Subject: BMJ - Decision on Manuscript ID BMJ-2019-049437

Body: 30-May-2019

Dear Dr. Walker

Manuscript ID BMJ-2019-049437 entitled "Variation in Responsiveness to Warranted Behaviour Change Among NHS Clinicians: a Novel Implementation of Change-Detection Methods in Longitudinal Prescribing Data"

Thank you for sending us your paper. We sent it for external peer review and discussed it at our manuscript committee meeting. We recognise its potential importance and would like to offer publication in the BMJ if you are able to revise to our satisfaction.

We hope very much that you will be willing and able to revise your paper as explained below in the report from the manuscript meeting and we are looking forward to reading the revised version in due course.

Please remember that the author list and order were finalised upon initial submission, and reviewers and editors judged the paper in light of this information, particularly regarding any competing interests. If authors are later added to a paper this process is subverted. In that case, we reserve the right to rescind any previous decision or return the paper to the review process. Please also remember that we reserve the right to require formation of an authorship group when there are a large number of authors.

When you return your revised manuscript, please note that The BMJ requires an ORCID iD for corresponding authors of all research articles. If you do not have an ORCID iD, registration is free and takes a matter of seconds.

John Fletcher
Dr John Fletcher
Associate Editor
The BMJ
jfletcher@bmj.com

*** PLEASE NOTE: This is a two-step process. After clicking on the link, you will be directed to a webpage to confirm. ***

https://mc.manuscriptcentral.com/bmj?URL_MASK=cf1b6c5866c3475898ff8bf34cd01139

Report from The BMJ's manuscript committee meeting

These comments are an attempt to summarise the discussions at the manuscript meeting. They are not an exact transcript.

Chair: John Fletcher, Statistical advisor: Tim Cole, Jose Merino, Tiago Villanueva, Tim Feeney, David Ludwig, Wim Weber, Helen Macdonald

Decision: Put points

Detailed comments from the meeting:

1. We found this quite a technical paper and were not quite sure where it would fit in the journal. You haven't written this as a "Research Methods and Reporting" and we don't suggest you do. As written the main point is to demonstrate the use of a method and we

don't usually publish this sort of paper. However, as editors we quite liked reading about this innovation and we think our readers will find this useful. Please can change the introduction and perhaps a little of the reporting so that you specify a "research question" and so that your manuscript addresses this question.

2. A couple of editors were not convinced that the examples you provide were good examples of where changes in clinical practice were necessary for all patients. There would seem to still be room for a good doctor to prescribe in Cerazette or trimethoprim in some circumstances, though we take your point that you would expect a large shift in average prescribing.

3. 25% of practices excluded for desogestrel and 14% for trimethoprim/nitrofurantoin. It should be possible to reduce values by improving the algortihm. Perhaps practices with all values of 1 should not be excluded.

4. Are examples in Figure 1 typical to illustrate a point, or specially selected to show a clear difference? The purple lines are dotted not dashed. Perhaps use las=1 in figures.

5. Figures 2 and 3 interesting, but still cross-sectional. Magnitude not particularly informative, as most shifting by same amount, particularly after run-in period.

6. It would be interesting to see how timing and gradient correlate across practices – i.e. see scatterplot. This would distinguish between early/late and shallow/steep changers.

7. The mean magnitude of change would be informative. Around 0.7 for desogestrel, but < 0.4 for trimethoprim/nitrofurantoin, indicating the proportion of "uncomplicated" UTI cases.

8. The discussion refers to 6000 practices, but the methods say 8078. Please resolve this diffeence.

9. Please revise your paper to respond to all of the comments by the reviewers. Their reports are available at the end of this letter, below.

In your response please provide, point by point, your replies to the comments made by the reviewers and the editors, explaining how you have dealt with them in the paper.

Comments from Reviewers

Reviewer: 1

Recommendation:

Comments:
Dr. Walker and colleagues present an automated statistical detection approach to detect changes in prescribing behavior in order to quantify variation in speed of adoption and magnitude of warranted changes at healthcare institutions. To evaluate the performance of their approach, they used two example time-series in a large study with English primary care prescribing data: the prescribing of generic desogestrel around the expiry of the Cerazette patent; and the prescribing of nitrofurantoin over trimethoprim around a change in prescribing guidelines. They found that the method was able to automatically and robustly detect changes in prescribing behaviors in both examples. Great variation exists in speed of implementation for these warranted changes. I commend the authors' efforts in creating this computational approach to automatically detect changes in clinical practice, which addresses an important question with rigorous and efficient methodology. Overall, the manuscript is well-written and can be informative to researchers interested in studying diffusion of change

in medical practice. I have some major and minor comments that hopefully will help strengthen the manuscript.
1. Methods, page 5 line 43 – One important feature of the approach is the choice of the level of significant for breaks to control the false-positive rate. The authors used p=0.000001 in the current study. Could you please include some discussion on how to select the level of significance in a given study/sample? It would be helpful for researchers who are interested in applying the method in their studies. Furthermore, how big an impact did this choice of level of significance have on the results in the current study?
2. Methods, page 5 line 14-15 – the investigation excluded "practices with incomplete time series, or those that did not vary during the time series". Please clarify the latter part. It appears that latter part refers to absolute change in value rather than trend in the time series based on the description of results in Data section on page 7 line 13-14.
3. Results, page 7 – This point is related to the second comment. As about 1/4 of practices were excluded from the Cerazette analyses, it would be informative to know the representativeness of the analytical sample to interpret the results. Please describe how these excluded practices compare to those included in the analysis.
4. Results, page 7 – The authors noted that the method could become hypersensitive to change and result in inappropriate detection when the initial variance of the time series was very low. They overcame this problem by "tweaking the maximum size of the block-partitioning". Please elaborate on how this parameter should to be set to avoid this problem of hypersensitivity. Moreover, would this method be able to differentiate larger structural changes from smaller changes such as seasonal variations if both were present?
5. Discussion, page 13 line 52-57 – While I understand the advantage of using the proportion of "undesirable" prescribing (over all prescribing) compared to its absolute volume, the meaning of the term "confounding by indication" is not very clear. As the term could be easily confused with its typical use, (i.e. referring to one type of confounding in studies of treatment-outcome relationship), I would suggest describing or replacing the term "confounding by indication" here.
6. Discussion, page 14 line 10 – there is a typo: the last word in "focused our analyses on practises".
7. The authors described their method and results well, but the manuscript can benefit from elaborating the discussion to include practical guidance for researchers in practice, for example, the requirements of this method on the size of an analysis unit (here, practice) and quality of data.

Additional Questions:
Please enter your name: Xiaojuan Li

Job Title: Research fellow

Institution: Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may

in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/declaration-competing-interests'target='_new'> (please see BMJ policy) please declare them here: None

Reviewer: 2

Recommendation:

Comments:
Thank you for the opportunity to review this manuscript. This was a very interesting topic - as someone who regularly uses interrupted time series to investigate the impact of medicine policies, I can see it being very useful for many applications. The manuscript was very well written and easy to follow, although I'd appreciate a few more details about certain things as described below:

1. From the perspective of a potential user of this method who isn't familiar with trend-indicator saturation, I would be interested in a few more details about the underlying regression model. Presumably it can account for autocorrelation, seasonality and cyclic or secular trends in the time series?

2. Additionally, I am curious if this method is broadly applicable, or are there specific requirements and/or assumptions of the time series data that must be met? (for example, a minimum number of time points).

3. I note that many practices did not observe a significant shift until well after the intervention(s). Obviously the further in time from the intervention the change occurs, the less sure you can be it is due to the intervention itself. Now, for the examples in this paper it may not matter so much, where the focus is more on improvements in prescribing and there are few other alternative explanations - but for other scenarios, it should be noted that an automated approach may identify changes potentially unrelated to the intervention and strategies would be needed to exclude these.

4. What about practices that experienced a change prior to the intervention date? Are these included in the summary measures in Table 1? It may be worth mentioning in how many cases the algorithm identified changes clearly unrelated to the interventions (i.e. beforehand). Also, how many practices experienced no change?

5. It wasn't entirely clear to me if the algorithm could potentially identify multiple changes within a practice, and if so how it deals with them. Looking at the graphs in the supplementary material there seems to always be one main change identified.

6. Can you clarify the Cerazette intervention? Prior to patent expiry and availability of the generic options, wouldn't only Cerazette be available and thus Cerazette prescribing be 100%?

Additional Questions:
Please enter your name: Dr Andrea Schaffer

Job Title: Research Fellow

Institution: Centre for Big Data Research in Health, UNSW Sydney

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/declaration-competing-interests'target='_new'> (please see BMJ policy) please declare them here:

Reviewer: 3

Recommendation:

Comments:
Thank you for the opportunity to review this paper. I believe this work adds substantially to the published literature, describing a novel, original method which can be applied across various domains of healthcare to capture behaviour change in response to new evidence or new developments in a field.

Although this paper is somewhat technical in its description of this novel method, behaviour change and implementation are relevant across any field of medicine or healthcare. In an era of increasing availability of routine health data, this approach has multiple applications which are well described in the discussion section (including differentiating between warranted and unwarranted variation in healthcare, identifying best practices in implementation, and driving quality improvement). I do believe it is important to multiple audiences and is best suited to a general medical journal, and has the potential to direct general practitioners to the team's web platform to examine these measures for their own practices.

The research question and study design are described appropriately. I have some minor suggestions of some elements of the methods could be clarified.
Page 5/19, Line 16 - It would be helpful to clarify here whether 'incomplete time series' complete data for all months during the study period, and perhaps include the reasons for missing values here, rather than at the beginning of the results.
Page 5/19, Line 55 - Although clear further in the paper, I would suggest mentioning here that each of these graphs related to an individual practice.
Page 6/19, Line 20 - The final line describing the slope measure ("until the mean of the time series at the end of the time series") is somewhat unclear.

Page 6/19, Line 26 - Refers to "mean proportion at the end of the time period", can the authors perhaps clarify if this is study time period rather than behaviour change time period? Also, should this be the proportion at the end of the study time period, or mean proportion over some period of time? For the final part of this sentence "at the time of the largest detected change", I would suggest amending to "at the start time of the largest detected change" or similar.

Overall the results do address the research question and I found them clear. One suggestion in relation to Table 1, would be to clarify whether the timing measure for "UTI antibiotics" relates to timing after the guidance change. It could also be interesting to add the equivalent metric for timing after the Quality Premium incentive.

Frank Moriarty.

Additional Questions:
Please enter your name: Frank Moriarty

Job Title: Senior research fellow

Institution: Royal College of Surgeons in Ireland

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/declaration-competing-interests'target='_new'> (please see BMJ policy) please declare them here: