PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	A systematic review of high-cost patients' characteristics and healthcare utilization
AUTHORS	Wammes, Joost; Van der Wees, Philip; Tanke, Marit; Westert, Gert; Jeurissen, Patrick

VERSION 1 – REVIEW

REVIEWER	Karthik Srinivasan University of Arizona Tucson
REVIEW RETURNED	11-Apr-2018

GENERAL COMMENTS	The paper is well-written and does a thorough job in summarizing the key findings from the 46 studies, screened from 7905 journal articles which included the terms 'high-cost', 'patients', and 'cost' and 'cost analysis' in the titles/abstracts of the respective articles (or including body?). The authors use the Andersen's behavioral model to categorize the characteristics of high-cost patients in the 46 studies into three categories - 'predisposing', 'enabling' and 'need' related factors. The authors conclude that high-cost patients are diverse populations and vary across payer types and countries, and hence tailored interventions are required. I congratulate the authors on conducting this study and highlighting important high-cost patient characteristics that may be relevant for clinical decision making. Table 2 of the paper indicates various demographic, disease related and admission related factors. The Pubmed keyword search strategy and the filtering strategies of the authors narrow down into 46 articles (Table 1) that primarily use descriptive and regression modeling-based research methodology. However, on the other spectrum of data analysis, there have been multiple data mining methodology-based studies on high-cost patient prediction that have identified important high-cost patient characterizing factors (often called as predictors in data mining terminology). Following are a few of the papers I was able to retrieve from PubMed that may be of interest to the authors (and which satisfy the author's filtering criteria): 1. Chechulin, Y., Nazerian, A., Rais, S., & Malikov, K. (2014). Predicting patients with high risk of becoming high-cost healthcare users in Ontario (Canada). Healthcare Policy. 2. Meenan, R. T., Goodman, M. J., Fishman, P. A., Hornbrook, M. C., O'Keeffe-Rosetti, M. C., & Bachman, D. J. (2003). Using risk-adjustment models to identify high-cost risks. Medical care. 3. Fleishman, J. A., & Cohen, J. W. (2010). Using information on clinical conditions to predict high-cost patients. Health services research.

	4. Shenas, S. A. I., Raahemi, B., Tekieh, M. H., & Kuziemsky, C. (2014). Identifying high-cost patients using data mining techniques and a small set of non-trivial attributes. Computers in biology and medicine. Such a category of data mining-based studies may exclude papers not listed in Pubmed such as Bertsimas et al. (2008)^1, or ones which are purely methodological and not providing clinical insights on high-cost patient characterization such as Maidman and Wang (2017)^2. Authors are requested to revisit the search stage and also include such studies in addition to descriptive and regression modeling-based studies. 1) Bertsimas, D., Bjarnadóttir, M. V., Kane, M. A., Kryder, J. C., Pandey, R., Vempala, S., & Wang, G. (2008). Algorithmic prediction of health-care costs. Operations Research 2) Maidman, A., & Wang, L. (2017). New semiparametric method for predicting high-cost patients. Biometrics.

REVIEWER	Bin Xie Parkland Center for Clinical Innovation, USA
REVIEW RETURNED	20-Apr-2018

GENERAL COMMENTS	This manuscript discussed an important topic that is of increasing importance to healthcare practitioners, policy makers, and the general public. It is well written and both the methodology and findings are clearly described. The authors acknowledged some of the limitations. One point that I believe has not received sufficient attention from the authors is the heavy concentrations of studies published in North America. Of the 46 studies included in the final analysis, only 3 were from outside of North America. While this is understandable and is a result of the authors' reasonable methodology, I believe the authors need to discuss the implications of such a fact. They can either re-orient the paper as a systematic review of high-cost patients in North America (removing the 3 non-NA studies), or maybe discuss some contrast of the findings in these two different groups of studies, or in some other way discuss this issue and how it affect their findings.

REVIEWER	Siddharth Singh University of California San Diego, La Jolla, California, USA
REVIEW RETURNED	23-Apr-2018

GENERAL COMMENTS	In this systematic review, Wammes and colleagues evaluated the distribution, characteristics, healthcare service utilization and drivers of high-cost patients, across health systems. The topic is meritorious as healthcare expenses balloon, and focus is shifting towards value-based care and population health. The study is well-designed and conducted and largely follows recommended reporting criteria such as PRISMA, though none are reported. The strengths of the study are: broad scope answering all pertinent questions for this population; use of Anderson behavioral model which allows categorization of attributes of high-cost patients into predisposing, enabling and needs factors; and evaluation of patterns of use across health systems, globally (though North America is over-represented).

	There are some weaknesses, inherent to individual studies, and some in the systematic review. These are highlighted below for the authors' consideration: From a systematic review perspective, 1. There seems to be over-reliance on a single person in study selection and data abstraction. While this is ideally not advisable, it may be an appropriate strategy provided checks and balances are in place. What was the qualification and expertise of the said 'Author A' for leading this effort? Where Authors B and C performed random assessment, what were the observations - kappa between agreements at different steps, reconciliatory measures, etc.? 2. The authors have not performed a risk of bias assessment for individual studies or for the body of evidence, citing differences in study design. While I acknowledge challenges in assessing, it is still a critical function of any knowledge synthesis to ascertain the underlying quality of literature. For example, how do we know the veracity of the fact that included studies were adept and accurate in truly identifying the top 5% or 10% of their cohort by expenses? In assessing risk factors, inability to adjust for confounders is vital since a lot of these factors likely interact (age, multimorbidity, etc.) I suggest the authors identify key factors required to ascertain the risk of bias in this literature, perform the said assessment, and discuss in manuscript, to help the reader understand the limitations of the literature better. 3. I'm not sure of the relevance of some older studies conducted in the 1980s. Those cost reflections are likely not reflective, and they are not assessed in terms of time-trends in high-cost patients. Similarly, some included studies do not appear to be representative (one key quality metric), for example, a study which included "all patients seen by a PCP", or "all employees of a bank". By refining inclusion criteria, these may be re-considered. 4. The authors have nicely attempted to meaningfully capture and synthesize data for all risk factors/characteristics, etc. from such a diverse group of studies. However, in the process, there seems to be subjectivity and selective presentation of 'positive' findings, wherein a few studies seem to be contributing the bulk of the 'evidence'. Moreover, qualitative statements are made on potential impact of a risk factor, without quantitative assessment - was the contribution of a risk factor, small vs. large, etc. While I do not have a great suggestion for how to overcome it, one approach may be to summarize key findings/results objectively from each study separately for an interested reader in a supplementary table. From a Discussion standpoint, - I believe this study has more implications for population health management, rather than 'quality improvement' as mentioned in Introduction. The authors may wish to expand on concept of population health management in the text - I'm not sure these 'very costly' patients are presented as 'very sick' in individual studies or by the review, yet the title seems to imply so. Besides multimorbidity assessed at a qualitative level, there is no objective assessment of patients being 'medically sick'. May be the review focuses only on high-cost patients, or perhaps high-need, high-cost patients, but not necessarily very sick patients - The Policy and Research implications sections seems to be disjointed from rest of the paper - the 4 key 'observations' that the authors inferred from the literature, are not apparent in the 'Results' section of the paper. I'm unsure how the authors arrived at those 4 observations.

	- The Abstract could be a little more quantitative and objective - for example, how much do the top 5% and 10% patients across health systems contribute, what proportion are generally deemed preventable; what is main contributor to cost in these patients; what are consistently modifiable and non-modifiable predisposing/enabling and needs factors

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Karthik Srinivasan
Institution and Country: INSITE: Center for Business Intelligence and Analytics, Department of Management Information Systems, Eller College of Management, University of Arizona, Tucson, AZ, USA Competing Interests: None declared

The paper is well-written and does a thorough job in summarizing the key findings from the 46 studies, screened from 7905 journal articles which included the terms 'high-cost', 'patients', and 'cost' and 'cost analysis' in the titles/abstracts of the respective articles (or including body?). The authors use the Andersen's behavioral model to categorize the characteristics of high-cost patients in the 46 studies into three categories - 'predisposing', 'enabling' and 'need' related factors. The authors conclude that high-cost patients are diverse populations and vary across payer types and countries, and hence tailored interventions are required. I congratulate the authors on conducting this study and highlighting important high-cost patient characteristics that may be relevant for clinical decision making.

Table 2 of the paper indicates various demographic, disease related and admission related factors. The Pubmed keyword search strategy and the filtering strategies of the authors narrow down into 46 articles (Table 1) that primarily use descriptive and regression modeling-based research methodology. However, on the other spectrum of data analysis, there have been multiple data mining methodology-based studies on high-cost patient prediction that have identified important high-cost patient characterizing factors (often called as predictors in data mining terminology).

Following are a few of the papers I was able to retrieve from PubMed that may be of interest to the authors (and which satisfy the author's filtering criteria):

1. Chechulin, Y., Nazerian, A., Rais, S., & Malikov, K. (2014). Predicting patients with high risk of becoming high-cost healthcare users in Ontario (Canada). Healthcare Policy.
2. Meenan, R. T., Goodman, M. J., Fishman, P. A., Hornbrook, M. C., O'Keeffe-Rosetti, M. C., & Bachman, D. J. (2003). Using risk-adjustment models to identify high-cost risks. Medical care.
3. Fleishman, J. A., & Cohen, J. W. (2010). Using information on clinical conditions to predict high-cost patients. Health services research.
4. Shenas, S. A. I., Raahemi, B., Tekieh, M. H., & Kuziemsky, C. (2014). Identifying high-cost patients using data mining techniques and a small set of non-trivial attributes. Computers in biology and medicine.

Such a category of data mining-based studies may exclude papers not listed in Pubmed such as Bertsimas et al. (2008)^1, or ones which are purely methodological and not providing clinical insights on high-cost patient characterization such as Maidman and Wang (2017)^2. Authors are requested to revisit the search stage and also include such studies in addition to descriptive and regression modeling-based studies.

1) Bertsimas, D., Bjarnadóttir, M. V., Kane, M. A., Kryder, J. C., Pandey, R., Vempala, S., & Wang, G.

(2008). Algorithmic prediction of health-care costs. Operations Research

2) Maidman, A., & Wang, L. (2017). New semiparametric method for predicting high-cost patients. Biometrics.

>> We agree with the reviewer that such predictive studies might identify important high-cost patient characterizing factors. As suggested, we revisited our earlier search stages and identified twenty-one of such articles, and studied these thoroughly. Twelve of those articles did include evaluations of individual predictors for high costs and were informative for our review. These studies were included in our review, and our analysis was updated accordingly. The remaining studies did not provide any insight in the characteristics and utilization of high-cost patients.

Reviewer: 2
Reviewer Name: Bin Xie
Institution and Country: Parkland Center for Clinical Innovation, USA Competing Interests: None declared

This manuscript discussed an important topic that is of increasing importance to healthcare practitioners, policy makers, and the general public. It is well written and both the methodology and findings are clearly described.

The authors acknowledged some of the limitations. One point that I believe has not received sufficient attention from the authors is the heavy concentrations of studies published in North America. Of the 46 studies included in the final analysis, only 3 were from outside of North America. While this is understandable and is a result of the authors' reasonable methodology, I believe the authors need to discuss the implications of such a fact. They can either re-orient the paper as a systematic review of high-cost patients in North America (removing the 3 non-NA studies), or maybe discuss some contrast of the findings in these two different groups of studies, or in some other way discuss this issue and how it affect their findings.

>> We agree with the reviewer that North-American studies are overrepresented and that this may affect our findings. We have discussed this issue in our research team, and we are convinced of the added value of the non-North-American studies. Our review's aim was to study high-cost patients from an international perspective. One of our findings was that most studies were performed in North-America, and we consider this a new and valuable insight from our work.
In our discussion section we cross the point of overrepresentation of North-American studies, and noted that this limits the generalisability of our findings.
We performed additional analyses to compare Dutch and Danish high-cost patients with high-cost patients in other total population studies. Based on these analyses, we added the following phrases to the results section and to the discussion section:
- "Besides, the prevalence of each of the chronic diseases in the Dutch study was comparable with the prevalence in other total population studies.".
- "The mortality among Danish and Dutch high-cost patients was comparable with the mortality in other total population studies.".
- "Although our comparison across countries did not reveal large differences in mortality or prevalence of common chronic diseases, these analyses were based on a limited number of variables, studies and countries. It is likely that the specific characteristics and utilization of high-cost patients vary across localizations due to a wide range of epidemiological and health system factors.".

Reviewer: 3

Reviewer Name: Siddharth Singh
Institution and Country: University of California San Diego, La Jolla, California, USA Competing Interests: None declared

In this systematic review, Wammes and colleagues evaluated the distribution, characteristics, healthcare service utilization and drivers of high-cost patients, across health systems.

The topic is meritorious as healthcare expenses balloon, and focus is shifting towards value-based care and population health. The study is well-designed and conducted and largely follows recommended reporting criteria such as PRISMA, though none are reported. The strengths of the study are: broad scope answering all pertinent questions for this population; use of Anderson behavioral model which allows categorization of attributes of high-cost patients into predisposing, enabling and needs factors; and evaluation of patterns of use across health systems, globally (though North America is over-represented).

There are some weaknesses, inherent to individual studies, and some in the systematic review. These are highlighted below for the authors' consideration:
From a systematic review perspective,
1. There seems to be over-reliance on a single person in study selection and data abstraction. While this is ideally not advisable, it may be an appropriate strategy provided checks and balances are in place. What was the qualification and expertise of the said 'Author A' for leading this effort? Where Authors B and C performed random assessment, what were the observations - kappa between agreements at different steps, reconciliatory measures, etc.?

>> Author A is a PhD-candidate with over six years experience in health services research, and considerable experience in literature studies. Author B and C are both senior-researchers and both performed a range of systematic reviews. At each round of our study selection process and data extraction we in-depth discussed (in-)consistencies until we reached consensus. Based on these consensus meetings, we refined our criteria and data extraction form, and prior work was repeated. Because of this iterative process, no kappa agreement was calculated. We added the following phrases to the manuscript:
- “At each step of this selection process, (in-)consistencies were discussed until consensus was reached. On basis of the discussions, the criteria were refined and the prior selection process was repeated.”.
- “On basis of this discussion, the data extraction form was refined and the prior data extraction was repeated.”.

2. The authors have not performed a risk of bias assessment for individual studies or for the body of evidence, citing differences in study design. While I acknowledge challenges in assessing, it is still a critical function of any knowledge synthesis to ascertain the underlying quality of literature. For example, how do we know the veracity of the fact that included studies were adept and accurate in truly identifying the top 5% or 10% of their cohort by expenses? In assessing risk factors, inability to adjust for confounders is vital since a lot of these factors likely interact (age, multimorbidity, etc.) I suggest the authors identify key factors required to ascertain the risk of bias in this literature, perform the said assessment, and discuss in manuscript, to help the reader understand the limitations of the literature better.

>> We agree with the reviewer that a risk of bias assessment is a critical function of knowledge synthesis. However, to our knowledge, no agreed upon framework exists for risk of bias assessment of the kind of studies included in our review.
One limitation in current frameworks for observation/cross-sectional studies is that these are primarily designed for studies that aim to assess intervention effects in comparative (retrospective) studies.

Besides, quality assessment frameworks exists for observation/cross-sectional studies, but such frameworks are not specifically designed to assess the risk of bias in the studies. To develop a new risk assessment framework for this particular issue might be risky, and might falsely credit or discredit the validity of included studies.
In our approach, we built on previously conducted research that is similar to ours (by Lehnert et al 2013^a), and Babitsch et al 2012^b) that have neither conducted risk assessments for the same reason.
To our opinion, the validity of the findings of each of included studies is contingent upon three main elements, including: internal validity: 1) analytical approach: how is dealt with possible confounding? Which explanatory variables and confounders were included in the modelling approaches? One problem in our included studies is that few used a comprehensive set of variables to study high-cost patients, and that no consensus exists about what factors should reasonably be controlled for; external validity: 2) the breadth of the population studied, and 3) the scope of costs included to establish total costs. All but one of these factors have been elaborately described in our manuscript. In the limitations section, we highlighted the possibility of confounding by design. We further elaborated on this issue:
“To our knowledge, no agreed upon framework exists for risk of bias assessment of the kind of studies included in our review. One limitation in current frameworks for observation/cross-sectional studies is that these are primarily designed for studies that aim to assess intervention effects in comparative studies. The internal validity of the findings in our included studies is mainly contingent upon its ability to control for relevant confounders. However, no consensus exists about what factors should reasonably be controlled for. The external validity of the findings of each of the studies depend upon the breadth of the population studied, and the scope of the costs considered for establishing total costs. Our study selection process was aimed at identifying studies with a broad population studied, and a wide range of costs considered.”

a. Review: health care utilization and costs of elderly persons with multiple chronic conditions. Med Care Res Rev. 2011 Aug;68(4):387-420. doi: 10.1177/1077558711399580. https://www.ncbi.nlm.nih.gov/pubmed/21813576
b. Re-revisiting Andersen's Behavioral Model of Health Services Use: a systematic review of studies from 1998-2011. Psychosoc Med. 2012;9:Doc11. doi: 10.3205/psm000089. https://www.ncbi.nlm.nih.gov/pubmed/23133505

3. I'm not sure of the relevance of some older studies conducted in the 1980s. Those cost reflections are likely not reflective, and they are not assessed in terms of time-trends in high-cost patients. Similarly, some included studies do not appear to be representative (one key quality metric), for example, a study which included "all patients seen by a PCP", or "all employees of a bank". By refining inclusion criteria, these may be re-considered.

>> We refined our inclusion criteria and removed all studies published prior 2000, and updated our analysis accordingly. We added the following phrase to our method section:
“and studies published in 2000 and later [were included.]”.

4. The authors have nicely attempted to meaningfully capture and synthesize data for all risk factors/characteristics, etc. from such a diverse group of studies. However, in the process, there seems to be subjectivity and selective presentation of 'positive' findings, wherein a few studies seem to be contributing the bulk of the 'evidence'. Moreover, qualitative statements are made on potential impact of a risk factor, without quantitative assessment - was the contribution of a risk factor, small vs. large, etc. While I do not have a great suggestion for how to overcome it, one approach may be to summarize key findings/results objectively from each study separately for an interested reader in a supplementary table.

>> Included studies rarely evaluated the relative contribution of each determinant for high costs, and the diversity of studies further complicated our efforts in gaining quantitative insight in the importance of individual factors. We agree with the reviewer that a summary of key findings may be highly informative for interested readers. We developed such an appendix accordingly, and referred to this in our manuscript in the method section:
“We also made a narrative summary of the findings per article (provided in appendix 2).”.

From a Discussion standpoint,
- I believe this study has more implications for population health management, rather than 'quality improvement' as mentioned in Introduction. The authors may wish to expand on concept of population health management in the text

>> We agree with the reviewer that our study has several implications for population health management. We firstly address this through our argument for segmentation analysis, which may be pivotal in population health management. In addition, population health management has proven to be beneficial for high-cost patients, especially for socially disadvantaged patients with multiple chronic morbidities and persistently high utilization. We added the following phrases to the discussion:
- “Such segmentation analysis may powerfully inform population health management initiatives.”.
- “Especially population health management approaches may be beneficial for these populations. Sherry et al. recently examined five community-oriented programs that successfully improved care for high-need, high-cost patients. The five programs shared common attributes, including a ‘whole person’ orientation, shared leadership, flexible financing and shared cross-system governance structures[85].”.

- I'm not sure these 'very costly' patients are presented as 'very sick' in individual studies or by the review, yet the title seems to imply so. Besides multimorbidity assessed at a qualitative level, there is no objective assessment of patients being 'medically sick'. May be the review focuses only on high-cost patients, or perhaps high-need, high-cost patients, but not necessarily very sick patients

>> We removed ‘Very sick’ from the title. We agree with the reviewer that the measure for multimorbidity was our main argument for choosing the title. However, a handful of studies also measured health status, and these studies showed high-cost patients were sick by every measure used.

- The Policy and Research implications sections seems to be disjointed from rest of the paper - the 4 key 'observations' that the authors inferred from the literature, are not apparent in the 'Results' section of the paper. I'm unsure how the authors arrived at those 4 observations.

>> We agree with the reviewer that the result section does not directly translate into the four major groups that we describe under ‘Policy and research implications’. This has to do with our breakdown of results according to the categories by Andersen. We based our four subgroups on themes that were reported most frequently (persistent vs incidental high costs, mental care), themes show most heterogeneity across settings (mental care, last year of life), and themes that require divergent policy approaches (all four subgroups). Aldridge et al. presented a similar breakdown of patient groups in high-cost populations, but did not identify mental health users as a subgroup, which we believe is highly justified given our findings. We rephrased the first sentence of the policy and research implications section:
- “Based on our findings, we deduced four major segments of high-cost patients for which separate policy may be warranted, including patients in their last year of life, patients experiencing a significant health event who return to stable health (episodically high-cost patients), patients with mental illness, and patients with persistently high costs characterized by chronic conditions, functional limitations and elder age.”

- The Abstract could be a little more quantitative and objective - for example, how much do the top 5% and 10% patients across health systems contribute, what proportion are generally deemed preventable; what is main contributor to cost in these patients; what are consistently modifiable and non-modifiable predisposing/enabling and needs factors

>> We rephrased the result section of our abstract accordingly:
- “Results: The studies pointed to a high prevalence of multiple (chronic) conditions to explain high-cost patients’ utilization. Besides, we found a high prevalence of mental illness across all studies, and a prevalence higher than 30% in US Medicaid and total population studies. Furthermore, we found that high costs were associated with increasing age, but that still more than halve of high-cost patients were younger than 65. High costs were associated with higher incomes in the US, but with lower incomes elsewhere. Preventable spending was estimated at maximally ten percent of spending. The top-10%, top-5% and top-1% high-cost patients accounted for respectively 68%, 55%, and 24% of costs within a given year. Spending persistency varied between 24% and 48%. Finally, we found that no more than 30% of high-cost patients are in their last year of life.”.

FORMATTING AMENDMENTS (if any)
Required amendments will be listed here; please include these changes in your revised version:
- Kindly re-upload Figure 1 with at least 300 dpi resolution.

>> We did so accordingly.

- Please include Figure 1 legend at the end of your main manuscript.

>> We did so accordingly.

- Kindly re-upload Appendix in PDF format.

>> We did so accordingly.

- Please remove tables uploaded separately and place it inside your main document where it was first cited.

>> We did so accordingly.

- Patient and Public Involvement:
Authors must include a statement in the methods section of the manuscript under the sub-heading 'Patient and Public Involvement'.

This should provide a brief response to the following questions:

How was the development of the research question and outcome measures informed by patients’ priorities, experience, and preferences?
How did you involve patients in the design of this study?
Were patients involved in the recruitment to and conduct of the study?
How will the results be disseminated to study participants?
For randomised controlled trials, was the burden of the intervention assessed by patients themselves?
Patient advisers should also be thanked in the contributorship statement/acknowledgements.
If patients and or public were not involved please state this.

>> We included the following phrase to the method section:
- “Patients and or public were not involved in the conduct of this study”.

VERSION 2 – REVIEW

REVIEWER	Bin Xie Parkland Center for Clinical Innovation
REVIEW RETURNED	12-Jun-2018

GENERAL COMMENTS	I believe this revised version addressed the concerns of the reviewers and is ready for publication. Thanks!

REVIEWER	Siddharth Singh University of California San Diego, La Jolla, CA
REVIEW RETURNED	16-Jun-2018

GENERAL COMMENTS	Thank you for systematically addressing all my concerns. I do not have any other questions, and commend the authors on a comprehensive review.

REVIEWER	Karthik Srinivasan University of Arizona, USA
REVIEW RETURNED	18-Jun-2018

GENERAL COMMENTS	The authors have addressed my revision comments. I have no other concerns/comments and would recommend the revised version to be accepted for publication.