Author’s response to reviews

Title: Post-Acute Care Referral in United States of America: A Multiregional Study of Factors Associated with Referral Destination in a Cohort of Patients with Coronary Artery Bypass Graft or Valve Replacement

Authors:

Ineen Sultana (ineensultana@tamu.edu)

Madhav Erraguntla (merraguntla@tamu.edu)

Hye-Chung Kum (kum@tamu.edu)

Dursun Delen (dursun.delen@okstate.edu)

Mark Lawley (malawley@tamu.edu)

Version: 1 Date: 14 Dec 2018

Author’s response to reviews:

Response to Reviewer(s)' Comments

Dear Editor,

We want to thank the reviewers for their valuable time and useful contribution in reviewing this manuscript. We appreciated all comments and inputs by the reviewers and made changes to the manuscript based on those. We feel that the manuscript is much improved as a result.

To improve clarity and understanding of the objective of our study, we made changes mainly in the background, discussion, and conclusion sections of our manuscript. All the changes made in the revised manuscript are highlighted using the Yellow color. Besides, point-by-point responses to reviewers’ specific comments are addressed separately in this response letter.

We sincerely hope these modifications address the feedbacks provided by the reviewers. Looking forward for your additional inputs and feedbacks.

Regards,

On behalf of all authors,

Ineen Sultana
Reviewer: 1

- Overall – The objective of this manuscript is not clear. The authors are looking at several issues: patient-, provider-, and region-level, without accounting for the effect of these on post-acute discharge settings. There is a good amount of literature, mostly coming from the Medicare claims data, that has shown the association between hospital and regional variables with post-acute care discharges. Use of EHR could have the benefits of looking at some of these variables (mostly patient-level) at a more granular level. May be this is not the case? In any case, clearly defining the objective is important.

The clustering effect of patients in hospitals in regions in well documents, thus the choice of single level analysis is not appropriate for this type of study. The authors should use analytical methods that can address the issue of data clustering.

Interpretation of the results needs to be focused on few key variables, that are pre-defined, backed up by a priori research questions.

Response:

1. In the revised version, the background, discussion, and conclusion section has been updated to clarify the specific objective of our study and interpretation.

2. We conducted a multilevel study in another work using the same data and final cohort, and observed very minimal clustering effect from acute care hospital ID. However, we didn’t include hospital ID as an independent variable in the model of this study since it was found as an insignificant variable for our model based on our variable selection criteria. Therefore, we also ignored the nested/ clustering from hospital level in this study and focused on bringing up a generic and comprehensive understanding of the post-acute referral pattern based on a single level study. We limited our focus on identifying the significant risk factors based on the variation of regional healthcare delivery and practice, patient demographics, and clinical conditions.
3. We didn’t check the clustering effects of patients within the same census division before and
the authors agree with the reviewer that the issue of data clustering for patients within the same
region should be checked. Therefore, we performed a multilevel analysis considering the random
effects from census division and found that the difference between the single and multilevel
results is negligible. For example, the difference of model misclassification errors is 0.21%; the
AIC difference between two models is 0.417%. It indicates that there is none or minimal
clustering effect of census divisions in our dataset and therefore, we want to limit the scope of
this study with single level analysis. However, we mentioned this finding in our revised
manuscript in the methods section (subsection 2.4) including some references in the literature
that performed multilevel analysis for similar type of study. Currently, we are working on CMS
claim data to extend this study including post-acute care costs and post-acute stay to further
improve our investigation and we will consider different set of multilevel models in that paper.

Specific Comments:

Background:

1. Clarify what authors meant by "joint trauma"? Are these fractures, replacements.
Response: In orthopedic terminology, “joint trauma” is actually used to refer to severe injuries to
musculoskeletal joint system of human body with significant damage. In our manuscript,
however, joint trauma has been used to indicate especially joint replacement like hip or knee
joint replacement surgery and usually, the patients with this type of joint replacement need post-
acute care for physical therapy and other support to accelerate the healing process. In the revised
version, the word “joint trauma” has been changed to “joint replacement” to make the statement
clear and specific.

2. Long-term acute care hospitals (LTC) needs to be replaced by long-term care hospitals
(LTCH).
Response: Done

Methods:

3. More information is needed on Cerner Health Facts data warehouse. What information,
variables, follow up.
Response: The Cerner’s HealthFacts data warehouse contains electronic medical records for over 63 Million unique patients obtained from more than 400 Cerner clients (hospital systems) in the US. This, de-identified fully HIPAA compliant relational database covers 16 years of longitudinal episodes, connected at the patient level using a unique patient identifier. The data contains information on patient specifics, hospital specifics, doctor specifics, diagnostic/clinical information, lab, pharmacy, and billing data.

In the revised version, the data source (section 2.1) portion has been updated to include the above-mentioned information about the Cerner Health Facts data warehouse.

4. Provide more details on access to data, since it was obtained from a third party.

Response: The Cerner’s HealthFacts data warehouse is donated to the Center for Health Systems Innovation at Oklahoma State University for medical and healthcare related academic research. The current form of the data is not publicly available, but Cerner makes collaborative arrangements on a case-by-case basis with universities and non-profits to share its data for meaningful research. For our study, the data is extracted by one of the co-authors on this paper (he is also the research director for the research center that provided the data). The details are updated in the manuscript in study design (section 2.2) paragraph.

5. Why only ICD9 procedure codes were used, and not CPT code or combination of the two.

Response: It was a matter of choice based on the way the data is organized in the data warehouse. That is, the data warehouse organized the records in tables and relationships that readily allows for data extraction on ICD-9 procedure codes. If needed, a more elaborate, multi-step data extraction procedure could be used to obtain the data using CPT codes. But as per our understanding, it would provide us with the same sample cohort. CPT codes are always used in conjunction with ICD-9-CM or ICD-10-CM numerical diagnostic coding during the electronic medical billing process. Therefore, using one identifier which is globally recognized for recording any diagnosis, diseases, and procedures is sufficient for data extraction for the purpose of this study.

6. Validation of Post-acute discharge destination: There is no mention of validation of this variable. From claims point of view, we know accuracy of such variable is a questionable, considering this is the primary variable of interest it is important to address this issue.
Response: The validation of post-acute discharge destination was beyond the scope of available information obtained from the data set. From our dataset, we just had the information, where a patient was referred to discharge after acute care hospitalization. We don’t have any further information of true enrollment to post-acute care or any follow-up information, which is actually one of the limitations of this study. The limitation is identified in the discussion section (Last Paragraph) of the paper.

7. Condition-specific comorbidity: Please provide more details on how these variables were identified/defined (ICD9 codes?).

Response: We have considered 13 condition specific comorbidities in our model and these 13 comorbid conditions were chosen based on top most frequency. That means these 13 comorbid diagnoses were noticed commonly in most of the patients in our sample cohort. The ICD-9 codes of all these 13 comorbid conditions are added in the revised manuscript (Last paragraph of Definition and Variable section 2.3)

8. Predictive Model: It is somewhat misleading to call these as "predictive models" while in reality the models are simply testing association (strength of association) between several covariates with post-acute discharge destination.

Response: The authors completely agree with this comment of the reviewer and therefore barely used ‘predictive model’ word to refer their own study. We had used the term ‘predictive model’ two times for our own study which has been replaced using the term ‘analytical model’ in the revised manuscript. However, the author used ‘predictive model’ in the literature review section to refer some of the existing similar type of relevant literature, where the literature mentioned those study as predictive model. In the revised manuscript, we removed those predictive modeling references and include some new references to support the changes and modification in our background and discussion section.

9. Clustering data structure: As mentioned in overall comments section, there is a good amount of literature that talks about implications for ignoring clustering/nesting in data structure, as is the case with this study. The authors should carefully evaluate primary variable(s) of interest and what level (patient, hospitals, or region). Analyzing all these variables at the same level can provide biased estimation.
Response: Thank you for your comment. Yes, the authors agree with the comment that careful evaluation of variables of interest and their level is required. But as we explained in the response of overall comment, for our dataset, we checked the clustering of hospitals before in another relevant work with the same dataset and while working for this revision, we performed multilevel analysis to consider the random effect of census division and found no/minimal effect. Therefore, the authors want to limit the scope of this research in single-level analysis.

Results:
Response: No changes in the result as there is no change in the analytical method

Discussion/Conclusion:
Response: Conclusion is modified to bring clarity.
Reviewer: 2

1. Abstract: You may spell out USA when first appears in the article. Add full stop after Charlson comorbidity index (page 1, line 33).

Response: Done

2. Background: a well-written study background and comprehensive literature review of relevant studies with clear identification of study gaps. The statistical data provided in paragraph one of background should be the USA data. You may specify that these are USA data.

Response: Thank you for your comment. In the revised manuscript, the background section has been updated to accommodate the suggestion. Yes, the statistical data provided in paragraph one of the background section is USA data, and it has been specified in the revised manuscript.

3. Methods: The study methods in term of data source, study population and variables are generally clearly described. You may add the study design such as a retrospective cohort study, and a little information on data collection procedure, who extracted the data, and in what way. Why patients with length of stay $>75$ days were excluded. Elaborate a bit on what do you mean that "patients with missing predictive variables (n=2685) were excluded." Do these patients miss all predictive variables or some variables?

Response: In the revised manuscript, study design has been added (section 2.2). Patients with length of stay $>75$ days were excluded because this is a very high and unusual prolonged stay and from our exploratory analysis, we identified this as an outlier. This indicates a potential data reporting error or very rare event. Therefore, to maintain data integrity, we excluded those patients. Patients with a missing variables refer to those patients who had a missing variable in our selected or sorted variables to include in the model. No, these patients do not miss all the predictor variables. For example, in our dataset we excluded 1022 patients who had ‘Discharge location’ information missing, 195 patients who had ‘race’ information missing and 1468
patients who had information missing about marital status. However, this breakdown is mentioned in the flow diagram of the final sample cohort processing through data analysis (Figure 1).

4. Normally, we include data analysis methods under the section of "Methods". You may move your "Descriptive Analysis and Model Development" to the "Methods" section. Please specify which statistical software you used for data analysis, e.g. SPSS.

Response: ‘Descriptive Analysis and Model Development’ has been moved to Methods section in the revised manuscript. The data analysis and all the statistical tests were carried out in R version 3.2.3 and this was mentioned in the last line of the ‘Descriptive Analysis and Model Development’ (section 2.4 in the revised manuscript).

5. If possible, please add your IRB reference number.

Response: The IRB reference number is IRB2016-0453M and it has been added in the revised manuscript in the last line of data source (section 2.1) portion.

6. "Variables with p-value less than 0.1 in the bivariate test were included as candidates in the multinomial logistic regression model." Please add a reference for this. Because the cut-off of 0.2 is also sometimes used by the researchers to select explorative independent variables being included in the model.

Response: Yes, the authors agree that the cut-off can be 0.2 also. In fact, we noticed three cut off values frequently used in the relevant existing literature for variable selection (0.05,0.1 and 0.2). It was our choice to consider the cut off as 0.1 and in the revised manuscript, we have added a reference that also used cut off value 0.1.

7. May not need to explain in so detail about the "relative risk ratio" (page 10, line 50 to page 11, line 18), because this is a common knowledge known to the researchers. Just simply go straight to interpret your results will be good enough.
Response: The detail of “relative risk ratio” has been removed as per reviewer’s suggestion in the revised manuscript.

8. The results are generally well-discussed. I am not sure whether the authors have found similar studies being conducted in other countries, and it would be good if you can do some kind of cross-cultural comparison so that the paper can be more interested to the international readers.

Response: Thank you for your thoughtful comment and suggestion. Yes, we noticed there are some studies on cardiac patient’s rehabilitation referral based Canadian data set. We have included one study in the discussion section and highlighted some cross-cultural comparison.

9. Good conclusion, maybe a few lines of the study limitations.

Response: Study limitations were mentioned in the last paragraph of the discussion section. Still, we have included a few more lines of comprehensive limitation to the conclusion as the reviewer suggested.