PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Multiple chronic conditions at a major urban health system: a retrospective cross-sectional analysis of frequencies, costs and comorbidity patterns
AUTHORS	Majumdar, Usnish; Hunt, Christophe; Doupe, Patrick; Baum, Aaron; Heller, David; Levine, Erica; Kumar, Rashi; Futterman, Robert; Hajat, Cother; Kishore, Sandeep

VERSION 1 - REVIEW

REVIEWER	Martha Pollock Beaumont Health USA
REVIEW RETURNED	26-Feb-2019

GENERAL COMMENTS	Interesting study. wondering about female over male prevalence of two or three chronic conditions, but possibly due to females being more likely to seek care. Overall very useful and well-written study

REVIEWER	Samuel Allemann EA 7425 HESPER Health Services and Performance Research Université Claude Bernard Lyon 1, France
REVIEW RETURNED	18-Mar-2019

GENERAL COMMENTS	The authors present a comprehensive analysis of co-occuring chronic diseases in a large urban health system. They use an interesting novel method of clustering patients by pairs of co-occuring diseases. I feel this is not adequately expressed in the title: I would suggest to change the title, e.g. "Multiple chronic conditions at a major urban health system: a descriptive analysis of frequencies, costs and co-occurence patterns". The authors provide a long list of references (12 citations, 3-14) for one claim. Most of these references are between 10 and over 25 years old. I would question their relevance to costs in today's healthcare setting. I would propose to omit all but the mose recent references. Additionnal small remarks:

	Page 7, line 18: "lives with" instead of "suffers from" [two or more chronic conditions] Table 3: add total yearly cost or number of individuals in cluster

REVIEWER	Ying-Chen Chi, Ph.D. Department of Healthcare Information and Management, Ming Chuan University, Taoyuan, Taiwan
REVIEW RETURNED	13-Apr-2019

GENERAL COMMENTS	This article revealed that the prevalence of multiple chronic conditions(MCC) in Medicaid population across selected boroughs in New York City is higher than the national wide in US, may cause the high expenditure of health care finance and poor health care quality, is an evidence of the challenges for clinical care system and financial utilization. Although the top clusters of MCC by age and gender segments with yearly cost in supplementary Table 1 has been shown, cause the prevalence of MCC among each boroughs in New York City is different also be revealed in Figure 2, the variable of spatial distribution combined with gender and age could provide more describing power by using with this proposed clustering methodology is suggested.

REVIEWER	Simon Brake University of Warwick Medical School United Kingdom
REVIEW RETURNED	02-Jun-2019

GENERAL COMMENTS	It would be useful to describe briefly in the abstract conclusion section the national comparable statsistic (42% of the national population with 2 or more chronic conditions as opposed to 52.7%, and also perhaps the range within the 5 boroughs). The discussion section might benefit from a discussion around the public health and prevention interventions that might benefit identified populations where they were previously unknown, and the tool's potential beneficial use by public authorities to reduce health inequalities and improve health outcomes.

REVIEWER	Kelsey Chalmers The Menzies Centre for Health Policy, Sydney School of Public Health, The University of Sydney
REVIEW RETURNED	28-Jun-2019

GENERAL COMMENTS	Thank you to the authors for their article and their approach in investigating a complex problem and claims data set. The authors have investigated a set of Medicaid claims from a population in New York (inpatient and outpatient claims). They aimed to describe this population based on groupings of patients using combinations of chronic condition diagnoses, the gender and age group of the patients. These groups are then ranked by the total cost (of medical services), frequency and how unexpected (observed/expected) the co-occurring chronic conditions were. These values and groups change when the authors change the

	criteria for a minimum number of patients in a cluster (30 versus 1000). The stated objective in the abstract “2) propose a methodology to identify sub-populations of interest for both clinical and financial interventions” hasn’t really been achieved since the link between these “interventions” is not clear. Restating this as identifying sub-populations based on diagnosis groups and high costs would be more accurate. The authors claim that this method is novel and more intuitive and actionable compared to other clustering methods. This critique of other methods in the introduction, however (page 5 line 7), is based on a discussion piece in 2001, which I don’t think reflects the current literature on the utility of these methods for MCC. The start of this paragraph “It remains difficult to compare and contrast the clinical and financial reforms enacted in different patient populations” does not really make sense (what has been enacted?), and is not supported by the rest of the paragraph. It’s not clear that this a novel method, or the justification for it. Reference 24 used similar methods for detecting multimorbidity (although they used triads of diseases). Why were pairs of diseases appropriate to use here? Is this really the development of a methodology (as stated page 5, line 20, and again page 5 line 38?). Abstract (line 17, page 2): it’s not clear that the clusters were defined iteratively – this would imply that there was some updated information or cycle of operations that adjusts the cluster definitions. Could the authors clarify what this step is? It appears all the clusters were predefined based on diagnosis pairs, age and sex (line 4 page 6). One of the age groups given in page 6 line 5 (methods section) is 0-18 years, but in the abstract the authors state this is an adult population over 18 years. The tests/p-values in Table 1 have not been explained in the methods and results – is this comparing across the number of chronic conditions, with years combined or across all three years? Methods, line 8 page 7: “Chi-squared tests were used to analyze differences in frequency between cluster groupings.” I cannot see where these were reported in the results. This might be the tests in Table 1? If so, this is confusing, because the authors have defined the “clusters” as referring to pairs of diagnoses. Results, line 30 page 8: “Significant disparities are observed between boroughs” implies that a statistical test was carried out, which I don’t think is the case here.

VERSION 1 – AUTHOR RESPONSE

Response to Reviewer 1:

Interesting study. wondering about female over male prevalence of two or three chronic conditions, but possibly due to females being more likely to seek care. Overall very useful and well-written study

Thank you for your comments.

Response to Reviewer 2:

The authors present a comprehensive analysis of co-occuring chronic diseases in a large urban health system. They use an interesting novel method of clustering patients by pairs of co-occuring diseases.

I feel this is not adequately expressed in the title: I would suggest to change the title, e.g. "Multiple chronic conditions at a major urban health system:

a descriptive analysis of frequencies, costs and co-occurence patterns".

We agree that the title could better reflect the manuscript content. We have changed it accordingly, to “Multiple chronic conditions at a major urban health system: a descriptive retrospective analysis of frequencies, costs and comorbidity patterns.”

The authors provide a long list of references (12 citations, 3-14) for one claim. Most of these references are between 10 and over 25 years old. I would question their relevance to costs in today's healthcare setting. I would propose to omit all but the mose recent references.

Thank you for noting the dated references. We have omitted the below references, but kept in the others above:

- Crystal S, Johnson RW, Harman J, et al. Out-of-pocket health care costs among older Americans. J Gerontol B Psychol Sci Soc Sci 2000;55:S51–62.
- Fishman P, Von Korff M, Lozano P, et al. Chronic care costs in managed care. Health Aff 1997;16:239–47.
- Moxey ED, O’Connor JP, Novielli KD, et al. Prescription drug use in the elderly: a descriptive analysis. Health Care Financ Rev 2003;24:127–41.
- Rice DP, LaPlante MP. Medical expenditures for disability and disabling comorbidity. Am J Public Health 1992;82:739–41.

Additionnal small remarks: Page 7, line 18: "lives with" instead of "suffers from" [two or more chronic conditions].

We agree that the prior language is stigmatizing, and have made this change throughout the document: see page 7 line 18, but also page 4, line 13, and page 13, line 15.

Table 3: add total yearly cost or number of individuals in cluster

We have added these data as the right-most three columns of Table 3. As detailed in the Results of the manuscript, these results demonstrate

Response to Reviewer 3:

This article revealed that the prevalence of multiple chronic conditions (MCC) in Medicaid population across selected boroughs in New York City is higher than the national wide in US, may cause the high expenditure of health care finance and poor health care quality, is an evidence of the challenges for clinical care system and financial utilization.

Although the top clusters of MCC by age and gender segments with yearly cost in supplementary Table 1 has been shown, cause the prevalence of MCC among each boroughs in New York City is different also be revealed in Figure 2, the variable of spatial distribution combined with gender and age could provide more describing power by using with this proposed clustering methodology is suggested.

Thank you for the comment. Although we agree it may be instructive to understand how the leading clusters within each age-gender cohort differ by borough of New York City, the small data set outside of Mount Sinai hospitals hampered this analysis: most patients lived in only 10 of 176 postal (ZIP) codes in New York City. We acknowledge this limitation explicitly in the Discussion section, with revised language as follows (new language in bold):

“The generalizability of our analysis is limited by the geospatial distribution of patients in the study population -- because provider attribution is accomplished regionally, our data set includes the subset of New York City patients who live near Mount Sinai practices. As a result, in the current data set, the majority of patients are located in just 10 of 176 ZIP codes. Future analyses using a data set such as an all-payer claims database would allow researchers to define clusters by region and ZIP code.”

Response to Reviewer 4:

It would be useful to describe briefly in the abstract conclusion section the national comparable statistic (42% of the national population with 2 or more chronic conditions as opposed to 52.7%, and also perhaps the range within the 5 boroughs).

Thank you for the suggestion – we agree that comparing our data to United States national data adds context.. We have included the 42% figure in the conclusion of the abstract, as follows:

“In this low-income, urban population, multiple chronic conditions are more prevalent (61%) than nationally (42%), motivating further research and implementation efforts in this population.”

The discussion section might benefit from a discussion around the public health and prevention interventions that might benefit identified populations where they were previously unknown, and the tool's potential beneficial use by public authorities to reduce health inequalities and improve health outcomes.

Thank you for this suggestion. We have added examples to illustrate the utility of this tool for measuring population risk in order to design novel health interventions. We describe a novel hospitalization-at-home program at Mount Sinai, whose ongoing expansion in both geographic coverage and scope of care will benefit from timely local identification of chronic disease clusters (see fourth paragraph, page 14). We also note the implications of this work for integrated chronic disease control programs in low- and middle-income countries (same paragraph) and for helping healthcare payors to create and adjust new reimbursement models (final paragraph, page 15). The full paragraph is as follows:

“Our findings apply not only to the reform of existing programs for low-income and vulnerable populations, but also the design of novel ones, in the Mount Sinai system and beyond. For example, Mount Sinai offers Healthfirst (and other) patients who require inpatient-level care an alternative: a Hospitalization-at-Home (HaH) program in lieu of inpatient admission [46,47]. Evaluation to date demonstrates that this HaH program delivers superior patient outcomes (including shorter length of stay) and greater patient satisfaction than in-hospital care, though costs have not yet been compared [46]. The HaH program focused on only nine diagnoses at its founding in 2014, but has since expanded in size and breadth of care across multiple New York hospitals, treating myriad other conditions across eight domains of care, such as post-surgical care, palliative care, and sub-acute rehabilitation, among others [47]. Rapid and timely data on the prevalence and overlap of these (largely chronic) diseases and their risk factors will be instrumental to the program’s ongoing cost-effective scale-up. Such data could prove even more valuable in low- and middle- income countries, where the burden of chronic disease is rapidly expanding, but models for the integrated care of more than one chronic condition are few and small in scope [48].”

Response to Reviewer 5:

Thank you to the authors for their article and their approach in investigating a complex problem and claims data set. The authors have investigated a set of Medicaid claims from a population in New York (inpatient and outpatient claims). They aimed to describe this population based on groupings of patients using combinations of chronic condition diagnoses, the gender and age group of the patients. These groups are then ranked by the total cost (of medical services), frequency and how unexpected (observed/expected) the co-occurring chronic conditions were. These values and groups change when the authors change the criteria for a minimum number of patients in a cluster (30 versus 1000).

The stated objective in the abstract “2) propose a methodology to identify sub-populations of interest for both clinical and financial interventions” hasn’t really been achieved since the link between these “interventions” is not clear. Restating this as identifying sub-populations based on diagnosis groups and high costs would be more accurate.

Thank you. We have made this change to be more accurate, as follows (change in bold):

“To (1) examine the burden of multiple chronic conditions (MCC) in an urban health system, and (2) propose a methodology to identify sub-populations of interest based on diagnosis groups and costs.”

The authors claim that this method is novel and more intuitive and actionable compared to other clustering methods. This critique of other methods in the introduction, however (page 5 line 7), is based on a discussion piece in 2001, which I don’t think reflects the current literature on the utility of these methods for MCC. The start of this paragraph “It remains difficult to compare and contrast the clinical and financial reforms enacted in different patient populations” does not really make sense (what has been enacted?), and is not supported by the rest of the paragraph.

It’s not clear that this a novel method, or the justification for it. Reference 24 used similar methods for detecting multimorbidity (although they used triads of diseases). Why were pairs of diseases appropriate to use here? Is this really the development of a methodology (as stated page 5, line 20, and again page 5 line 38?).

Thank you for this feedback. We have edited the manuscript to avoid overstating the novelty of our approach, and added references to justify the aspects we feel are innovative. We believe that our approach provides a simple descriptive method for large datasets that use health insurance claims - and we therefore focused our initial work on low-income patients as part of Medicaid insurance:

“While there exist numerous sophisticated statistical methods for clustering populations of patients - such as random forests, single decision trees, k-means, and hierarchical cluster analysis - these methods suffer from limited interpretability, result instability, immense computing overhead and/or tendency for overfitting (Breiman 2001; Nicholson et al. 2017; Ng et al. 2018). Rather than relying on complex statistical models that require significant computing overhead, we propose a simple descriptive method that can be applied to any population for whom medical claims are available.

Because its requisites are computationally simple, this methodology can be easily scaled to larger populations."

We chose to use pairs (as well as triads) of diseases given prior research from US Health and Human Services defines multi-morbid states as including two or more chronic conditions (dyads) . Although other definitions of multi-morbidity, such as Guys & St Thomas Charity, use a definition of three or more chronic conditions (triads) , we felt that with our current approach this presented a prohibitively large number of clusters to explore (419,152 as opposed to 18,768). We intend to expand the analysis to present further combinations as our work deepens.

Abstract (line 17, page 2): it's not clear that the clusters were defined iteratively – this would imply that there was some updated information or cycle of operations that adjusts the cluster definitions. Could the authors clarify what this step is? It appears all the clusters were predefined based on diagnosis pairs, age and sex (line 4 page 6).

Thank you for your observation – we agree that the word 'iteratively' implies the iterative inclusion or exclusion of certain variables (as in a regression model) – and our study did not approach exploring patient clusters iteratively, in this sense. We have updated that section to use the word 'combinatorically' instead, which more accurately describes the method of exploring a large number of pre-defined combinations of disease.

One of the age groups given in page 6 line 5 (methods section) is 0-18 years, but in the abstract the authors state this is an adult population over 18 years.

Thank you for bringing this error to our attention. Our authorship team re-ran analyses to ensure consistent exclusion of those under 18 years of age. These updated variables are reflected in the results and discussion sections of the manuscript, as well as the relevant tables and figures.

The tests/p-values in Table 1 have not been explained in the methods and results – is this comparing across the number of chronic conditions, with years combined or across all three years?

Thank you for noting this omission. We initially generated these p-values comparing between number of chronic conditions, but then noted that the inclusion of the "0 chronic conditions" column would make all of these comparisons statistically significant, defeating the purpose of the test. We have thereforenow removed all of the p-values from the table.

Methods, line 8 page 7: "Chi-squared tests were used to analyze differences in frequency between cluster groupings." I cannot see where these were reported in the results. This might be the tests in Table 1? If so, this is confusing, because the authors have defined the "clusters" as referring to pairs of diagnoses.

Thank you for noting this ambiguity. As justified above, we have removed the chi-squared tests from the methods section and do not report any p-values in our analyses.

Results, line 30 page 8: “Significant disparities are observed between boroughs” implies that a statistical test was carried out, which I don’t think is the case here.

Thank you for this observation - you are correct, we do not perform any statistical testing and so have removed this language from the manuscript.

VERSION 2 – REVIEW

REVIEWER	Kelsey Chalmers The Menzies Centre for Health Policy, Sydney School of Public Health, The University of Sydney. Australia.
REVIEW RETURNED	20-Aug-2019

GENERAL COMMENTS	Thank you to the authors for their response to my previous review, and for correcting their analysis. On rereading the article, I have some additional comments. Since this is the second round of reviews (apologies for not pointing out these in the first round), these can be addressed at the editors/authors discretion. The word ‘clusters’ usually refers to data points that have some natural grouping, so the use of the word here to describe set definitions of combinations of diseases is a bit jarring. I would suggest changing the description of the groups to ‘combinations’ or similar. In Table 2 – are the shown chronic condition groups the top by membership volume, or selected in some other way? The impact of the ‘minimum threshold’ on the types of diseases in clusters and the highest cost groups has not really been included in the discussion, and shows some of the potential problems providers might have in using this in the ‘real world’ as the authors propose – how should they choose what their threshold should be? Page 10, “Table 3 shows all clusters segmented”… I don’t think this is all clusters, just the top ten.

VERSION 2 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 5

Reviewer Name: Kelsey Chalmers

Institution and Country: The Menzies Centre for Health Policy, Sydney School of Public Health, The University of Sydney. Australia.

Please state any competing interests or state ‘None declared’: None declared.

Please leave your comments for the authors below

Thank you to the authors for their response to my previous review, and for correcting their analysis. On rereading the article, I have some additional comments. Since this is the second round of reviews (apologies for not pointing out these in the first round), these can be addressed at the editors/authors discretion.

The word ‘clusters’ usually refers to data points that have some natural grouping, so the use of the word here to describe set definitions of combinations of diseases is a bit jarring. I would suggest changing the description of the groups to ‘combinations’ or similar.

Thank you for pointing this out. Our research team was also concerned about disambiguating our use of the world, so we appreciate the opportunity to clarify. We have chosen to change the word ‘cluster’ to the word ‘segment’.

In Table 2 – are the shown chronic condition groups the top by membership volume, or selected in some other way?

These groups are the top 10 by age-adjusted frequency. We have updated the manuscript to reflect this change.

The impact of the ‘minimum threshold’ on the types of diseases in clusters and the highest cost groups has not really been included in the discussion, and shows some of the potential problems providers might have in using this in the ‘real world’ as the authors propose – how should they choose what their threshold should be?

Thank you for pointing this out. We have updated the discussion to include the below commentary clarifying the utility and purpose behind selecting different thresholds:

It is clear that the threshold itself - small, medium or large - for the volume of patients to analyze can be modified with effect on the resultant segments. While senior executives and health services

analysts in population health may be interested in overall patterns, costs and adjusted risk of co-morbidity, specialty service lines may be focused more on tailored, smaller patient segments with unique disease patterns requiring integrated care. For example, the development of a value-based healthcare program in the US Navy involved the creation of integrated practice units to treat low back pain and osteoarthritis [45]. Our analysis across multiple thresholds animates how the thresholds can affect the resultant patterns produced.

Page 10, “Table 3 shows all clusters segmented”… I don’t think this is all clusters, just the top ten.

That is correct, we have corrected the text to: Table 3 shows

the top 10 segments including age…”

Typographic Changes:

In our careful review of the paper and analysis code, we found that we had accidentally mis-labeled the Y-axis of Figure 2 to state “Percentage with >=1 Chronic Conditions”. After verifying that the value plotted was actually the % of patients with 2 or more chronic conditions (concordant with the definition of MCC we utilized in the paper), we have amended the Y-axis label to “Percentage with >=2 Chronic Conditions”.