Author’s response to reviews

Title: Expenditure Variations Analysis Using Residuals for Identifying High Health Care Utilizers in a State Medicaid Program

Authors:

Chengliang Yang (ximen14@ufl.edu)
Chris Delcher (cdelcher@ufl.edu)
Elizabeth Shenkman (eshenkm@ufl.edu)
Sanjay Ranka (ranka@cise.ufl.edu)

Version: 1  Date: 10 Oct 2018

Author’s response to reviews:

Dear MIDM Editor and Reviewers,

We are really grateful that you provided the great comments and suggestions. They are extremely helpful to improve this manuscript. Please see below for our responses to your concerns and the highlighted changes in the manuscript to reflect your suggestions.

Technical Comments:

1. Please cite Table 4 in the manuscript.
Response: We fixed it in Page 11, last paragraph.

2. Please provide a 'Declarations' heading.
Response: We added declarations in the end.

Reviewer 1:

First, the results are very interesting but little or even not discussed. A first paragraph briefly highlights the added value of their method, which is good but the following paragraphs focus on
the limitations and the next steps of their research project. Instead, it may be helpful to provide more context and discuss the findings with respect to available literature in this field. Also, the reasons why high utilizers spend more than other people should be further investigated. In addition, the authors should focus greater attention on the potential reasons why certain limitations exist and how they may be overcome. In sum, the manuscript lacks of a proper discussion both on the method and the findings.

Response: We enriched the discussion section with comparisons to existing literatures and recommendations as suggested. (Last paragraph of Page 16)

Second, the particular issue of comorbidity has been raised by the authors and highlighted as crucial to take into account. In contrast, comorbidity seems poorly documented or discussed in the paper.

Response: We added discussion of the two comorbidities in second paragraph of Page 17 as suggested.

Third, while the conclusion efficiently summarizes the study contributions, it lacks of lessons learned for improved and effective policy and practice, especially regarding the group of patients that are potentially "over-utilizing" the health system. Are there specific conclusions to improve the supply or the use of health care services for hypertension (including the result of "the major source of variation was found within one ICD code - ICD-9-CM 4019 of unspecified essential hypertension")? Similarly, what can be learned for the Medicaid management of chronic kidney diseases, for which ICD-9-CM 5856 End-stage renal disease was reported with the main driver of expenditure? Authors should also consider to further detail the discussion and conclusion in order to formulate concrete recommendations regarding both inpatient and EDs.

Response: We made more recommendations for policy implications throughout the discussion section as suggested. Also we stated clearer why we need to examine more things before we can conclude with any policy changes in the limitations.

Finally, I have a few other observations:

a. The method reports on a nonzero expenditures group that has been isolated for analysis. What percent of zero expenditures was found?

Response: About 16% population in Texas Medicaid have zero expenditures.
b. The various charts are not readable while printed.
Response: We will improve the readability of the charts with larger font size.

c. Be more specific in the title as the paper focused on the two medical diagnosis of hypertension and chronic kidney diseases. Also consider to add a geographical scope in the title i.e., Texas, USA.
Response: We changed the title to reflect the recommendation.

Reviewer 2:

Manuscript authors used residual distributions for the top 5% super-utilizers v/s others, based on alternative statistical methodologies (linear regression, or LR, and tree-based GBM models), to investigate health care expenditure variations based on ICD-9 disease codes, using administrative data of one U.S. state's Medicaid Program patients. Their results show expenditure disparities between super-utilizers and others to vary within and across ICD-9 disease categories. There is current research and policy interest for payers, state Medicaid programs included, to ID super-utilizers with a view towards crafting out strategies to reduce such expenditures to better manage limited program funds. This said, the authors should benefit from reviewing a more expansive and richer literature published in high quality economics (e.g., J. Health Economics, Health Economics, etc), psychology (e.g., J. Mental Health Policy and Economics, ..., on mental health ED visits) and related discipline journals on the subject matter researched. Statistical modelling has its place; however, as the authors stated their paper has no public health policy to prescribe. To fill this void, please research factors that are more recently known to generate high utilizers' high health care expenditures. Surprisingly, they include: homelessness, lack of reliable transportation modes to obtain timely health care, lack of nutritious meals, living alone, etc. If your manuscript controlled for these factors the huge residual variance identified in your LR and GBM models are bound to shrink. Do you have these variables in your administrative database? If so, control for them and reexamine your residuals and ICD-9 differences for super-utilizers and others. Really need you to tease out some public health policy implications of findings. This way, your revised work is more likely to be value-adding. Also, administrative data have limitations when used for research. Include these in your study limitations. Finally, your results canNOT be generalized to all Medicaid programs since they are based on one (1) state. Make this another of your study's limitations.

Response: We added related work in health economics in introduction as suggested (last second paragraph, Page 3). As the current study is limited to administrative claims data, we are not able to include socio-economical variables in the model. We stated clearer of this limitation and plan
to improve it in next steps. Also we tried to make more health policy recommendations in the discussion (Page 16-18).

Reviewer 3

There are extensive health services research studies that address the high expenditure problem in health care. The authors need to thoroughly review these papers before claiming the validity of their own approach. For example, it is a naive approach to use the OLS to analyze the health care expenditure. Manning's two-step econometric approach is more appropriate to handle the highly skewed distributions of medical expenditures with many zeros. Moreover, including the patients from 18-60 is not appropriate since elder patients consume the majority of the health care resources. These limitations can create systematic bias in the model estimation, which makes the residual an unreliable dependent variable to analyze.

Response: We included more related research in introduction as suggested (last second paragraph, Page 3). We looked at Manning's two-step GGM approach and had a discussion in methods selection in last second paragraph of Page 8. We limited to patients from 18-60 years old because this is the main population of the Medicaid program. We stated clearer about this limitation in discussion.