Reviewer’s report

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

Version: 0 Date: 23 Aug 2018

Reviewer: ALBERT Ade OKUNADE

Reviewer's report:

Please include all comments for the authors in this box rather than uploading your report as an attachment. Please only upload as attachments annotated versions of manuscripts, graphs, supporting materials or other aspects of your report which cannot be included in a text format.

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 can NOT be generalized to all Medicaid programs since they are based on one (1) state. Make this another of your study's limitations.
Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

No

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?
If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

"I declare that I have no competing interests"

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal