PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Impact of Electronic Health Records on Pre-defined Safety Outcomes in Patients Admitted to Hospital. A Scoping Review
AUTHORS	Subbe, Christian; Tellier, Genevieve; Barach, paul

VERSION 1 – REVIEW

REVIEWER	Justin Keen University of Leeds, England
REVIEW RETURNED	15-Jul-2020

GENERAL COMMENTS	Thank you for the opportunity to review this draft article on an important topic, the effects of electronic health records on hospital patient outcomes. The current draft raises a number of questions. The first concerns the relationship of this literature review to published systematic reviews. Many reviews of the effects of hospital information technologies have been published in the last few years, and it is not clear how this review relates to them - does it seek to provide an overview (including earlier reviews), or to fill a gap left by those reviews? It appears to be the former, but cited reviews such as that by Brenner et al - Brenner, S.K., et al., Effects of health information technology on patient outcomes: a systematic review. J Am Med Inform Assoc, 2016. 23(5): p. 1016-36. - are not included. Admittedly Brenner and colleagues' review covered a wide range of technologies, not just EHR, but their review did include hospital EHRs, and it is not clear why this or other reviews were included. It may also be worth noting here that several reviews included larger numbers of studies. Even allowing for different search parameters, the total of 22 included studies for this review seems lower than one might expect. This leads to another issue, concerning the focus of the review. The title indicates that the focus is on the relationship between EHR functions and patient outcomes, but the abstract (page 4 of file) states that the focus is on identifying gaps in the literature, and the research question (page 7) does not mention either functionality or gaps. Different sections of the article appear to focus on one or more topics - functionality, interventions, metrics, outcomes. This led to some confusion for this reader, exemplified by the status of Table 2.

	The text suggests that it is concerned with 'impact' of EHR, but the Zogers article cited focuses on interventions, not on impacts or outcomes. This leads to a broader point, that the basis for identifying gaps in the literature - if this was the principal objective - was not clear. Is the presence or absence of patient outcome evidence the right basis for identifying gaps? Or, if the focus was on EHR functionality - as implied by the title - should the article focus, instead, on functions that need further study? Following on from this point, it is not clear why the quality of included studies was not reviewed. If the evidence found, and listed in Table 2, was of poor quality, isn't this essentially the same as finding no evidence, given the small number of studies found under each heading?

REVIEWER	Philip Scott University of Portsmouth, UK
REVIEW RETURNED	23-Jul-2020

GENERAL COMMENTS	This is a clear and mostly well written paper on an important topic, following an established review type. Within its defined scope, the findings seem reasonable. My principal reservation is about the lack of discussion grounding the study in previous research, such as https://pubmed.ncbi.nlm.nih.gov/24159271/ and perhaps most notably Black et al. 2011: https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000387 There are several papers I would have expected the review to have included (even only in the discussion) but I do not see them in the references, for example: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3236066/ https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3797550/ The other main weakness is that the paper does not follow the (admittedly recent) PRISMA extension for scoping reviews: https://www.equator-network.org/reporting-guidelines/prisma-scr/ Minor comments: Strengths and limitations: second bullet text is poor grammar. First sentence of the Introduction is ungrammatical. First sentence of Study Selection Process should be "that" were reviewed (not "who"). Table 1 first column says it is author and year but actually it is only author. Page 13, line 12-13 would read better as "almost exclusively"; line 48-49 would read better as "a single" or "any" rather than "one" trial. Page 18, line 26ff - I do not understand why there is a need for such a long direct quotation rather than a summary; line 40-41 - I cannot see the term "System level two metrics" explained anywhere. Page 19, line 48 - typo on "scoping".

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Justin Keen

Thank you for the opportunity to review this draft article on an important topic, the effects of electronic health records on hospital patient outcomes. See the authors responses below in red.

Reviewer: The current draft raises a number of questions. The first concerns the relationship of this literature review to published systematic reviews. Many reviews of the effects of hospital information technologies have been published in the last few years, and it is not clear how this review relates to them - does it seek to provide an overview (including earlier reviews), or to fill a gap left by those reviews? It appears to be the former, but cited reviews such as that by Brenner et al - Brenner, S.K., et al., Effects of health information technology on patient outcomes: a systematic review. J Am Med Inform Assoc, 2016. 23(5): p. 1016-36. - are not included. Admittedly Brenner and colleagues' review covered a wide range of technologies, not just EHR, but their review did include hospital EHRs, and it is not clear why this or other reviews were included.

Reply: Many thanks: we have to apologise as we did indeed miss Brenner's review. Of the 69 papers included in this review the majority proceed our inclusion period of 2008 to 2018; only four addressed Electronic Health Records, and two of these fulfilled inclusion criteria for our review (Cook et al., 2011; Dowding, Turley and Garrido, 2012). These papers have now been added to our manuscript.

We apologize but it is our understanding that scoping reviews can be used for four reasons: To (1) map fields of study, (2) determine the value of a full systematic review, (3) summarise research findings or (4) identify research gaps in the existing literature (O'Malley and Arksey, 2005). Our focus was on mapping the existing field of study and identify the key gaps. We do believe that the heterogeneity of the provided evidence would at current preclude a full systematic review.

Reviewer: It may also be worth noting here that several reviews included larger numbers of studies. Even allowing for different search parameters, the total of 22 included studies for this review seems lower than one might expect.

Reply: We agree that the number of reviewed studies is much lower than we expected. In our search we encountered a large number of observational or feasibility studies that described concepts or implementation efforts but we struggled to identify interventional studies that actually measured our pre-defined selection of safety outcomes in contemporaneous studies. A recently published review of Personal Health Records encountered very similar problems (Kelly, Coller and Hoonakker, 2017).

Reviewer: This leads to another issue, concerning the focus of the review. The title indicates that the focus is on the relationship between EHR functions and patient outcomes, but the abstract (page 4 of file) states that the focus is on identifying gaps in the literature, and the research question (page 7) does not mention either functionality or gaps.

Reply: Many thanks for the comment. We used the term 'functionality' to describe the process by which the EHR might effect changes in patient safety outcomes. In response to the comment by the reviewer we have simplified the title of the paper: "The impact of Electronic Health Records on pre-defined safety outcomes in patients admitted to hospital". We have additionally revised the abstract.

It is our understanding that a summary of the evidence and identification of gaps is inherent to the

methodology of a scoping review. The identified gaps in the evidence for impact of EHRs on the pre-defined safety outcomes are summarised in table 2 and the subheadings of the paper have been revised accordingly.

Reviewer: Different sections of the article appear to focus on one or more topics - functionality, interventions, metrics, outcomes. This led to some confusion for this reader, exemplified by the status of Table 2. The text suggests that it is concerned with 'impact' of EHR, but the Zegers article cited focuses on interventions, not on impacts or outcomes.
This leads to a broader point, that the basis for identifying gaps in the literature - if this was the principal objective - was not clear. Is the presence or absence of patient outcome evidence the right basis for identifying gaps? Or, if the focus was on EHR functionality - as implied by the title - should the article focus, instead, on functions that need further study?

Reply: We apologise to the reviewer if the listed terms were not clear in our manuscript. Impact is in our understanding defined as ‘effect or influence’. Zegers et al state as the objective of their study to search for “effective interventions aimed at reducing rates of adverse events in hospitals” with adverse events being the relevant outcome. They extracted data on “study population, study design, intervention characteristics and adverse patient outcomes”. Their protocol specifies that “interventions should contain 1 or more components (described in the article) that aimed to reduce adverse patient outcomes”, and, the outcomes were specified in the search strategy in the appendices of the manuscript and are the same outcomes that we used.

In an analogous way, we searched for interventions (as evidenced by functionality of the EHR) that had impact on safety outcomes. This leads to a somewhat broader discussion: the focus of any interventional study might be the relationship between an intervention as the cause and outcome as the effect, with the effect achieved through faithful application of the intervention and usually evidenced by process measures. There is however evidence from your and other groups (Dixon-Woods et al., 2011, 2013; Randell et al., 2019) that the effects of interventions might not be mediated purely through the functionality of the intervention but perhaps by the cultural context and trust: i.e., a shift in organisational culture might be a side effect of an intervention and conversely where organisations were willing to invest into EHRs and the efforts in implementing them this might improve patient mortality and other patient outcomes in these organisations. For the purpose of this paper we examined the evidence for processes of the EHR that might explain their effects on patient safety outcomes.

Following the critique of the reviewer we have removed the word ‘functionality’ from the title. We have revised the subheadings of the result section to read: ‘Processes by which EHRs aimed to effect changes in Safety Outcomes’, ‘Impact on Safety Outcomes’, and ‘ Additional gaps in understanding of impact of EHRs on safety outcomes ‘. We have reorganised all content in line with the new subheadings.

Reviewer: Following on from this point, it is not clear why the quality of included studies was not reviewed. If the evidence found, and listed in Table 2, was of poor quality, isn't this essentially the same as finding no evidence, given the small number of studies found under each heading?

Reply: We apologise but it is our understanding that a review of the quality of studies is not within the usual methodology of scoping reviews (O’Malley and Arksey, 2005; Peters et al., 2015).

Reviewer: 2
Reviewer Name: Philip Scott

Reviewer: This is a clear and mostly well written paper on an important topic, following an established

review type. Within its defined scope, the findings seem reasonable. My principal reservation is about the lack of discussion grounding the study in previous research, such as https://pubmed.ncbi.nlm.nih.gov/24159271/ and perhaps most notably Black et al. 2011: https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000387

Reply: Many thanks for your insightful comments and calling our attention to other papers. We have added these paper's conclusions and references to the discussion in our manuscript. The argument that Bowman makes about the challenges of implementing EHRs safely are highly relevant to our work. Black et al, seem to come to a similar conclusion to our review that 'there is a large gap between the postulated and empirically demonstrated benefits of eHealth technologies'. Their review focused on manuscripts published from 1997 up to 2007 in contrast to our sampling period of 2008 to 2018.

Reviewer: There are several papers I would have expected the review to have included (even only in the discussion) but I do not see them in the references, for example:
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3236066/
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3797550/

Reply: Many thanks for the additional references – we have added them to the discussion. Both underline the points that we have made in the comments above – investment in EHRs might change the perceptions of patient safety but not always result in demonstrable outcomes that are measurable at the patient level.

Reviewer: The other main weakness is that the paper does not follow the (admittedly recent) PRISMA extension for scoping reviews: https://www.equator-network.org/reporting-guidelines/prisma-scr/

Reply: Thank you for highlighting our omission. We have added a completed PRIMSA inspired checklist to the uploaded documents. We have added additional information about protocol and searches. We believe that we have covered the items of the PRISMA extension in the revised manuscript. We have added to the manuscript that the study was not registered.

Minor comments:

Strengths and limitations: second bullet text is poor grammar.

Authors: Many thanks. We have corrected the text.

First sentence of the Introduction is ungrammatical.

Authors: We have changed the sentence.

First sentence of Study Selection Process should be "that" were reviewed (not "who").

Authors: We have corrected the sentence.

Table 1 first column says it is author and year but actually it is only author.

Authors: We have corrected the label.

Page 13, line 12-13 would read better as "almost exclusively"; line 48-49 would read better as "a single" or "any" rather than "one" trial.

Authors: The wording has been corrected.

Page 18, line 26ff - I do not understand why there is a need for such a long direct quotation rather than a summary; line 40-41 - I cannot see the term "System level two metrics" explained anywhere.

Authors: We have reviewed the text, summarised the quotation and corrected the sentence that you highlighted.

Page 19, line 48 - typo on "scoping".

Authors: We have been unable to find this typo. Please clarify.

VERSION 2 – REVIEW

REVIEWER	Justin Keen University of Leeds, England
REVIEW RETURNED	15-Sep-2020

GENERAL COMMENTS	The revised manuscript addresses some of the issues raised in the first round of reviewer comments. There are, though, two substantive issues that still merit attention. The first concerns the main focus of the article, which in different passages is on patient outcomes, on process measures and on gaps in the literature. Two examples may help to illustrate this point. (1) The title mentions pre-defined safety outcomes, but the conclusions in the Abstract do not state how many of the pre-defined outcomes were supported by empirical evidence (or indeed where the gaps in the literature lie). (2) The Methods section does not state how the 14 pre-defined safety outcomes were identified, or why Zegers' list was used rather than a different one (several have been published). The main source provided is Table 2, but most of the measures listed appear to be process measures. So, is the focus on processes or outcomes? The second substantive issue concerns the 'diffuse' nature of EHRs. Observational studies show that they can be, and are, used in myriad different ways in different hospitals. Linking the EHR intervention to a specific outcome will always be difficult, however ingenious the research team. Indeed, as some of the reported studies show, it can be difficult to design robust studies even where process changes are the end-points. The significance of this point lies in the authors' conclusions that we need better metrics, and deeper insights into the poor quality of EHRs that clinicians are expected to use. An alternative conclusion is that randomised trials alone cannot provide definitive evidence about the effects of EHRs. Rather, we need carefully designed suites of experimental and observational studies - the latter, in particular, to understand quite why EHRs are still so clunky, in the age of Amazon, WhatsApp and so on. That, arguably, is why some systematic reviews have sought to piece together evidence about both processes and outcomes, using Realist and Meta-Narrative methods.

	A broader point follows. Comments are made to the effect that clinicians need high quality information to treat patients - this is undeniable - but it appears to be assumed that EHRs are the main (only?) source of information about patients. Can this be right? Isn't it also reasonable to suggest that patients themselves are important sources. So are clinicial colleagues, who may have seen the patient before you, or who may have specialist insights that help you to work out a treatment plan. Indeed, isn't a key skill of any doctor the ability to integrate otherwise fragmented information, from many sources, and make sense of it? Is the deep assumption about the key role of EHRs in clinical practice simply wrong? On a more detailed - but important - point the text refers to search dates in both 2018 and 2019. It is not entirely clear when the initial and subsequent searches were undertaken.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Justin Keen
Institution and Country: University of Leeds, England
Competing interests: None declared

Reviewer: The revised manuscript addresses some of the issues raised in the first round of reviewer comments. There are, though, two substantive issues that still merit attention.

The first concerns the main focus of the article, which in different passages is on patient outcomes, on process measures and on gaps in the literature. Two examples may help to illustrate this point.

Reviewer (1): The title mentions pre-defined safety outcomes, but the conclusions in the Abstract do not state how many of the pre-defined outcomes were supported by empirical evidence (or indeed where the gaps in the literature lie).
Reply: Thanks for this helpful comment. In the Study Selection section of the abstract we state a search for “interventions targeting a list of fourteen pre-defined safety outcomes”. The last line of the Results section of the abstract states “We found no or limited evidence in 13 of 14 pre-defined safety areas.”. We have added “with good evidence limited to medication safety” to address your comment.

Reviewer (2): The Methods section does not state how the 14 pre-defined safety outcomes were identified, or why Zegers' list was used rather than a different one (several have been published).
Reply: Many thanks for the comment. The pre-defined safety outcomes were taken from the paper by Zegers(1) published in BMJ open with Charles Vincent as a leading UK authority on patient safety as one the senior authors. A key reason for the choice of Zegers as the reference paper was the usage of a validated search algorithm: The search algorithm that Zegers used had been developed and validated in a previous study by Tanon(2). Tanon and co-workers aimed to develop an optimised search strategies for identifying papers on patient safety in MEDLINE, EMBASE and CINAHL. The resulting search algorithm was highly sensitive, specific and outperformed other strategies found in the literature. We are not aware of other algorithms that have been validated for searches of the peer-reviewed literature on patient safety in this robust manner. We have changed the Method section to better reflect our rational.
We agree with the reviewer that there are a significant number of other useful systems to classify safety outcomes. Literature reviews using the Global trigger tool have shown a low sensitivity and

specificity for safety related issues(3). The World Health Organisation endorses the list of patient safety 'domains' below(4). There is significant overlap between the outcomes described and those that we used for our search: Health care-associated infections, adverse drug events, falls. Other terms are processes: unsafe injection, unsafe blood products – both would result in infections or adverse drug events. 'Maternal care' is not a safety outcome but an area of care.

Table: WHO patient safety domains
1. Health care-associated infection (HCAI)
2. Maternal care
3. Adverse drug events
4. Adverse events devices
5. Unsafe injections
6. Unsafe blood products
7. Misdiagnosis
8. Surgical and anesthetic error
9. Falls

Reviewer (3): The main source provided is Table 2, but most of the measures listed appear to be process measures. So, is the focus on processes or outcomes?
Reply: Thanks for your comment. Table 2 lists the "impact of implementation of EHR on pre-defined patient safety areas". We added additional classification of the study results in the right column of the table to specify where the evidence exists for assessing the impact on process and outcome measures.

Reviewer (4): The second substantive issue concerns the 'diffuse' nature of EHRs. Observational studies show that they can be, and are, used in myriad different ways in different hospitals. Linking the EHR intervention to a specific outcome will always be difficult, however ingenious the research team. Indeed, as some of the reported studies show, it can be difficult to design robust studies even where process changes are the end-points.
Reply: Thanks. We would absolutely agree with the reviewer's comment above. We have re-emphasised this in the discussion. Nevertheless, it has been part of the advertising and mythology of EHRs that they will allow clinicians to drive outcomes and safety. In fact, as we started the reply to the reviewer, our hospital trust announced the introduction of a new EHR – and the company website boasts improved safety – again without any evidence to support this claim. We, like the reviewer, are absolutely convinced that EHRs with the right human factors-informed design, introduced in the right way and with the right quality assurance can absolutely support improvement of clinical outcomes including safety outcomes. At this moment in time it appears that the evidence for the current generation of EHRs is still weak.
Another piece of anecdotal evidence illustrate the face-validity of our findings: We have reviewed the vital sign charts from some of the globally leading EHR providers and found challenging designs, with an inability to see related vital signs in the same view, and lack of recommended colour schemes etc. User-interfaces in commonly used EHRs might not always focus on finding the most safety relevant information in the fastest and most effective manner. In our view there remain important human factor and system design issues that require further exploration.

Reviewer (5): The significance of this point lies in the authors' conclusions that we need better metrics, and deeper insights into the poor quality of EHRs that clinicians are expected to use. An alternative conclusion is that randomised trials alone cannot provide definitive evidence about the effects of EHRs. Rather, we need carefully designed suites of experimental and observational studies - the latter, in particular, to understand quite why EHRs are still so clunky, in the age of Amazon, WhatsApp and so on. That, arguably, is why some systematic reviews have sought to piece together evidence about both processes and outcomes, using Realist and Meta-Narrative methods.

Reply: We absolutely and wholeheartedly agree with the suggestion by the reviewer and have added the comment to the discussion. We believe there is an urgent need for better evidence about both processes and outcomes and research using adaptive Realist and Meta-Narrative methods.

Reviewer (6): A broader point follows. Comments are made to the effect that clinicians need high quality information to treat patients - this is undeniable - but it appears to be assumed that EHRs are the main (only?) source of information about patients. Can this be right? Isn't it also reasonable to suggest that patients themselves are important sources. So are clinicial colleagues, who may have seen the patient before you, or who may have specialist insights that help you to work out a treatment plan. Indeed, isn't a key skill of any doctor the ability to integrate otherwise fragmented information, from many sources, and make sense of it? Is the deep assumption about the key role of EHRs in clinical practice simply wrong?

Reply: Many thanks for raising these points. We can't agree more with the reviewer that patients are an important source of information and the examination of systems for patients and their carers to contribute to their own safety in hospital—this is the theme of my fellowship with The Health Foundation. We added relevant references to the discussion(5,6).
Based on what we have learned from our work on patient safety we would challenge the hypothesis that it is the skill of a doctor to integrate poorly designed and fragmented information. Since the accident in the nuclear reactor in Three Mile Island in 1979, copious evidence from a wide range of industries and including healthcare have been provided for the need to reduce the cognitive workload on operators to enable safe decision making. Unstructured and fragmented information is at the core of countless serious adverse events and the links between fragmented information and patient harm are well established in the literature (7). Human factors and ergonomics design are part of the safety assurance of medical devices (8): On a regulatory level there are stringent test regimes to detect problems in ergonomics that facilitate wrong usage for medical devices; however, and remarkably, electronic health records are not currently covered by these stringent safety and design standards. We have added this to the discussion.

Reviewer (7): On a more detailed - but important - point the text refers to search dates in both 2018 and 2019. It is not entirely clear when the initial and subsequent searches were undertaken.
Reply: Many thanks: The search included studies published between 2008 and 2018. The initial search was conducted in March 2019 with the supplementary searches run in December 2019. We have clarified this in the method section.

VERSION 3 – REVIEW

REVIEWER	Justin Keen University of Leeds, England
REVIEW RETURNED	06-Nov-2020

GENERAL COMMENTS	I have read the third version of the article carefully. My conclusion is that the authors have not addressed the two substantive issues raised in my last review - cliarifying the focus of the article, and the extent to which experimental study designs can reasonably be expected to shed light on the research question. There is still a tension, throughout the article, between evidence and argument about process and outcome measures. And, the Discussion and Conclusions still don't properly discuss the extent to which the 14 pre-defined measures really focus on processes (e.g. staffing, safety culture) rather than outcomes.