Author’s response to reviews

Title: A systematic review of the effects of care provided with and without diagnostic clinical prediction rules

Authors:

Sharon Sanders (ssanders@bond.edu.au)
John Rathbone (jrathbon@bond.edu.au)
Katy Bell (katy.bell@sydney.edu.au)
Paul Glasziou (pglaszio@bond.edu.au)
Jenny Doust (jdoust@bond.edu.au)

Version: 1 Date: 22 Mar 2017

Author’s response to reviews:

Ann Van Den Bruel
Associate Editor
Diagnostic and Prognostic Research

15/3/17

Dear Dr Van Den Bruel

Re: Manuscript DAPR-D-16-00015

Thank you for inviting us to resubmit our paper titled ‘A systematic review of the effects of care provided with and without diagnostic clinical prediction rules’.

We sincerely thank the reviewers for their considered and constructive comments on our paper. Our responses to the reviewer comments, and our proposed changes to the manuscript, are presented on the following pages.
Reviewer #1

In this paper the authors conducted a systematic review to assess the effects of diagnostic clinical prediction rules (CPRs) on patient and process of care outcomes. This is a rigorously conducted, well-written review performed by very experienced team of systematic reviewers and covering an important topic.

Over the past decades, the number of diagnostic studies focusing on developing and validating CPRs has sharply increased, but far fewer studies have focused on evaluating the effects of implementing diagnostic CPRs on patient and process of care outcomes. This latter is of utmost importance since diagnostic accuracy does not necessarily translate into improved patient and process of care outcomes. Although supporting publication of this paper, I have some comments to be addressed by the authors.

1. Abstract, Results; The authors should consider adding the results for suspected (bacterial) pneumonia, e.g. “For suspected pneumonia, diagnostic CPRs reduced antibiotic prescriptions (3 studies) with no difference in unfavourable outcomes between interventions (2 studies)” or provide justification why these results were not presented in the abstract.

RESPONSE: We agree that these results should be included in the abstract

CHANGE: We have added the following text (in red) to the Abstract, Results section; For suspected Group A Streptococcus throat infection, diagnostic CPRs reduced symptoms (1 study) and antibiotic prescriptions (5 studies, RR 0.86 95%CI 0.75 to 0.99). For suspected cardiac chest pain, diagnostic strategies incorporating a CPR improved early discharge rates (1 study), decreased objective cardiac testing (1 study) and decreased hospitalisations (1 study). For ankle injuries, Ottawa Ankle Rules reduced radiography when used with clinical examination (1 study) but had no effect on length of stay as a triage test (1 study). For suspected acute appendicitis, CPRs had no effect on rates of perforated appendix (1 study), or the number of nontherapeutic operations (5 studies, RR 0.68 95%CI 0.43 to 1.08). For suspected pneumonia, CPRs reduced antibiotic prescribing with no difference in unfavourable outcomes (3 studies). For children with possible serious bacterial infection, diagnostic CPRs did not improve process of care outcomes (3 studies).
2. a. Abstract, Conclusion; I would recommend authors to emphasize the apparent lack of studies evaluating the effects of diagnostic CPRs on patient and process of care outcomes. Despite including “a range of conditions commonly encountered in clinical medicine”, the authors identified only 27 RCTs for 14 conditions. In the light of the large number of diagnostic CPRs published over the past decades, this is a rather low number and should be regarded as important/key finding of this review.

RESPONSE: We agree that this finding should be clearly articulated and emphasised in the review Abstract

CHANGE: We have made the following changes to the Abstract, Conclusion section; Conclusion: Studies of diagnostic CPRs infrequently reported patient outcomes. There are few randomised trials of diagnostic CPRs and patient outcomes are infrequently reported. Diagnostic CPRs had a positive effect on process outcomes in some clinical conditions however many studies were at unclear or high risk of bias and but these results may be context specific. Future studies should seek to detail how the CPR might alter the diagnostic pathway, report effects on both patient and process outcomes, and to improve reporting of the study interventions and implementation.

b. Abstract, Conclusion; Although authors rigorously assessed risk of bias for all included studies and most were judged at high or unclear risk of bias, it is not apparent from the Abstract that results should be interpreted with caution. Although this has been highlighted in the Discussion section: “The conclusions drawn in this review…., caution is advised in interpretation of their results”, such a statement is lacking in the Conclusion section of the abstract.

RESPONSE: Agreed

CHANGE: We have added the following statement (in red) to the Abstract, Conclusion; Diagnostic CPRs had a positive effect on process outcomes in some clinical conditions however many studies were at unclear or high risk of bias and but these results may be context specific.

3. Methods, Risk of bias assessment; Authors should provide additional information how attrition bias was classified into low, unclear or high risk of bias and how reporting bias was assessed.

RESPONSE: Agreed.
CHANGE: The following text has been added to the Methods, Risk of bias section. Assessments of risk of bias arising from incomplete outcome data were based on both the number of missing outcomes and the reasons for the missing outcomes with studies considered to be at low risk of attrition bias if there was no missing outcome data, or if the missing outcome data was balanced between study arms and the reasons given for the missing data did not have different implications for each study arm. Selective reporting within each study was assessed by comparing outcomes listed in the study protocol and the outcomes presented in the published report. When a study protocol was not available, the outcomes described in the method section of the published report were compared to the outcomes presented in the results section.

4. Results, Risk of bias; Table 2 is rather lengthy and somewhat difficult to interpret. The authors may consider to replace this Table by the Cochrane “Risk of bias” summary format.

RESPONSE: We agree that a summary risk of bias table will assist readers of the review.

CHANGE: We have removed the existing Table 2 which shows the risk of bias for each domain of bias for each included study, and replaced it with a summary risk of bias table. The summary risk of bias table shows our judgments on the risk of bias for each domain of bias presented as percentages across all included studies (part a of the figure). Part b of the figure shows our judgments on the risk of bias for domains relevant to the included cluster randomized trials. The original Table 2 has been placed in Additional File 5.

5. a. Results, Assessment of the reporting of interventions; This section is better placed after the Risk of bias section (and prior to describing the effects of diagnostic CPRs).

RESPONSE: Agreed

CHANGE: The section titled “Assessment of the reporting of interventions” has been moved so it now appears after the RESULTS, Risk of bias section, and before the Effects of diagnostic strategies incorporating diagnostic clinical prediction rules section.

b. Results, Assessment of reporting of interventions; Table 3 is missing from the manuscript

RESPONSE: We apologise for labelling this incorrectly in the text of the manuscript

CHANGE: The table referred to as Table 3 appears as a table in Additional File 6

6. Results, Effects of diagnostic strategies incorporating diagnostic CPRs; The sections relating to suspected Group A streptococcus throat infections and suspected acute appendicitis are
lengthy. Authors may consider to shorten these sections by highlighting the key findings and refer to the tabulated results in Additional File 5 for the other results.

RESPONSE: Agreed

CHANGE: The following text has been removed from the Results, Effects of diagnostic strategies incorporating diagnostic CPRs, Studies of diagnostic CPRs for suspected Group A streptococcus throat infection; Two trials reported the appropriateness of antibiotic prescribing. The first [33] reported no difference between the study arms in unnecessary antibiotic prescriptions, defined as a prescription for an antibiotic in a patient with a throat culture (which was performed in all patients) negative for Group A streptococcus (OR adjusted 0.76 95%CI 0.42 to 1.240). In the second study [32] appropriateness of prescribing was determined by assessing whether the prescribing decisions of the experimental group were more like the management recommendations of the CPR than the control group, and by comparing the proportions of patient prescribed antibiotics at different levels of probability of GAS infection. In the experimental group antibiotic prescribing and throat culture use corresponded more closely to suggested management recommendations, and a greater reduction in prescribing was observed for patients with low probability of GAS infection. The text has been replaced with the following statement referring the reading to the results table in Additional File 7. The results of studies reporting the effects of diagnostic CPRs on other clinical decisions including test use and the appropriateness of antibiotic prescribing are shown in Additional File 7.

The following text has been removed from the Results, Effects of diagnostic strategies incorporating diagnostic CPRs, Studies of diagnostic CPRs for suspected acute appendicitis; The effect of care provided with CPRs on admission rates was conflicting. Admission rates were reduced in one study compared to clinical judgment without a diagnostic aid (RR 0.90 95%CI 0.82 to 0.99). [19] CPRs did not reduce admission rates in 2 studies compared to a control group where clinicians used a standard data collection form (RR 1.0 95%CI 0.91 to 1.12) [19] and usual clinical assessment (RR 0.72 95%CI 0.49 to 1.05). [26] There was an increase in delayed treatment in association with perforation in the experimental arm of two trials, but the difference was not statistically significant (RR 3.0 95%CI 0.13 to 69.7, [26] and RR 2.22 95%CI 0.44 to 11.26). [24]. In one small trial (n=42) reporting time to surgery (from randomisation to skin preparation), time to surgery was significantly shorter in the experimental group (Median 2.05 hrs vs 8.35 hrs p=0.03). [26]. The text has been replaced with the following statement referring the reading to the results table in Additional File 7. The results of studies reporting the effects of diagnostic CPRs on other process of care outcomes including admission rates and time to surgery are shown in Additional File 7.

7. Results, Effects of diagnostic strategies incorporating diagnostic CPRs, Process of care outcomes; In some sections, information about the impact on patients is lacking, e.g. for
suspected Group A Streptococcus, diagnostic CPRs reduced antibiotic prescriptions (5 studies) but this is only relevant if no difference in unfavourable outcomes between the two strategies were observed. Did the trial authors report on these findings? If so, this information should be added. If not, a short statement should be added that no unfavourable outcome data were provided in the included studies. The same applies to some other related sections.

RESPONSE: We agree this information will contribute to the interpretation of results.

CHANGE: The following text (in red) has been added to the section Results, Effects of diagnostic strategies incorporating diagnostic CPRs, Studies of diagnostic CPRs for suspected Group A streptococcus throat infection, Process of care outcomes; All 5 studies reported clinical decisions to prescribe antibiotics. [18, 19, 32-34] In the 3 studies reporting clinical decisions to prescribe as the primary study outcome, CPRs were associated with a non-significant reduction in prescribing. (19, 32, 33) One of these studies reported no difference in emergency department or outpatient visits as a proxy for the appropriateness of care. (32) Complications were either did not occur or were not reported in these studies. In pooled analysis, CPRs reduced antibiotic prescriptions compared to care provided without a CPR (pooled RR 0.86 95% CI 0.75 to 0.99) (Figure 3. Meta-analysis of sore throat studies for the outcome antibiotic prescriptions).

The following text (in red) has been added to the section Results, Effects of diagnostic strategies incorporating diagnostic CPRs, Studies of diagnostic CPRs for suspected acute appendicitis, Process of care outcomes; The results of two studies providing data on the effect of CPRs on duration of hospitalisation among patients with suspected appendicitis are conflicting. One small study of the Alvarado score reported significantly shorter duration of hospitalisation in the intervention group (Median 37.00 hrs vs 60.40 hrs p=0.03) ‘without significant increase’ in perforation (1 perforation occurred in the intervention group and 2 perforations in the control group) , [27] while the other study reported no difference in mean duration of hospital stay between a diagnostic protocol incorporating the Alvarado score and graded compression ultrasound, and the control group (Mean 53.4 hrs vs 54.5 hrs p=0.84). [25]

The results sections for other clinical conditions have been checked and we believe provide sufficient information on both the benefits and harms of the interventions.

8. Results, Studies of diagnostic CPRs for suspected acute appendicitis; The authors should justify the reason for presenting the accuracy results for the studies on suspected acute appendicitis. These results don’t match with the objectives of this review, i.e. to evaluate the effect of diagnostic CPRs studies patient and process of care outcomes.

RESPONSE: We agree that the accuracy data presented is not relevant
Three trials report indices of accuracy. [19, 29, 30] The reference standard was a diagnosis obtained at surgery or at discharge in 1 trial [19] or at 1 month follow-up in 2 trials. [29, 30] In one trial in children, there was no difference in sensitivity and specificity of the initial examination between the experimental and control groups (sensitivity 83% vs 85 p value not significant and specificity 69% vs 52 % p value not significant respectively) [29]. In one trial in adults there was no difference in sensitivity (87% vs 89% p value not significant) but specificity of initial clinical assessment was higher (59% vs 80% p=0.03). [30] In the third trial, sensitivity and specificity in the experimental arm was higher than clinical judgment with no diagnostic aids (sensitivity 48% vs 28% p=0. 01 and specificity 98% vs 96% p=0.04) and clinical judgment with standardised data collection (sensitivity 48% vs 42% p=0.48 and specificity 98% vs 96% p=0.01). [19]

9. Results, Studies of diagnostic CPRs for suspected bacterial infection in children with fever; A description of the risk of bias of included studies is missing in this section.

RESPONSE: We agree a statement about the risk of bias in these studies should be provided.

CHANGE: We have added the following text (in red) to the section Results, Studies of diagnostic CPRs for suspected bacterial infection in children with fever; Three studies evaluated 3 different directive CPRs. [35, 42, 43] All 3 were considered at high risk of bias arising from a lack of blinding of care providers and at unclear risk of bias for allocation concealment.

10. Results, Studies of diagnostic CPRs for suspected (bacterial) pneumonia; A description of the risk of bias of included studies is missing in this section.

RESPONSE: We agree a statement about the risk of bias in these studies should be provided.

CHANGE: We have added the following text (in red) to the section Results, Studies of diagnostic CPRs for suspected (bacterial) pneumonia; Three studies evaluated different CPRs in patients with suspected pneumonia, [31] and in patients (unvaccinated and vaccinated against pneumococcus) with non-severe community acquired pneumonia (according to clinical criteria) of unknown aetiology. [40, 44] All 3 studies were judged to be at unclear or high risk of bias on 3 or more domains of bias.
11. Results, Studies of diagnostic CPRs for suspected ankle or mid-foot fracture; A description of the risk of bias of included studies is missing in this section.

RESPONSE: We agree a statement about the risk of bias in these studies should be provided.

CHANGE: We added the following text (in red) to the section Results, Studies of diagnostic CPRs for suspected ankle or mid-foot fracture. Two studies evaluated the impact of the Ottawa Ankle Rules (OARs). [22, 25] In 1 trial, the OARs were used as a triage test. [25] This trial was judged to be at unclear or high risk of bias arising from inadequate randomisation sequence generation, incomplete data and selective reporting. In this trial, standard departmental care was compared to a pathway in which the OARs were applied at presentation: if positive, the patient was x-rayed, and if negative, the patient underwent usual clinical assessment. In the second trial, [22] clinicians in hospitals randomised to the intervention were encouraged to use the OARs as part of their clinical assessment. This study was judged to be at high or unclear risk of bias on 5 of the 6 domains of bias assessed and at unclear risk from baseline differences between the randomised groups. When used as a triage test, the OARs did not decrease total length of stay in the emergency department (mean difference -6.7 minutes 95% CI -20.9 to 7.4), and there was no difference in patient satisfaction ratings and radiography requests between the study groups. [25]

Minor comments: 1a. Abstract, Results; For cardiac chest pain should read as “For suspected cardiac chest pain”. 1b. For acute appendicitis should read as “For suspected acute appendicitis”.

2. Results, Studies of suspected acute coronary syndrome. This heading should be changed to “Studies of diagnostic CPRs for suspected acute coronary syndrome”.

RESPONSE: AGREED

CHANGE: The word ‘suspected’ before ‘cardiac chest pain’ and ‘acute appendicitis’ has been added to the abstract as suggested by the reviewer. The heading in the results section for studies of acute coronary syndrome now reads; Studies of diagnostic CPRs for suspected acute coronary syndromes

Reviewer #2

This is an excellent systematic review from a highly experienced team. The focus of this article is a review of current evidence of the effects of clinical prediction rules on patient centered outcomes and clinical outcomes. This is a timely and useful article, and contributes to our understanding of the general lack of impact or evidence for impact for most clinical prediction rules, and suggests and urgent need to re-evaluate our methods and approach in this field of diagnostic research.
1. Introduction: This is fine, and the authors make the case appropriately. One recent paper to consider citing that might add to their argument is IEEE J Transl Eng Health Med. 2016 Jun 13;4:2800208. Doi: 10.1109/JTEHM.2016.2570222. eCollection 2016. More Than Just Accuracy: A Novel Method to Incorporate Multiple Test Attributes in Evaluating Diagnostic Tests Including Point of Care Tests.

RESPONSE: We thank the reviewer for suggesting this relevant article

CHANGE: We have included this reference (in red text) in the Introduction in support of the statement “However, diagnostic accuracy does not necessarily translate into patient benefits, [2] nor is it a necessary prerequisite for improved patient health as a CPR may alter patient outcomes through other non-decisional routes including by changing the timing of decisions and actions relative to the existing pathway, or through direct effects of the CPR itself. [3, 4] The reference has also been included in the discussion section.

2. Methods: These are good. My only comment is whether the inclusion of only RCTs is too restrictive? They found in the end only 27 trials which is a surprisingly small number. It would be interesting to see whether expanding study design to controlled clinical studies and other designs, albeit less rigorous than RCTs, would produce a much larger body of literature. The small number of trials suggests that RCTs of CPRs are perhaps difficult to perform/fund/undertake in clinical practice, which I would concur with, particularly for conditions with potentially serious outcomes such as serious infection in children or acute appendicitis. Therefore I wonder if their restriction to RCTs is justifiable? However, the methods they used to find, appraise and synthesize the trials were excellent, so this is perhaps a commentary point.

RESPONSE: We agree that our decision to limit the review to randomised comparisons of care with and without a diagnostic CPR should be justified.

CHANGE: We have added the following text (in red) to the discussion section; To our knowledge, this is the first review specifically of diagnostic CPRs across a range of clinical conditions. We are aware of one recently published systematic review evaluating the impact of diagnostic and prognostic clinical prediction rules (as a stand-alone tool) for conditions encountered in primary care. [10] Similarly, this review identified only a small number of studies evaluating the impact of CPRs in a limited number of clinical domains, and found that these studies focused primarily on assessing the effect of CPR use for changing clinician behaviour. Our review was limited to studies randomly allocating participants to care with or without a diagnostic CPR. However, it is likely, as examination of our table of excluded studies and the review of Wallace [10] suggests, that the use of non-randomised and uncontrolled study designs to assess the impact of diagnostic CPRs are reasonably common. This may reflect the practical and methodological challenges in performing randomised trials in the clinical setting.
particularly for conditions with potentially serious and often infrequent outcomes. While examining the evaluations using non-randomised and observational designs would have contributed to the overall evidence base, only randomised controlled trials can rule out the possibility that an observed association between an intervention and outcome is caused by a third factor linked to both. The findings of our review, that CPRs reduce prescribing and test ordering for some conditions, are also generally consistent with existing research evaluating clinical decision support tools more broadly, which has found that some systems can improve test ordering and antibiotic prescribing behaviour. [5-7]

3. Discussion; I have little to add to the detailed discussion. The authors discuss at length the implications of their findings appropriately, and call into question our current methods and approach to evaluation of CPRs in clinical research. In particular they note the lack of evidence for impact of CPRs across the wide range of conditions they identified, and the lack of data on patient centered outcomes. It would be interesting to speculate on whether the current approach researchers have used (as the authors note, deriving more and more CPRs and limited validation and limited impact assessment) should be radically changed – what exactly has been the benefit to clinicians, and policy makers of the considerable investment on research on CPRs with so few implemented worldwide? On the other hand, one might argue that for other diagnostic tests, e.g. new biomarkers or other lab tests, the typical standard is to demonstrate comparative accuracy to existing tests (and perhaps some other advantage e.g. cheaper/more acceptable etc.), and that is considered sufficient for implementation. It is reasonable to expect CPRs to not only show comparative accuracy to some other diagnostic test/process, but then also show impact on clinical or patient centered outcomes? Based on the findings of this review, clearly this does not happen, suggesting these different standards have impacted adoptions of CPRs.

RESPONSE: We agree with the reviewer that this would be a very interesting and worthwhile discussion and have raised this point in the Discussion section of the manuscript.

CHANGE: We have added the following text to the Discussion section; To our knowledge, this is the first review specifically of diagnostic CPRs across a range of clinical conditions. We are aware of one recently published systematic review evaluating the impact of diagnostic and prognostic clinical prediction rules (as a stand-alone tool) for conditions encountered in primary care. [11] Similarly, this review identified only a small number of studies evaluating the impact of CPRs in a limited number of clinical domains. This lack of evidence about the effect of use of diagnostic CPRs might be one factor contributing to their limited uptake into practice. Conventionally, CPRs are required to go through the stages of derivation, validation and impact analysis prior to full implementation in practice. This rigorous requirement is seemingly inconsistent with the situation for other diagnostic tests (such as point of care tests) which are frequently and fervently adopted into practice only on the basis of demonstrated practical
advantage or accuracy against a reference standard. Our review was limited to studies randomly allocating participants to care with or without a diagnostic CPR.