs in the UK increasingly record more and more coded demographic, diagnostic, and therapeutic data, providing a basis for future clinical decision support systems.

Several factors could explain the disappointingly small effect sizes typically achieved by clinical decision support systems. Firstly, researchers and leaders in clinical informatics and human factors have recognised for years the importance of informing this work with a rich sociotechnical model. This model includes the human-computer interface, hardware and software computing infrastructure, clinical content, people, workflow and communication, internal organisational culture, external regulations, and system measurement.<sup>75-81</sup> Yet, those who develop and implement clinical decision support systems typically do not take into account the full range of these dimensions, or the complex interplay between them. In addition to using such considerations to inform the design and evaluation of clinical decision support systems, future studies should report specific potential effect modifiers, including physician EHR training, number of alerts for each visit, physician visit volume, and broader contextual factors, such as validated measures of burnout and local safety culture. A recent publication guideline may foster such improved reporting of trials of clinical decision support systems.<sup>82</sup>

Alert fatigue is a second possible explanation for the small to moderate effect sizes typically achieved in our analysis. Clinicians could encounter the same alert for many patients, or a given clinical information system could have several clinical decision support systems operating concurrently. In either case, clinicians could become less receptive to alerts from these support systems. Surprisingly, only 19 (15%) of 129 trials in this review mentioned alert fatigue in their design. Although only a minority of trials considered alert fatigue, studies focusing on this topic outside of controlled trials have highlighted the problem, and strategies to minimise the burden to users from clinical decision support systems are under investigation. Growing concerns that alert fatigue contributes to dissatisfaction with EHRs, and even clinician burnout, underscore the importance of considering this problem in future studies of support systems.<sup>68</sup>

Finally, as has happened with other common improvement interventions, from bundles and checklists to performance report cards and financial incentives, clinical decision support systems have become a “go to” off-the-shelf intervention. Many reflexively reach for these interventions without considering the way they work or the degree to which they would be likely to help with the cause of the target problem.<sup>83 84</sup> The decision to employ a clinical decision support system often reflects ease of deployment rather than its usefulness in dealing with the problem. A pop-up computer screen reminding clinicians of the approved indications for a broad spectrum antibiotic may ignore the psychological reality of the clinician’s primary concern at that moment—avoiding further

deterioration of a very sick patient, rather than the public health consequences of excessive antibiotic use. An antimicrobial stewardship programme<sup>85</sup> could hold far greater promise in achieving this goal than an interruptive alert, which many clinicians will probably ignore, or worse, could result in alert fatigue, thereby undermining the potential effectiveness of other support system interventions in a given healthcare setting. Future interventions may also seek to harness potential synergies between clinical decision support systems and other well known interventions, such as performance feedback.<sup>86</sup>

### Risk of bias

The only included studies with high risk of bias were 18 (16%) of 113 clustered trials exhibiting unit of analysis errors, which typically underestimate the width of confidence intervals. We were able to include such studies without replicating this bias because we used their primary data, rather than the reported effect sizes, together with reported or imputed intraclass correlation coefficients for each clustered trial in a multilevel model. The lack of other trials with a high risk of bias among the included studies probably reflects the nature of the intervention (that is, clinical decision support systems delivered at the point-of-care) and the endpoints in our analysis (that is, the degree to which patients received recommended care as documented in the EHR), as many of the methodological shortcomings that can undermine evaluations of other interventions are easier to avoid. For instance, systematic differences in outcome ascertainment can hardly arise. There is also no clear way in which exposures other than the intervention of interest (the clinical decision support system) could systematically differ between intervention and control patients. Losses to follow-up (that is, attrition bias) are also virtually impossible given that the EHR captures whether or not the clinician’s orders or documentation complied with recommended care in response to the clinical decision support system triggered.

Attrition due to loss of entire clusters is the one possible exception and occurred in several trials in which one or more clinics assigned to the control arm dropped out of the trial. In a representative example of such a trial,<sup>87</sup> participating clinics received an EHR system with or without additional guideline based decision support. Some clinics assigned to the control group dropped out owing to lack of motivation to implement a new system with no chance of benefitting their patients. The decision of clinics in the control arm not to follow through with an information technology intervention, which is well known to take personnel time and cause frustration to clinicians, did not seem to us to carry a clear risk of bias. We labelled the few such trials as “unclear risk.”

### Limitations

Heterogeneity among the included studies is the main limitation to our analysis. Recommendations to refrain from meta-analysis when the  $I^2$  statistic



exceeds 50%<sup>22</sup> stem from the desire to avoid spurious precision in the estimated effect size. For instance, the extreme heterogeneity we report—again, indicating substantial variation in effects across trials beyond those expected from chance alone—suggests that some subsets of trials might have achieved substantially larger (or smaller) effects than denoted by the 95% confidence interval for the pooled effect across all trials. Indeed, this is one of the central findings in our study. Yet, even with a meta-regression model incorporating objective features of either the clinical decision support systems or the studies evaluating them, heterogeneity remained unchanged. Thus we conducted this meta-analysis despite extreme levels of heterogeneity because the results highlight that, while clinical decision support systems sometimes achieve large effects, current literature does not adequately identify the circumstances under which worthwhile improvements occur.

### Conclusions

Despite publication of over 100 randomised and quasi-randomised trials involving over one million patients and 10000 providers, the observed improvements display extreme, unexplained variation. Achieving worthwhile improvements therefore remains largely a case of trial and error. Future research must identify new ways of designing clinical decision support systems that reliably confer larger improvements in care while balancing the threat of alert fatigue. Head-to-head trials comparing design features of different support systems will also be important.<sup>88</sup> In the meantime, a critical consideration in deciding to implement clinical decision support systems is the strength of the connection between the targeted process of care and patient outcomes. Achieving small improvements in care with only a presumptive connection to patient outcomes may not be worth the risk of contributing to alert fatigue or the growing concern of physician burnout attributable to EHRs.<sup>68 89</sup>

We thank Julia Worswick, Sharlini Yogasingam, and Claire Chow for their support with data abstraction; Alex Kiss for statistical support; and Michelle Fiander for assistance with the bibliographic search.

**Contributors:** JLK, LL, JMG, and KGS led the study design. JLK, LL, JF, HG, and KGS extracted the data. JLK, JPD, GT, and KGS analysed the data. JLK and KGS drafted the manuscript. All authors provided critical revision of the manuscript for important intellectual content and approval of the final submitted version. JLK and KGS are the guarantors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

**Funding:** JMG holds a Canada Research Chair in health knowledge transfer and uptake. The funders had no role in considering the study design or in the collection, analysis, interpretation of data, writing of the report, or decision to submit the article for publication.

**Competing interests:** All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/doi\\_disclosure.pdf](http://www.icmje.org/doi_disclosure.pdf) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

**Ethical approval:** Ethical approval for this evidence synthesis was not required.

**Data sharing:** No additional data available.

The lead author (the manuscript's guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

**Dissemination to participants and related patient and public communities:** Dissemination of the study results to study participants is not applicable. We plan to use media outreach (eg, press release) and social media to disseminate our findings and communicate with the population at large.

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

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 and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

- Gillum RF. From papyrus to the electronic tablet: a brief history of the clinical medical record with lessons for the digital age. *Am J Med* 2013;126:853-7. doi:10.1016/j.amjmed.2013.03.024.
- Jha AK, DesRoches CM, Campbell EG, et al. Use of electronic health records in US hospitals. *N Engl J Med* 2009;360:1628-38. doi:10.1056/NEJMs0900592.
- Adler-Milstein J, DesRoches CM, Kralovec P, et al. Electronic health record adoption in US hospitals: progress continues, but challenges persist. *Health Aff (Millwood)* 2015;34:2174-80. doi:10.1377/hlthaff.2015.0992.
- Adler-Milstein J, Jha AK. HITECH Act drove large gains in hospital electronic health record adoption. *Health Aff (Millwood)* 2017;36:1416-22. doi:10.1377/hlthaff.2016.1651.
- Sheikh A, Jha A, Cresswell K, Greaves F, Bates DW. Adoption of electronic health records in UK hospitals: lessons from the USA. *Lancet* 2014;384:8-9. doi:10.1016/S0140-6736(14)61099-0.
- NHS. NHS long term plan. 2019. <https://www.longtermplan.nhs.uk>.
- Loo TS, Davis RB, Lipsitz LA, et al. Electronic medical record reminders and panel management to improve primary care of elderly patients. *Arch Intern Med* 2011;171:1552-8. doi:10.1001/archinternmed.2011.394.
- Tamblyn R, Huang A, Perreault R, et al. The medical office of the 21st century (MOXXI): effectiveness of computerized decision-making support in reducing inappropriate prescribing in primary care. *CMAJ* 2003;169:549-56.
- McGinn TG, McCullagh L, Kannry J, et al. Efficacy of an evidence-based clinical decision support in primary care practices: a randomized clinical trial. *JAMA Intern Med* 2013;173:1584-91. doi:10.1001/jamainternmed.2013.8980.
- Chertow GM, Lee J, Kuperman GJ, et al. Guided medication dosing for inpatients with renal insufficiency. *JAMA* 2001;286:2839-44. doi:10.1001/jama.286.22.2839.
- 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;13:385-90. doi:10.1197/jamia.M1945.
- Linder JA, Rigotti NA, Schneider LI, Kelley JH, Brawarsky P, Haas JS, et al. An electronic health record-based intervention to improve tobacco treatment in primary care: a cluster-randomized controlled trial. *Arch Intern Med* 2009;169:781-7. doi:10.1001/archinternmed.2009.53.
- Wright A, Pang J, Febloowitz JC. Improving completeness of electronic problem lists through clinical decision support: a randomized, controlled trial. *J Am Med Inform Assoc* 2012;19:555-61. doi:10.1136/amiajn-2011-000521.
- Garg AX, Adhikari NK, McDonald H, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA* 2005;293:1223-38. doi:10.1001/jama.293.10.1223.
- 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. doi:10.1136/bmj.38398.500764.8F.
- Hunt DL, Haynes RB, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. *JAMA* 1998;280:1339-46. doi:10.1001/jama.280.15.1339.
- Shojania KG, Jennings A, Mayhew A, Ramsay C, Eccles M, Grimshaw J. Effect of point-of-care computer reminders on physician behaviour: a systematic review. *CMAJ* 2010;182:E216-25. doi:10.1503/cmaj.090578.
- Holt TA, Dalton A, Marshall T, et al. Automated software system to promote anticoagulation and reduce stroke risk: cluster-randomized controlled trial. *Stroke* 2017;48:787-90. doi:10.1161/STROKEAHA.116.015468.

- 19 Bright TJ, Wong A, Dhurjati R, et al. Effect of clinical decision-support systems: a systematic review. *Ann Intern Med* 2012;157:29-43. doi:10.7326/0003-4819-157-1-201207030-00450.
- 20 Roshanov PS, Fernandes N, Wilczynski JM, et al. Features of effective computerised clinical decision support systems: meta-regression of 162 randomised trials. *BMJ* 2013;346:f657. doi:10.1136/bmj.f657.
- 21 Moja L, Kwag KH, Lytras T, et al. Effectiveness of computerized decision support systems linked to electronic health records: a systematic review and meta-analysis. *Am J Public Health* 2014;104:e12-22. doi:10.2105/AJPH.2014.302164.
- 22 Higgins JPT, Thomas J, Chandler J, et al. *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0 (updated July 2019). <https://www.training.cochrane.org/handbook>.
- 23 Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535. doi:10.1136/bmj.b2535.
- 24 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;CD001096. doi:10.1002/14651858.CD001096.pub2.
- 25 Gonzales R, Anderer T, McCulloch CE, et al. A cluster randomized trial of decision support strategies for reducing antibiotic use in acute bronchitis. *JAMA Intern Med* 2013;173:267-73. doi:10.1001/jamainternmed.2013.1589.
- 26 Higgins JPT, Savović J, Page MJ, et al. Assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, et al, eds. *Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)*. Cochrane, 2019:Chapter 8. doi:10.1002/9781119536604.ch8.
- 27 Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898. doi:10.1136/bmj.l4898.
- 28 University of Aberdeen Health Services Research Unit Research Tools. Database of intra-correlation coefficients (ICCs). <https://www.abdn.ac.uk/hsru/what-we-do/tools/index.php> accessed March 21, 2019.
- 29 McKenzie JE, Brennan SE. Synthesizing and presenting findings using other methods. In: Higgins JPT, Thomas J, Chandler J, et al, eds. *Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)*. Cochrane, 2019:Chapter 12. doi:10.1002/9781119536604.ch12.
- 30 Grimshaw JM, Thomas RE, MacLennan G, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess* 2004;8:iii-iv, 1-72. doi:10.3310/hta8060.
- 31 Walsh JM, McDonald KM, Shojania KG, et al. Quality improvement strategies for hypertension management: a systematic review. *Med Care* 2006;44:646-57. doi:10.1097/01.mlr.0000220260.30768.32.
- 32 Ranji SR, Steinman MA, Shojania KG, Gonzales R. Interventions to reduce unnecessary antibiotic prescribing: a systematic review and quantitative analysis. *Med Care* 2008;46:847-62. doi:10.1097/MLR.0b013e318178eabd.
- 33 Burnham K, Anderson DR. *Model selection and multimodel inference: a practical information-theoretic approach*. 2nd ed. Springer, 2002.
- 34 Coté GA, Rice JP, Bulsiewicz W, et al. Use of physician education and computer alert to improve targeted use of gastroprotection among NSAID users. *Am J Gastroenterol* 2008;103:1097-103. doi:10.1111/j.1572-0241.2008.01907.x.
- 35 Eccles M, McColl E, Steen N, et al. Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care: a cluster randomised controlled trial. *BMJ* 2002;325:941. doi:10.1136/bmj.325.7370.941.
- 36 Flottorp S, Oxman AD, Håvelsrud K, Treweek S, Herrin J. Cluster randomised controlled trial of tailored interventions to improve the management of urinary tract infections in women and sore throat. *BMJ* 2002;325:367. doi:10.1136/bmj.325.7360.367.
- 37 Myers JS, Gojraty S, Yang W, Linsky A, Airan-Javia S, Polomano RC. A randomized-controlled trial of computerized alerts to reduce unapproved medication abbreviation use. *J Am Med Inform Assoc* 2011;18:17-23. doi:10.1136/jamia.2010.006130.
- 38 Sequist TD, Holliday AM, Orav EJ, Bates DW, Denker BM. Physician and patient tools to improve chronic kidney disease care. *Am J Manag Care* 2018;24:e107-14.
- 39 Tamblyn R, Winslade N, Qian CJ, Moraga T, Huang A. What is in your wallet? A cluster randomized trial of the effects of showing comparative patient out-of-pocket costs on primary care prescribing for uncomplicated hypertension. *Implement Sci* 2018;13:7. doi:10.1186/s13012-017-0701-x.
- 40 Taveras EM, Marshall R, Kleinman KP, et al. Comparative effectiveness of childhood obesity interventions in pediatric primary care: a cluster-randomized clinical trial. *JAMA Pediatr* 2015;169:535-42. doi:10.1001/jamapediatrics.2015.0182.
- 41 van Wyk JT, van Wijk MA, Sturkenboom MC, Mosseveld M, Moorman PW, van der Lei J. Electronic alerts versus on-demand decision support to improve dyslipidemia treatment: a cluster randomized controlled trial. *Circulation* 2008;117:371-8. doi:10.1161/CIRCULATIONAHA.107.697201.
- 42 Persell SD, Doctor JN, Friedberg MW, et al. Behavioral interventions to reduce inappropriate antibiotic prescribing: a randomized pilot trial. *BMC Infect Dis* 2016;16:373. doi:10.1186/s12879-016-1715-8.
- 43 Smith JR, Noble MJ, Musgrave S, et al. The at-risk registers in severe asthma (ARRISA) study: a cluster-randomised controlled trial examining effectiveness and costs in primary care. *Thorax* 2012;67:1052-60. doi:10.1136/thoraxjnl-2012-202093.
- 44 Abdel-Kader K, Fischer GS, Li J, Moore CG, Hess R, Unruh ML. Automated clinical reminders for primary care providers in the care of CKD: a small cluster-randomized controlled trial. *Am J Kidney Dis* 2011;58:894-902. doi:10.1053/j.ajkd.2011.08.028.
- 45 Gupta A, Gholami P, Turakhia MP, Friday K, Heidenreich PA. Clinical reminders to providers of patients with reduced left ventricular ejection fraction increase defibrillator referral: a randomized trial. *Circ Heart Fail* 2014;7:140-5. doi:10.1161/CIRCHEARTFAILURE.113.000753.
- 46 Gill JM, Mainous AG 3rd, Koopman RJ, et al. Impact of EHR-based clinical decision support on adherence to guidelines for patients on NSAIDs: a randomized controlled trial. *Ann Fam Med* 2011;9:22-30. doi:10.1370/afm.1172.
- 47 Safran C, Rind DM, Davis RB, et al. Guidelines for management of HIV infection with computer-based patient's record. *Lancet* 1995;346:341-6. doi:10.1016/S0140-6736(95)92226-1.
- 48 Arts DL, Abu-Hanna A, Medlock SK, van Weert HC. Effectiveness and usage of a decision support system to improve stroke prevention in general practice: a cluster randomized controlled trial. *PLoS One* 2017;12:e0170974. doi:10.1371/journal.pone.0170974.
- 49 Schriever SP, Landis SE, Turbow DJ, Patch SC. Effect of a computerized body mass index prompt on diagnosis and treatment of adult obesity. *Fam Med* 2009;41:502-7.
- 50 Downs M, Turner S, Bryans M, et al. Effectiveness of educational interventions in improving detection and management of dementia in primary care: cluster randomised controlled study. *BMJ* 2006;332:692-6. doi:10.1136/bmj.332.7543.692.
- 51 Tang JW, Kushner RF, Cameron KA, Hicks B, Cooper AJ, Baker DW. Electronic tools to assist with identification and counseling for overweight patients: a randomized controlled trial. *J Gen Intern Med* 2012;27:933-9. doi:10.1007/s11606-012-2022-8.
- 52 Hicks LS, Sequist TD, Ayanian JZ, et al. Impact of computerized decision support on blood pressure management and control: a randomized controlled trial. *J Gen Intern Med* 2008;23:429-41. doi:10.1007/s11606-007-0403-1.
- 53 Dregan A, van Staa TP, McDermott L, et al. Trial Steering Committee, Data Monitoring Committee. Point-of-care cluster randomized trial in stroke secondary prevention using electronic health records. *Stroke* 2014;45:2066-71. doi:10.1161/STROKEAHA.114.005713.
- 54 Meigs JB, Cagliero E, Dubey A, et al. A controlled trial of web-based diabetes disease management: the MGH diabetes primary care improvement project. *Diabetes Care* 2003;26:750-7. doi:10.2337/diacare.26.3.750.
- 55 Beeler PE, Eschmann E, Schumacher A, Studt JD, Amann-Vesti B, Blaser J. Impact of electronic reminders on venous thromboprophylaxis after admissions and transfers. *J Am Med Inform Assoc* 2014;21(e2):e297-303. doi:10.1136/amiajnl-2013-002225.
- 56 Karlsson LO, Nilsson S, Bång M, Nilsson L, Charitakis E, Janzon M. A clinical decision support tool for improving adherence to guidelines on anticoagulant therapy in patients with atrial fibrillation at risk of stroke: a cluster-randomized trial in a Swedish primary care setting (the CDS-AF study). *PLoS Med* 2018;15:e1002528. doi:10.1371/journal.pmed.1002528.
- 57 Kucher N, Koo S, Quiroz R, et al. Electronic alerts to prevent venous thromboembolism among hospitalized patients. *N Engl J Med* 2005;352:969-77. doi:10.1056/NEJMoa041533.
- 58 Spirk D, Stuck AK, Hager A, Engelberger RP, Aujesky D, Kucher N. Electronic alert system for improving appropriate thromboprophylaxis in hospitalized medical patients: a randomized controlled trial. *J Thromb Haemost* 2017;15:2138-46. doi:10.1111/jth.13812.
- 59 Boustani MA, Campbell NL, Khan BA, et al. Enhancing care for hospitalized older adults with cognitive impairment: a randomized controlled trial. *J Gen Intern Med* 2012;27:561-7. doi:10.1007/s11606-012-1994-8.
- 60 Dean NC, Jones BE, Jones JP, et al. Impact of an electronic clinical decision support tool for emergency department patients with pneumonia. *Ann Emerg Med* 2015;66:511-20. doi:10.1016/j.annemergmed.2015.02.003.
- 61 Mann DM, Palmisano J, Lin JJ. A pilot randomized trial of technology-assisted goal setting to improve physical activity among primary care patients with prediabetes. *Prev Med Rep* 2016;4:107-12. doi:10.1016/j.pmedr.2016.05.012.
- 62 Tierney WM, Overhage JM, Murray MD, et al. Effects of computerized guidelines for managing heart disease in primary care. *J Gen Intern Med* 2003;18:967-76. doi:10.1111/j.1525-1497.2003.30635.x.

- 63 Tierney WM, Overhage JM, Murray MD, et al. Can computer-generated evidence-based care suggestions enhance evidence-based management of asthma and chronic obstructive pulmonary disease? A randomized, controlled trial. *Health Serv Res* 2005;40:477-97. doi:10.1111/j.1475-6773.2005.0t369.x.
- 64 Murray MD, Harris LE, Overhage JM, et al. Failure of computerized treatment suggestions to improve health outcomes of outpatients with uncomplicated hypertension: results of a randomized controlled trial. *Pharmacotherapy* 2004;24:324-37. doi:10.1592/phco.24.4.324.33173
- 65 Bennett GG, Steinberg D, Askew S, et al. Effectiveness of an app and provider counseling for obesity treatment in primary care. *Am J Prev Med* 2018;55:777-86. doi:10.1016/j.amepre.2018.07.005.
- 66 Furukawa TA, Guyatt GH, Griffith LE. Can we individualize the 'number needed to treat'? An empirical study of summary effect measures in meta-analyses. *Int J Epidemiol* 2002;31:72-6. doi:10.1093/ije/31.1.72.
- 67 Diaz MCG, Wysocki T, Crutchfield JHJR, Franciosi JP, Werk LN. Provider-focused intervention to promote comprehensive screening for adolescent idiopathic scoliosis by primary care pediatricians. *Am J Med Qual* 2019;34:182-8. doi:10.1177/1062860618792667.
- 68 Downing NL, Bates DW, Longhurst CA. Physician burnout in the electronic health record era: are we ignoring the real cause? *Ann Intern Med* 2018;169:50-1. doi:10.7326/M18-0139.
- 69 Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. *JAMA* 2018;320:2199-200. doi:10.1001/jama.2018.17163.
- 70 Ruppel H, Liu V. To catch a killer: electronic sepsis alert tools reaching a fever pitch? *BMJ Qual Saf* 2019;28:693-6. doi:10.1136/bmjqs-2019-009463.
- 71 Desautels T, Calvert J, Hoffman J, et al. Prediction of sepsis in the intensive care unit with minimal electronic health record data: a machine learning approach. *JMIR Med Inform* 2016;4:e28. doi:10.2196/medinform.5909.
- 72 Churpek MM, Yuen TC, Winslow C, Meltzer DO, Kattan MW, Edelson DP. Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. *Crit Care Med* 2016;44:368-74. doi:10.1097/CCM.0000000000001571.
- 73 Ye C, Wang O, Liu M, et al. A real-time early warning system for monitoring inpatient mortality risk: prospective study using electronic medical record data. *J Med Internet Res* 2019;21:e13719. doi:10.2196/13719.
- 74 Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med* 2018;24:1716-20. doi:10.1038/s41591-018-0213-5.
- 75 Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. *Qual Saf Health Care* 2010;19(Suppl 3):i68-74. doi:10.1136/qshc.2010.042085.
- 76 Ancker JS, Miller MC, Patel V, Kaushal R, HITEC Investigators. Sociotechnical challenges to developing technologies for patient access to health information exchange data. *J Am Med Inform Assoc* 2014;21:664-70. doi:10.1136/amiajnl-2013-002073.
- 77 Carayon P, Wetterneck TB, Cartmill R, et al. Medication safety in two intensive care units of a community teaching hospital after electronic health record implementation: sociotechnical and human factors engineering considerations. *J Patient Saf* 2017. doi:10.1097/PTS.0000000000000358.
- 78 Castro GM, Buczkowski L, Hafner JM. The contribution of sociotechnical factors to health information technology-related sentinel events. *Jt Comm J Qual Patient Saf* 2016;42:70-6. doi:10.1016/S1553-7250(16)42008-8.
- 79 Furniss D, Garfield S, Husson F, Blandford A, Franklin BD. Distributed cognition: understanding complex sociotechnical informatics. *Stud Health Technol Inform* 2019;263:75-86. doi:10.3233/shti190113.
- 80 Meeks DW, Takian A, Sittig DF, Singh H, Barber N. Exploring the sociotechnical intersection of patient safety and electronic health record implementation. *J Am Med Inform Assoc* 2014;21(e1):e28-34. doi:10.1136/amiajnl-2013-001762.
- 81 Odukoya OK, Chui MA. e-Prescribing: characterisation of patient safety hazards in community pharmacies using a sociotechnical systems approach. *BMJ Qual Saf* 2013;22:816-25. doi:10.1136/bmjqs-2013-001834.
- 82 Singh H, Sittig DF. A sociotechnical framework for safety-related electronic health record research reporting: the SAFER reporting framework. *Ann Intern Med* 2020;172(11\_Supplement):S92-100. doi:10.7326/M19-0879.
- 83 Davidoff F, Dixon-Woods M, Leviton L, Michie S. Demystifying theory and its use in improvement. *BMJ Qual Saf* 2015;24:228-38. doi:10.1136/bmjqs-2014-003627.
- 84 Foy R, Ovretveit J, Shekelle PG, et al. The role of theory in research to develop and evaluate the implementation of patient safety practices. *BMJ Qual Saf* 2011;20:453-9. doi:10.1136/bmjqs.2010.047993.
- 85 Peragine C, Walker SAN, Simor A, Walker SE, Kiss A, Leis JA. Impact of a comprehensive antimicrobial stewardship program on institutional burden of antimicrobial resistance: a 14-year controlled interrupted time series study. *Clin Infect Dis* 2019;ci21183. doi:10.1093/cid/ciz1183.
- 86 Gulliford MC, Prevost AT, Charlton J, et al. Effectiveness and safety of electronically delivered prescribing feedback and decision support on antibiotic use for respiratory illness in primary care: REDUCE cluster randomised trial. *BMJ* 2019;364:l236. doi:10.1136/bmj.l236.
- 87 Goud R, de Keizer NF, ter Riet G, et al. Effect of guideline based computerised decision support on decision making of multidisciplinary teams: cluster randomised trial in cardiac rehabilitation. *BMJ* 2009;338:b1440. doi:10.1136/bmj.b1440.
- 88 Van de Velde S, Heselmans A, Delvaux N, et al. A systematic review of trials evaluating success factors of interventions with computerised clinical decision support. *Implement Sci* 2018;13:114. doi:10.1186/s13012-018-0790-1.
- 89 Gregory ME, Russo E, Singh H. Electronic health record alert-related workload as a predictor of burnout in primary care providers. *Appl Clin Inform* 2017;8:686-97. doi:10.4338/ACI-2017-01-RA-0003.

### Web appendix: Supplementary appendix