

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)	Trends in High Deductible Health Plan Enrollment and Spending Among Commercially Insured Members with and without Chronic Conditions: A Natural Experiments for Translation in Diabetes (NEXT-D2) Study
AUTHORS	Garabedian, Laura; Zhang, Fang; LeCates, Robert; Wallace, Jamie; Ross-Degnan, Dennis; Wharam, JF

VERSION 1 – REVIEW

REVIEWER	Sarah R Haile Epidemiology, Biostatistics and Prevention Institute, University of Zurich
REVIEW RETURNED	12-Oct-2020

GENERAL COMMENTS	<p>Thank you for the opportunity to review this very interesting manuscript. I have a few questions related to statistical methodology.</p> <p>What proportion of the subjects in each of the three groups had a higher deductible?</p> <p>Was the association between income and size of deductible explored, overall and in the three groups?</p> <p>Were subjects with a deductible between 500 and 100 dollars just excluded from the analysis?</p> <p>I couldn't find much mention of missing data. How many of the variables were missing and where? GEE requires complete data, so I'm wondering how many subjects were removed from the analysis due to this? It's also described how the deductible is imputed, but it appears that single imputation was performed. Is that correct? If that's true, please use multiple imputation instead, as single imputation underestimates the variance in any models.</p>
-------------------------	--

REVIEWER	Betsy Cliff University of Illinois Chicago, USA
REVIEW RETURNED	13-Oct-2020

GENERAL COMMENTS	<p>Trends in High Deductible Health Plan Enrollment and Spending Among Commercially Insured Members with and without Chronic Conditions: A Natural Experiments for Translation in Diabetes (NEXT-D2) Study</p> <p>Reviewed Oct 12, 2020</p>
-------------------------	---

This study describes trends in enrollment and costs for those in high-deductible health plans (HDHPs), comparing healthier enrollees to those with two common chronic conditions, diabetes and cardiovascular disease. It uses data from a large national insurer from 2005-2013. While some of the findings, particularly those on out-of-pocket costs—are well-established in the literature, enrollment trends among people with different health risks are not, making this paper a potentially important addition to the literature. I have outlined several concerns with this project below, notably with the presentation and interpretation of results. I appreciated the chance to read this paper and all comments are intended to help the authors improve upon their work.

Major Concerns

- 1) The most important issue that needs to be addressed in this study is the interpretation of enrollment patterns given by Figure 1 and results in “Rates of HDHP Enrollment Over Time” section. The authors note that the “rate of enrollment of HDHPs increased markedly over the study period for all disease categories, increasing by approximately 6 percentage points per year....” The figure and the percentage point result both show a constant rate of increase, not an increasing rate. An increasing rate would have a convex shape, whereas the figure appears roughly linear. Further, you note that you used GEE models to model trends and annual rates, but never give results that suggest differences year-over-year, or test whether these differences are increasing at a statistically significant rate. It also does not appear that you tested year-to-year differences in prevalence of healthier vs chronic disease enrollees—only differences in average rates. The methodology notes you test the average difference in rates between disease groups and healthier members (p. 5; Statistical Analysis). Your research question implies a test of whether the composition of enrollees is changing over time, which implies a formal analysis of not just average differences but also year-to-year rates of enrollment.
- 2) Related to the first concern, much of the paper will need to be rewritten. The abstract Conclusion and manuscript Discussion sections are currently not supported by the results of this study.
- 3) Odds ratios have limitations not noted in the paper. Notably, they lack external validity because odds ratio are quite dependent on the number of explanatory variables and sample. (EC Norton et al. JAMA. July 2018). Additionally, odds ratios can be hard to interpret, and are often mistaken as probabilities. Perhaps consider a more intuitive measurement in Table 1, such as predicted prevalence.
- 4) I am confused about exactly how the sample is drawn—what percent of enrollees in your sample actively select into plans vs are subject to rollover by employers? Some of my confusion is likely semantic – in the Intro you describe the rigorous methodology of the NEXT-D1 study, and deem this the NEXT-D2 study—but do not appear to use same methodology here. I think this confusion could be cleared up by deleting description of methodology for NEXT-D1 in Intro. However, the issue of selection into these plans remains and is never explicitly addressed. Chronic disease patient selecting into these plans is very different behaviorally from being ‘forced’ in by an employer decision. I think your enrollment results will be much more meaningful if you can tease that out.

	<p>5) Citations to other work need to be included. These authors primarily cite their own work. While they have contributed greatly to the literature on this topic, there are a number of researchers who have also contributed to enrollment in HDHPs (Kaiser Family Foundation); selection into HDHPs (e.g. Lave, Men et al., HSR, 2010) and spending in HDHPs (e.g. Melinda Buntin; Paul Fronstin (EBRI))</p> <p>Minor Issues</p> <p>1) Put years (2005-2013) of sample into Abstract</p> <p>2) How do you handle enrollees who turn 65 during plan year?</p> <p>3) The finding that OOP costs decreased over time seems counter to the idea of rising health care costs and rising share of cost shifting to enrollees. Any idea what's going on?</p> <p>4) The finding that members with chronic disease paid a lower share of total costs is also interesting. Why is that? Is it because of higher costs in the first place? Are they more likely to hit their deductible? If so, what's their incentive for spending less (in the standardized spending algorithm) implying lower utilization? Some of this may be beyond the scope of this article, but any insights you can put into Discussion along these lines would be welcome.</p> <p>5) You note in the Discussion that HSA-eligible members spend more OOP. This finding is interesting and should be highlighted more, perhaps with context of recent finding (Kullgren et al, JAMA Open, 2020) that >50% of people with accounts do not put any money into HSAs.</p> <p>6) Tables in Appendix 4 are hard to interpret. I think they could use better labeling or more exposition in notes.</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Comments to the Author

Thank you for the opportunity to review this very interesting manuscript. I have a few questions related to statistical methodology.

1. What proportion of the subjects in each of the three groups had a higher deductible?

Our original analysis focused on the percent of members with deductibles >\$1000 (and we called this a HDHP). For the resubmission, we split our HDHP cohort into two groups - those with deductibles \$1000-2499 and those with deductibles greater or equal to \$2500 – for the analyses that examines the percentage of members enrolled in a HDHP. These results appear in Appendix 5 and are mentioned in the results section (under “Percentage of Members Enrolled in a HDHP Over Time”): “The percentage of members enrolled in a very high deductible health plan (≥\$2500) increased over the study period for all disease categories, from less than 1% in 2005 to 14-15% in 2015 (Appendix 5). In the last three years of the study period, the percentage of members in a HDHP with a deductible between \$1000 and \$2499 remained relatively flat, and the increase observed in HDHP plan enrollment overall was driven by enrollment in very high deductible health plans.”

2. Was the association between income and size of deductible explored, overall and in the three groups?

This is an insightful question, but answering it is beyond the scope of the current paper. We are

currently working on another project that examines socio-economic disparities in HDHP uptake and spending and we will add this suggested analysis to that paper.

3. Were subjects with a deductible between 500 and 1000 dollars just excluded from the analysis? Members with deductibles of \$501-\$999 were excluded from the analyses that examine OOP and total costs since we are comparing members in HDHP and low deductible health plans head to head. However, they are included in the models that assess percentage of members enrolled in a HDHP (they fall in the denominator). We updated the methods section to clarify this, as follows: "In the models to assess percentage of members enrolled in a HDHP and predictors of HDHP enrollment, the denominator was all members in that disease category. The analyses that examined OOP and total costs compared members in HDHP (\geq \$1000) to members in low deductible plans (\leq \$500) and therefore excluded members with deductibles of \$501-\$999."

4. I couldn't find much mention of missing data. How many of the variables were missing and where? GEE requires complete data, so I'm wondering how many subjects were removed from the analysis due to this? It's also described how the deductible is imputed, but it appears that single imputation was performed. Is that correct? If that's true, please use multiple imputation instead, as single imputation underestimates the variance in any models.

A very small percent of the members had missing values for the neighborhood level variables; we now include the n and % missing in the demographic table (Appendix 3). We excluded members with missing variables from the GEE models. We added a statement about missingness and the GEE variables to the limitations section of the paper.

The question about imputation of the deductible level is one that we have thought about at length and have addressed in other peer-reviewed publications that use the same deductible imputation. We did not use multiple imputation for 3 reasons:

1. A key requirement for multiple imputation – that missingness be random – is not met. Missingness in deductible level is in fact highly systematic given that it is almost entirely among larger employers. The reason for this systematic missingness is that we have a benefits table (list of exact deductible levels per employer ID) that includes mostly small employers and fewer large employers.
2. The uncertainty about the deductible we inferred from claims is low. Moreover, as we now demonstrate below, sensitivity and specificity improve substantially at larger employers (where missingness in deductible level predominates) because such employers generate more claims, providing better evidence about actual deductible levels at their firms.
3. Carrying out formal multiple imputation would be a practical challenge. The programming steps that make up and follow our imputation algorithm are enormously time consuming due to our large sample size (the process incorporates analysis of every single claim from every one of the many millions of members in the entire dataset) and the complexity of SAS programs.

We previously validated our imputation algorithm in a single category of employer size that has substantial overlap with gold standard data (i.e., 75 to 100 enrollees per employer). However, we do have gold standard data on a smaller number of large employers. Therefore, we created 3 additional categories of employer size (101-400, 401-700, and 701-1000) and tested sensitivity and specificity, finding a range from 96% to 100% that increased across employer size category. We now include these results as Appendix Table 1.a-1.d.

Furthermore, in the limitations section of the Discussion, we now include the following sentences: “Although we knew the exact deductible level of most smaller employers, we had to infer it from claims at large employers. However, the sensitivity and specificity of our algorithm was high and increased across employer size category, ranging 96% to 100% (Appendix 1). We expect adjustment for the uncertainty of the imputation process would have a negligible effect on confidence bounds.”

Reviewer: 2

Comments to the Author

This study describes trends in enrollment and costs for those in high-deductible health plans (HDHPs), comparing healthier enrollees to those with two common chronic conditions, diabetes and cardiovascular disease. It uses data from a large national insurer from 2005-2013. While some of the findings, particularly those on out-of-pocket costs—are well-established in the literature, enrollment trends among people with different health risks are not, making this paper a potentially important addition to the literature. I have outlined several concerns with this project below, notably with the presentation and interpretation of results. I appreciated the chance to read this paper and all comments are intended to help the authors improve upon their work.

Major Concerns

1. The most important issue that needs to be addressed in this study is the interpretation of enrollment patterns given by Figure 1 and results in “Rates of HDHP Enrollment Over Time” section. The authors note that the “rate of enrollment of HDHPs increased markedly over the study period for all disease categories, increasing by approximately 6 percentage points per year....” The figure and the percentage point result both show a constant rate of increase, not an increasing rate. An increasing rate would have a convex shape, whereas the figure appears roughly linear. Further, you note that you used GEE models to model trends and annual rates, but never give results that suggest differences year-over-year, or test whether these differences are increasing at a statistically significant rate. It also does not appear that you tested year-to-year differences in prevalence of healthier vs chronic disease enrollees—only differences in average rates. The methodology notes you test the average difference in rates between disease groups and healthier members (p. 5; Statistical Analysis). Your research question implies a test of whether the composition of enrollees is changing over time, which implies a formal analysis of not just average differences but also year-to-year rates of enrollment.

We apologize that the terminology in our original submission was confusing. Throughout the manuscript, we replaced “rate” with “percentage of members in a HDHP,” to describe our HDHP enrollment analysis. In addition to examining the percentage of members in a HDHP each year, we also measured the slope of the enrollment (i.e., year-over-year differences) – we now refer to these analyses as “trend” (e.g., the trend in HDHP enrollment increased by 5 percentage points a year). In the revised manuscript, we tested whether these trends were statistically significant (they all were) and added the 95% confidence intervals of the slopes. We conducted these two analyses (i.e., examined the percentage of members in a HDHP and the trend of HDHP enrollment) separately for each disease category. Since HDHP enrollment in each disease category had similar year-to-year trends (i.e., slopes), we did not test the difference in trends, as you noted (although the reader can examine the trends head to head by looking at the confidence intervals for each disease area provided in the results section). Rather, we focused on the average difference in the percentage of members in a HDHP for diabetes vs. healthy members and CVD vs. healthy members.

For the OOP and total cost analyses, we provide the following analyses to compare HDHP and low-deductible health plans within each disease category.

- Average absolute difference in cost between HDHP and low-deductible health plans over the study period
- Average relative difference in cost between HDHP and low-deductible health plans over the study period
- Slope (i.e., year-over-year change, aka “trend”) of average costs for HDHP and low-deductible health plan members

(Note that we do not explicitly compare costs across the different disease categories since the key analysis is HDHP vs. low deductible plans, although the differences between disease categories are apparent in the Figures and manuscript.)

2) Related to the first concern, much of the paper will need to be rewritten. The abstract Conclusion and manuscript Discussion sections are currently not supported by the results of this study. With the changes made above, we now more clearly describe the results and the conclusion and discussion sections are supported by the results.

3) Odds ratios have limitations not noted in the paper. Notably, they lack external validity because odds ratio are quite dependent on the number of explanatory variables and sample. (EC Norton et al. JAMA. July 2018). Additionally, odds ratios can be hard to interpret, and are often mistaken as probabilities. Perhaps consider a more intuitive measurement in Table 1, such as predicted prevalence.

Per your suggestion, we have replaced the odds ratios in Table 1 with average adjusted predictions, which we refer to as “predicted probability of HDHP enrollment.” We also now use this method to calculate the percentage of members in a HDHP. We have updated the methods section and all of the results in the manuscript and Table 1 and Figure 1.

4) I am confused about exactly how the sample is drawn—what percent of enrollees in your sample actively select into plans vs are subject to rollover by employers? Some of my confusion is likely semantic – in the Intro you describe the rigorous methodology of the NEXT-D1 study, and deem this the NEXT-D2 study—but do not appear to use same methodology here. I think this confusion could be cleared up by deleting description of methodology for NEXT-D1 in Intro. However, the issue of selection into these plans remains and is never explicitly addressed. Chronic disease patient selecting into these plans is very different behaviorally from being ‘forced’ in by an employer decision. I think your enrollment results will be much more meaningful if you can tease that out.

You are correct. The study population is the entire population of members in the study categories (i.e., diabetes, cardiovascular disease, healthy) in the database; it is not limited to the NEXTD-1 study population (i.e., members forced into HDHPs by their employer). We see how this led to confusion and removed the details of the NEXTD-1 study design from the intro.

The issue of members’ *choice* of HDHPs is a very interesting one and a topic of another paper on which we are working. We alluded to this limitation and future work in the discussion section of the original submission (“Finally, our data combined HDHP members whose employers offered only a

HDHP with members who were offered a choice by their employer to enroll in a HDHP or a lower-deductible plan. Future research should examine HDHP enrollment among members who have plan choice to better understand factors associated with selecting HDHPs.”)

However, in the revision, we have decided to include a few of the choice calculations: (1) the percentage of members who are offered a choice of a HDHP and a non-HDHP from their employer and (2) the percentage of members with employer-level choice who choose to enroll in a HDHP. These now appear in the methods and results sections and in Appendix 6 and we allude the results in the discussion section: “Among the subset of members who were offered a choice of a HDHP or lower deductible plan from their employer, most members opted for a lower deductible health plan and members with chronic diseases were less likely to choose a HDHP than healthier members.”

5) Citations to other work need to be included. These authors primarily cite their own work. While they have contributed greatly to the literature on this topic, there are a number of researchers who have also contributed to enrollment in HDHPs (Kaiser Family Foundation); selection into HDHPs (e.g. Lave, Men et al., HSR, 2010) and spending in HDHPs (e.g. Melinda Buntin; Paul Fronstin (EBRI)) Thank you for these suggestions. We added references from Lave, Buntin and Fronstin to the introduction and discussion sections of the paper. We had already included references to the Kaiser Family Foundation in the original submission.

Minor Issues

1) Put years (2005-2013) of sample into Abstract

The years are in the abstract (under participants).

2) How do you handle enrollees who turn 65 during plan year?

We assigned age at the annual level using member age at the anniversary month, which is the last month of each member’s 12 month cycle. Anyone with age >64 at the anniversary month would have been excluded so there was no opportunity for anyone to turn 65 during the study period.

3) The finding that OOP costs decreased over time seems counter to the idea of rising health care costs and rising share of cost shifting to enrollees. Any idea what’s going on?

OOP costs remained relatively flat over the time period for the HDHP members and decreased slightly for the low deductible plan members. There are multiple explanations for this. One, our cost figures are adjusted for medical cost inflation (unlike the figures in reports like this one from Kaiser Family Foundation: <https://www.healthsystemtracker.org/brief/tracking-the-rise-in-premium-contributions-and-cost-sharing-for-families-with-large-employer-coverage/>). Also, unlike the KFF report, our OOP figures are adjusted for changes in demographic and health characteristics and employer characteristics and stratified by disease category (meaning that increases in the prevalence of diabetes and cardiovascular disease over time cannot impact the OOP costs in our study).

4) The finding that members with chronic disease paid a lower share of total costs is also interesting. Why is that? Is it because of higher costs in the first place? Are they more likely to hit their deductible? If so, what’s their incentive for spending less (in the standardized spending algorithm) implying These are all very interesting and important questions. Yes, members with chronic diseases paid OOP for a lower share of total costs than healthier members because members with chronic diseases have much higher total costs and these members either hit the deductible or OOP max limits. Fronstin et al. (EBRI, October 24, 2019) showed that half of high cost members (which they

define as members with the top 10% of spending) hit their deductible and OOP max. We added this to the discussion section of the article.

Also, you are correct - there is evidence that once members hit a deductible they have little incentive to spend less. HDHP members who exceed the deductible increase utilization of low value care and screenings (e.g., Fronstin EBRI Issue Brief October 22, 2020, No. 516) and adding cost-containing measures (copayments/co-insurance) after the deductible is met decreases spending (Fronstin EBRI Issue Brief November 19, 2020, No 519). However, analyzing spending before and after a member hits their deductible is beyond the scope of this study.

There is evidence from multiple studies (mentioned in the intro) that deductibles provide a financial disincentive to seek needed care – this is likely the driver of lower costs we observe in HDHP members compared to members with lower deductibles.

lower utilization? Some of this may be beyond the scope of this article, but any insights you can put into Discussion along these lines would be welcome.

5) You note in the Discussion that HSA-eligible members spend more OOP. This finding is interesting and should be highlighted more, perhaps with context of recent finding (Kullgren et al, JAMA Open, 2020) that >50% of people with accounts do not put any money into HSAs.

Thank you for this suggestion. We added the Kullgren et al. reference to the discussion section.

6) Tables in Appendix 4 are hard to interpret. I think they could use better labeling or more exposition in notes.

Thank you for this comment. We split Appendix 4 into two appendices (Appendix 4a and 4b), renamed the Appendix title and figure titles within the appendix, and added a note to explain the denominator. We also added to the results section of the manuscript that “non-account HDHPs were the most common HDHP type for HDHP members in all three disease areas” (this refers to data in Appendix 4a) and then made it clear that the next result mentioned (“the higher percentage of enrollment in HDHPs among healthy members, compared to members with chronic diseases, was associated with higher enrollment in HSA-eligible HDHPs”) specifically refers to the figures in Appendix 4b.

VERSION 2 – REVIEW

REVIEWER	Haile, Sarah University of Zurich, Epidemiology, Biostatistics and Prevention Institute
REVIEW RETURNED	06-Mar-2021

GENERAL COMMENTS	Thank you for your response. No further comments.
-------------------------	---

REVIEWER	Cliff, Betsy University of Illinois at Chicago, Health Policy & Administration
REVIEW RETURNED	04-Mar-2021

GENERAL COMMENTS	<p>I really appreciate the authors' responsiveness to my previous concerns. They have assuaged all my concerns about the validity of their analyses and interpretation of results. In my opinion, this article is ready to be accepted. In re-reading the manuscript, I had 2 minor changes suggested to increase clarity:</p> <p>1) In "Results:Percentage of Members Enrolled in an HDHP Over Time", sentence reading "Members with chronic diseases were less likely than healthier members to be in an HDHP throughout the entire study period (Figure 1)." I interpreted this statement as implying as longer spells of enrollment by healthier HDHP members. Based on Figure 1, I think what you mean is that, throughout the study period, healthier members had consistently higher levels of HDHP enrollment. Change the wording to reflect that interpretation. Same issue in second sentence of the Discussion. Make it clear that you are talking about the level of enrollment not the length of spells. Perhaps something like "In all study period years, members with diabetes and CVD had lower levels of enrollment."</p> <p>2) In notes for Figure 1, add more details about how the percentage is adjusted, i.e. the exact regression model used.</p>
-------------------------	---

VERSION 2 – AUTHOR RESPONSE

We responded to Reviewer 1's comments and made the following minor changes:

1. Revised two sentences (one in the results section and one in the discussion section) to make it clear that members with chronic diseases had lower levels of HDHP enrollment than healthier members (and not shorter spells of HDHP enrollment).
2. Updated the notes in all figures (not just figure 1 as the reviewer requested) to provide more details about how we calculated the adjusted estimates.

Review 2 did not have any suggested revisions.

In response to the editor's comments, we removed one of the "strengths and limitations" bullets and now have 5 bullets. We also added an ethics statement just before the references list.

We look forward to having our work published in BMJ Open! Please let me know if you need me to make any further changes.