PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Choosing a model to predict hospital admission. An observational study of new variants of predictive models for case finding.
AUTHORS	Georghiou, Theo; Billings, John; Blunt, Ian; Bardsley, Martin

VERSION 1 - REVIEW

REVIEWER	Maria C. Raven, MD, MPH, MSc Assistant Professor of Emergency Medicine Department of Emergency Medicine University of California, San Francisco School of Medicine USA I have no competing interests.
REVIEW RETURNED	26-Jun-2013

GENERAL COMMENTS	This manuscript is well written and with a few exceptions easy to follow. Because currently, policy makers are focused on targeting high cost patients and those at risk of 30-day readmission, this analysis offers some valuable and also some novel data that may help those struggling with how to best identify at-risk populations. Introduction: No issues Methods: No issues Results: 1. Top of page 10: The authors describe that with additional groups of data elements the models identify more, and slightly less ill individuals. What is not clear is if the additional data elements are forcing some of the most ill patients out of the model, or if the new patients are simply lowering the morbidity of the identified group as a whole. It would be nice to have this clarified. If the most high-risk patients are being excluded, please describe this group as this would potentially influence which models policy makers/planners might decide to explore using to target patients for intervention.

	2. Top of page 14 under “Testing alternative population denominators” header-- What is the “GP list denominator”? This paragraph does not make the findings clear. 3. Page 12: The authors state that individual site/locally calibrated models do not outperform the pooled model. The authors state their findings here are not definitive and that future analyses are planned. In terms of other regions that might be planning to develop similar models, it would be helpful if the authors could comment on the population variability (or lack thereof) amongst the sites and if they believe the models performed at a similar level, whether pooled or site specific, because the site populations themselves do not vary much. 4. Important point that GP codes contributed to improved case findings but also to variation due to inconsistency of documentation and highlighting the importance of creating incentives for GPs to document thoroughly. <u>Discussion</u>: 1. It would be helpful to get a sense of how these predictive models compare with other efforts to identify high-risk patients—either commercial risk prediction tools or other tools that have been put forth in the literature to identify patients at risk of future/high utilization. Do these tools outperform any other models for predicting risk? 2. Page 16—the finding that including entire GP registries in the risk model denominator enhanced case finding at potentially earlier intervention points was quite interesting. The authors posit that this is because larger population denominators identify greater numbers of high risk patients, because the larger at-risk pool provides more basis for comparison of high and low risk patients. To me, this is one of the more novel findings and the fact that it is highlighted in the abstract and key findings is appropriate. 3. Limitations should include that for many diagnostic variables related to behavioral health, there may be under-reporting/under-coding. 4. Regarding intervention strategy discussion, consider referencing the current JAMA article by Karen E. Joynt et al, Contribution of Preventable Acute Care Spending to Total Spending for High-Cost Medicare Patients.

	Tables: Appendix B: page 21 lines 10-11—correct table formatting-line formatting is off for 2nd column (CW) on for multiple lines.

REVIEWER	Thierry Chaussalet Professor of Healthcare Modelling University of Westminster UK
REVIEW RETURNED	17-Jul-2013

THE STUDY	This paper compares various predictive models for case finding. The main difference between the models lies in the datasets used (e.g. inpatient only, inpatient + outpatient, inpatient + outpatient + A&E, inpatient + outpatient + A&E + GP, etc) rather than the modelling techniques (logistic regression throughout). This could/should have been made clear at the outset. There could also be more and clearer description of previous models, which probably differ from those proposed in terms of data used rather than modelling methodology. The limitations of the models could also be clearer. I would also suggest to provide more description of variables in a separate table. Also linking GP and SUS/HES data should be explained in more detail, especially as this is one of the interesting features of the analysis. All the above are only suggestions for minor revisions to the paper, which is very well written and presented. A lot of useful data analysis is done. The experiments done on data are clear. The results are well presented and compared using appropriate methodology.

VERSION 1 – AUTHOR RESPONSE

Reviewer: Maria C. Raven

[Results 1: Top of page 10: The authors describe that with additional groups of data elements the models identify more, and slightly less ill individuals. What is not clear is if the additional data elements are forcing some of the most ill patients out of the model, or if the new patients are simply lowering the morbidity of the identified group as a whole. It would be nice to have this clarified. If the most high-risk patients are being excluded, please describe this group as this would potentially influence which models policy makers/planners might decide to explore using to target patients for intervention.]

This is an interesting point. The additional data sets in general did not lower risk scores of higher risk patients, but raised risk scores of lower risk patients putting more above risk score threshold of 50. We have added text to indicate this.

[Results 2: Top of page 14 under “Testing alternative population denominators” header-- What is the “GP list denominator”? This paragraph does not make the findings clear.]

Text altered for clarity: “models using GP list denominator” - changed to “models using the full GP register”.

[Results 3: Page 12: The authors state that individual site/locally calibrated models do not outperform the pooled model. The authors state their findings here are not definitive and that future analyses are planned. In terms of other regions that might be planning to develop similar models, it would be helpful if the authors could comment on the population variability (or lack thereof) amongst the sites and if they believe the models performed at a similar level, whether pooled or site specific, because the site populations themselves do not vary much.]

We did not see any clear patterns between out limited sample of sites – our feeling is that the differences are a product of local coding and information practices – we have added some text on this to the discussion

[Results 4: Important point that GP codes contributed to improved case findings but also to variation due to inconsistency of documentation and highlighting the importance of creating incentives for GPs to document thoroughly.]

Good point – see point above

[It would be helpful to get a sense of how these predictive models compare with other efforts to identify high-risk patients—either commercial risk prediction tools or other tools that have been put forth in the literature to identify patients at risk of future/high utilization. Do these tools outperform any other models for predicting risk?]

We have added some further text on this.

[Page 16—the finding that including entire GP registries in the risk model denominator enhanced case finding at potentially earlier intervention points was quite interesting. The authors posit that this is because larger population denominators identify greater numbers of high risk patients, because the larger at-risk pool provides more basis for comparison of high and low risk patients. To me, this is one of the more novel findings and the fact that it is highlighted in the abstract and key findings is appropriate.]

We agree

[Limitations should include that for many diagnostic variables related to behavioral health, there may be under-reporting/under-coding.]

We have added text on this

[Regarding intervention strategy discussion, consider referencing the current JAMA article by Karen E. Joynt et al, Contribution of Preventable Acute Care Spending to Total Spending for High-Cost Medicare Patients.]

We have added this to our discussion

[Tables: Appendix B: page 21 lines 10-11—correct table formatting-line formatting is off for 2nd column (CW) on for multiple lines.]

We have fixed the formatting

Reviewer: Thierry Chaussalet

[The main difference between the models lies in the datasets used (e.g. inpatient only, inpatient + outpatient, inpatient + outpatient + A&E, inpatient + outpatient + A&E + GP, etc) rather than the modelling techniques (logistic regression throughout). This could/should have been made clear at the outset.]

We have added some text to the introduction to emphasise this point

[There could also be more and clearer description of previous models, which probably differ from those proposed in terms of data used rather than modelling methodology. The limitations of the models could also be clearer.]

We have added some text to the introduction to emphasise this point

[I would also suggest to provide more description of variables in a separate table.]

A further table has been added as appendix D

[Also linking GP and SUS/HES data should be explained in more detail, especially as this is one of the interesting features of the analysis.]

A note and reference has been added.

Other changes:
Author names reordered; reference 12 onwards renumbered after new reference 12 added; updated word count