PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Multimorbidity and health care utilization among high-cost patients in the U.S. Veterans Affairs Health Care System
AUTHORS	Donna M. Zulman, Christine Pal Chee, Todd H. Wagner, Jean Yoon, Danielle M. Cohen, Tyson H. Holmes, Christine Ritchie,Steven M. Asch

VERSION 1 - REVIEW

REVIEWER	Susan Smith RCSI Dublin, Ireland
REVIEW RETURNED	13-Feb-2015

GENERAL COMMENTS	This is a clearly written and well-conducted study presenting a cross sectional analysis of 5.2 million patients from the VA system in the US. There have been several previous studies examining this issue but few have had such large numbers and such comprehensive healthcare utilization data. The focus on high cost patients is interesting and allows important distinctions between high cost patients with single conditions such as schizophrenia and those with multimorbidity. I would suggest the following minor revisions: • Need more setting information for an international readership. The VA system seems quite unique and there are issues around external validity due to the gender bias and the higher likelihood in veterans of conditions such as post traumatic stress disorder and other mental health conditions. There also seem to be a high proportion of high cost patients who were homeless during the year of the study (14%) • Abstract: While it is a group of patients the term cohort usually implies longitudinal data and this is really cross sectional data reflecting utilization over a one year period • Multimorbidity patterns – please provide some summary of the methods for coding condition triads rather than simply referring the reader to previous publications • Discussion: I think would be improved by more specific consideration of what this study tells us about who to target for interventions, while also acknowledging that the evidence base for interventions is limited

REVIEWER	Juan F Orueta Osakidetza (Basque National Health Service) Spain
REVIEW RETURNED	13-Feb-2015

GENERAL COMMENTS	The objectives of the paper are to describe the morbidity patterns observed in the highest cost patients and the relationship between

	multimorbidity and costs. Although there are a growing number of publications about multimorbidity, I think this is an interesting paper because it is specifically devoted to such subgroup of patients. It responds to the defined objectives ant its conclusions have implications for health policy and health services management. I have no major points and only I suggest some minor revisions. 1.- One important result of the study is that patients with very serious pathologies, such as cancer and schizophrenia, present less comorbid conditions than other patients. However, as the authors recognize in the discussion section, under-diagnosis (or under-coding) may have influenced such finding. Consequently, its inclusion in the Abstract needs qualifying (for example, patients with cancer and schizophrenia were less likely to have "registered" comorbid conditions) 2.- I am not an expert in statistics and I am not familiar with fractional logit regression models. For me, there is something confusing in results ("Association between multimorbidity and health care utilization and costs"; table 3; appendix: tables 3 and 4). The increase in absolute costs for additional chronic conditions is more pronounced in the outpatient than in inpatient ones, so it seems rational that the predicted share of outpatient costs rises as the inpatient share decreases. However, values are somehow different. I understand that predicted absolute costs and predicted shares are estimated by employing different statistical models but, for example, in table 3 predicted inpatient absolute costs are $38,901 for 1 body system (approx. 56% of total) and the predicted share is 48.2 while for outpatient values are, respectively, $17,560 (25%) and 30.7 . In my view, some explanation is needed to add clarity for non-expert readers. 3.- Regression to the mean is a well-known phenomenon that affects health care costs and other variables. The authors are aware of that because they accept that only a portion of high-cost patients (20-60%) will remain in such situation in the next year. However, in my opinion, they should avoid comments to this fact (in introduction and discussion): they are not describing characteristics of persistent high users, but only one year users. 4.- Other researchers have considered that a zero-inflated binomial negative (ZINB) regression is more suitable than a Poisson distribution to evaluate differences in health care utilization, because data are strongly overdispersed. In my opinion, the authors should include a comment, explaining their choice for Poisson distributions. 5.- There are some minor discrepancies between figures in Appendix table 1 and text (percentage of hypertension 63 – 65; mental health condition: 48 -47) 6.- There are not references about ethics approval . Also I have not found a STROBE checklist

VERSION 1 – AUTHOR RESPONSE

Reviewer Name Susan Smith, RCSI Dublin, Ireland

1. Need more setting information for an international readership. The VA system seems quite unique and there are issues around external validity due to the gender bias and the higher likelihood in veterans of conditions such as post traumatic stress disorder and other mental health conditions. There also seem to be a high proportion of high cost patients who were homeless during the year of the study (14%)

The reviewer is right that VA patients have higher rates of psychosocial problems and are more likely to be male (92%) than the U.S. population. However, the VA Health Care System is still an interesting place to examine the relationship between multimorbidity and utilization. First, it is worth studying in and of itself as it is America's largest integrated health care system and serves nearly six million veterans of the U.S. military per year. Moreover, VA patients are sicker in many ways than the average U.S. patient, thus increasing the variability in multimorbidity for our models to explain. The large number of mentally ill and homeless patients are typical of safety net institutions in the United States (and around the world), and these systems will no doubt be able to draw lessons from our work.

To address this point, we have added a brief description of the VA context for international readers:

"The VA Health Care System serves close to six million veterans of the U.S. military per year through 150 medical centers and 820 community based outpatient clinics. The VA is considered a leader in the delivery of high quality care in the U.S..11,12 Patients in the VA are predominantly male and typically have higher rates of physical and mental illness, and are poorer, than age-matched non-Veterans. In addition, younger Veterans (individuals deployed after 9/11) have high rates of physical and emotional trauma, including post-traumatic stress disorder. Patients are eligible not only for health benefits, but also for substantial social support including disability payments, pensions, educational benefits, housing support, and vocational rehabilitation." (page 4/para 2)

In addition, on page 15 (para 1), we acknowledge the limitations to the generalizability: "Finally, the relationships that we observed may be limited to the VA population, however the VA is a very large healthcare system and its patients' high rates of physical and psychosocial health problems are similar to those observed in safety net systems worldwide."

2. Abstract: While it is a group of patients the term cohort usually implies longitudinal data and this is really cross sectional data reflecting utilization over a one year period

We have removed the word cohort from the paper to prevent confusion, and have modified the Abstract so that it now reads:

"In this retrospective cross-sectional study of all patients in the United States' Veterans Affairs (VA) Health Care System, we aggregated costs of individuals' outpatient and inpatient care, pharmacy services, and VA-sponsored contract care received in 2010." (page 2)

3. Multimorbidity patterns – please provide some summary of the methods for coding condition triads rather than simply referring the reader to previous publications

We have modified the Methods to provide more detail about how we coded condition triads:

“We also coded chronic condition triads, as described in previously published studies of VA patient populations.17,19 All coded conditions were searched for all patients and all existing trios of comorbid conditions were identified. The quantity of high-cost patients with each of these identified trios was tabulated to determine the trios that were present in at least 5% of patients.” (page 7/para 1)

4. Discussion: I think would be improved by more specific consideration of what this study tells us about who to target for interventions, while also acknowledging that the evidence base for interventions is limited

The question of which high-cost patients should be targeted for interventions is a very important one, and one for which little evidence has surfaced to date. While many intensive management/complex care programs identify patients using risk prediction tools or criteria such as frequent hospitalizations, it is not clear that these approaches alone identify the patients who are most likely to benefit from the intervention. The reviewer is also correct that the evidence base for intensive management programs for high-cost patients remains limited. We have augmented the discussion to address some of these points as follows:

“Finally, this study highlights the heterogeneity of patients with high-costs, and suggests that a one-size-fits-all intervention may not be the optimal approach. The evidence base for intensive interventions is still limited, and important questions remain about how best to identify patients and match services to their needs. There may be individuals with stable chronic conditions and strong social support who are unlikely to benefit substantially from intensive management programs. Other patients may have highly specialized needs. For example, while the majority of high-cost patients in our study had multiple chronic conditions, we identified subgroups of patients with cancer and schizophrenia who have lower levels of multimorbidity. This finding is notable because intensive primary care programs that are focused on coordinating care for multiple chronic conditions may not be the best fit for patients whose care needs are dominated by these types of conditions. Health systems may instead choose to ensure that high-cost patients who have cancer or a serious mental illness but few comorbidities receive care within multidisciplinary programs that are focused on these diseases.” (page 13/para 3)

Reviewer Name Juan F Orueta, (Basque National Health Service), Spain

1. One important result of the study is that patients with very serious pathologies, such as cancer and schizophrenia, present less comorbid conditions than other patients. However, as the authors recognize in the discussion section, under-diagnosis (or under-coding) may have influenced such finding. Consequently, its inclusion in the Abstract needs qualifying (for example, patients with cancer and schizophrenia were less likely to have “registered” comorbid conditions)

We appreciate this comment and have modified the Abstract results section to qualify the finding, as suggested by the reviewer:

“Patients with cancer and schizophrenia were less likely to have documented comorbid conditions than other high-cost patients.” (page 2)

2. I am not an expert in statistics and I am not familiar with fractional logit regression models. For me, there is something confusing in results (“Association between multimorbidity and health care utilization and costs”; table 3; appendix: tables 3 and 4). The increase in absolute costs for additional chronic conditions is more pronounced in the outpatient than in inpatient ones, so it seems rational

that the predicted share of outpatient costs rises as the inpatient share decreases. However, values are somehow different. I understand that predicted absolute costs and predicted shares are estimated by employing different statistical models but, for example, in table 3 predicted inpatient absolute costs are $38,901 for 1 body system (approx. 56% of total) and the predicted share is 48.2 while for outpatient values are, respectively, $17,560 (25%) and 30.7 . In my view, some explanation is needed to add clarity for non-expert readers.

The reviewer has correctly noted a difference in the values produced by two different statistical approaches that are seeking to estimate the same underlying quantity. One of the estimates amounts to taking a ratio of averages: the average inpatient cost across patients divided by the average total cost across patients. The other estimate amounts to taking an average of ratios: the average across patients of the ratio of inpatient cost to total cost. These two estimation approaches need not produce the same numerical value (see TC Rao. Means of ratios or ratios of means or both? J of Statistical Planning and Inference. 2002:102;129-138). The two methods essentially amount to different implicit weighting schemes for the patients, with the mean of ratios more directly estimating share at the patient level (which is why we modeled these ratios using fractional logistic regression). The important point, however, is that both values are similar and produce the same qualitative interpretation. We have added a sentence of clarification in the footnotes of Table 2 and Appendix Tables 3 and 4:

"Note that dividing predicted component costs by predicted total costs does not equal the predicted share of total. The former is a ratio of means and the latter is a mean of ratios, and although both procedures estimate the same quantity, they are not guaranteed to produce the same result. The two procedures do, however, produce comparable results."

3. Regression to the mean is a well-known phenomenon that affects health care costs and other variables. The authors are aware of that because they accept that only a portion of high-cost patients (20-60%) will remain in such situation in the next year. However, in my opinion, they should avoid comments to this fact (in introduction and discussion): they are not describing characteristics of persistent high users, but only one year users.

We appreciate this point and have revised the introduction to make it clear that we are focusing on individuals with high costs in a single year:

"In order to inform program development, we sought to investigate chronic conditions, multimorbidity patterns, and utilization of services among individuals with high costs in a single year in the Veterans Affairs (VA) Health Care System, the largest integrated healthcare system in the United States." (page 3/para 3)

We also revised the opening sentence in the Discussion so that it now reads:

"The 5% highest-cost patients in the U.S. VA Health Care System in 2010 accounted for approximately half of total health care spending." (page 11/para 2)

In addition, in the Discussion we included the following sentence under limitations:

"First, our findings reflect patterns observed among patients who were identified as high-cost over a one-year period, and might not be generalizable to the 20-60% of high-cost patients who persist in this category over two years.4-6" (page 14/para 2)

4. Other researchers have considered that a zero-inflated binomial negative (ZINB) regression is

more suitable than a Poisson distribution to evaluate differences in health care utilization, because data are strongly overdispersed. In my opinion, the authors should include a comment, explaining their choice for Poisson distributions.

To address this comment we tested both models: ZINB (zero-inflated negative binomial) and ZIP (zero-inflated Poisson) for all analyses of utilization outcomes. We used the Vuong test to confirm that zero-inflation was indicated for all models. For all outpatient utilization models, ZINB gave higher log-likelihood values and the additional shape parameter for the ZINB (absent in ZIP) was statistically significant, suggesting that ZINB was the better-fitting model. We therefore now report results from the ZINB models for these outcomes in Table 3 and text. For hospitalizations, convergence appeared to benefit from the one fewer parameter in the ZIP model so we reports results from ZIP.

Of note, for most utilization outcomes there were minimal changes when using ZINB instead of ZIP. The exception to this occurred for mental health clinic visits when we examined this outcome for all patients. This is likely due to the fact that half of the study population did not have mental health conditions and most of these patients likely had low mental health utilization. When we restricted the analysis to patients with mental health conditions, the ZINB and ZIP models are similar (results are reported in the table footnote). We have revised the methods (page 8/para 1) and results (page 11/para 2) to incorporate our use of and findings from the ZINB models, respectively.

5. There are some minor discrepancies between figures in Appendix table 1 and text (percentage of hypertension 63 – 65; mental health condition: 48 -47)

Thank you- we have fixed these numbers.

6. There are not references about ethics approval. Also I have not found a STROBE checklist

As described at the end of the Methods section, this study was approved by the Stanford University Institutional Review Board. We have also completed the STROBE checklist and will upload this with our revised manuscript.

“Analyses used de-identified data and were approved by the Stanford University Institutional Review Board.” (page 8/para 2)