Author’s response to reviews

Title: Disease Associations Depend on Visit Type: Results from a Visit-Wide Association Study

Authors:

Mary Regina Boland (bolandm@pennmedicine.upenn.edu)
Snigdha Alur-Gupta (snigdha.alur-gupta@uphs.upenn.edu)
Lisa Levine (lisa.levine@uphs.upenn.edu)
Peter Gabriel (peter.gabriel@uphs.upenn.edu)
Graciela Gonzalez-Hernandez (gragon@pennmedicine.upenn.edu)

Version: 1 Date: 23 May 2019

Author’s response to reviews:

Reviewer Comments with Author Responses

We thank the reviewers for their thoughtful comments. We have made substantial edits to the manuscript and have improved the readability of our work. Below is a point-by-point response to the reviews.

Reviewer 1

Reviewer 1, Comment 1: The authors should give a definition of EMR, what it stands for? The authors should also give a brief explanation of ICD-9 and ICD-10 for those reads are not familiar with these terms.

Author Response: We thank the reviewer for this comment. We realized that we used the acronym ‘EMR’ in two locations in our initial submission where we should have used ‘EHR’, which stands for Electronic Health Records. We have corrected this in our current version. We have also included a brief explanation of ICD-9 and ICD-10 for those not familiar with these terminologies. On page 4 it now states,

“For defining diagnoses, we used the International Classification of Diseases (ICD) version 9 (ICD-9) and version 10 (ICD-10). The ICD-9 and ICD-10 codes are used primarily for billing purposes to describe the conditions, symptoms, illnesses and diseases for a given patient. We collectively term ICD-9 and ICD-10 diagnoses as conditions in our study.”
Reviewer 1, Comment 2: What are those 7186 conditions? Are they different kinds of disease?

Author Response: We thank the reviewer for this comment. We have provided some more details and clarity on what the 7186 conditions contain. On page 5 it now states,

“Overall, there were a total of 7,186 conditions coded for in our dataset. Conditions include diseases (e.g., breast cancer), symptoms (e.g., lump in breast), and infections (e.g., flu) provided that there is a distinct ICD-9 or ICD-10 for that condition.”

Reviewer 1, Comment 3: In Table S1, how the numbers in Ob./Gyn Visit vs. Cancer Visit calculated?

Author Response: Some information on the three different comparison groups could be found in section 5.4. However, based on the reviewer’s comment, we have added additional information into the results section, specifically near the reference to Table S1 to help clarify this issue. On page 5 it now states,

“In our VisitWAS study, we performed three different comparisons: 1) those treated at an Ob/Gyn clinic versus any visit (i.e., the entire population), 2) those treated at a cancer clinic versus any visit (i.e., the entire population) and 3) those treated at an Ob/Gyn clinic versus those treated at a cancer clinic. Because patients visiting the Ob/Gyn and those visiting a cancer clinic often visit frequently it made sense to compare these two groups to each other to highlight conditions that were more strongly associated with one group versus the other.”

Reviewer 1, Comment 4: In 2.3 Pain Depends on Visit Type section, Only 129 pain conditions were in our dataset …, there should be 119 pain conditions according to the numbers given in Ob/Gyn visits and cancer visits.

Author Response: We thank the reviewer for this comment. There were only 129 pain conditions in our dataset. Many of the pain conditions associated with cancer were also associated with Ob/Gyn visits therefore adding those two numbers to equal 119 is not correct. We realized this was confusing and therefore, we altered Table S2 for clarity to show the numerators and denominators for all percents shown. We also clarified the text on page 6, which now states,

“Overall 129 pain conditions were found in our dataset. We looked at the proportion of associated conditions that were pain related and found that 43 or 2.00% (43/2150) of unique conditions associated with Ob/Gyn visits were pain related (Table S2).”
Reviewer 1, Comment 5: What are the blue and red colors mean in Figure 2 and Figure 3? Are the results of Figure 2 and Figure 3 are similar since one is based on -log(p) and the one is based on log(OR)?

Author Response: We thank the reviewer for this comment. We used blue and red just to distinguish between the various ICD-9 disease categories by alternating the color. We have added notation of this to the Figure legends. Both Figure 2 and Figure 3 now include the following sentence, “We alternate between blue and red to highlight the different disease categories.” Yes the results are similar because Figure 2 shows the –log(p value) and Figure 3 shows the log(OR). However, we feel that these 2 figures illustrate different facets of our results. One shows the strength of the association in terms of statistical significance and the other shows the size of the effects.

Reviewer 2

Reviewer 2, Comment 1: This study investigated the association between outpatient diagnosis codes and certain visit types (Ob/Gyn clinics and cancer clinics), as well as compared the difference of "pain-related" conditions between these two clinics. A new concept of VisitWAS was established based on Phenome-Wide Association Studies (PheWAS). The manuscript is quite straightforward with clear results and conclusions. I have several concerns and suggestions about the contents.

Author Response: We thank the reviewer for this comment. We hope we have addressed all of the suggestions and comments in this revision.

Reviewer 2, Comment 2: From the introduction part, I firstly learned that PheWAS analyze the association between phenotypes and disease (line 9). According to my understanding, PheWAS examine the effects of many phenotypes on a single disease/genetic variant. However, in the following sentences, PheWAS became a study that examines the "gene-disease associations" (line 13-19), which sounded the same as the well-known GWAS. And later, the author wrote that PheWAS investigated genes associated with "structure disease codes" (line 19-21). I thought the description here needs to be clarified: what is the exact definition of PheWAS (e.g.: Does it investigate the association of multi-phenotypes with a single trait such as gene/disease? Or the association of multi-genes with a single disease/phenotype?) Current words are confusing. It may also make the readers feel confused when they learn about the definition of VisitWAS: What is the full name of "VisitWAS"? If it were "Visit-Wide Association Studies", will it examine the relationship between many visit types and a single disease trait? Or will it examine the relationship between many disease traits and a single visit type?
Author Response: We thank the reviewer for this question. We have revised how we introduce the concept of PheWAS and VisitWAS. PheWAS involves the comparison of high-throughput exploration of the relationship between multiple SNPs and distinct diseases extracted from EHRs. We have added clarifying text to the Introduction section of the manuscript on page 3. We have added a new paragraph on page 3, which states,

“With the expansion of the number of EHR datasets available from across the country and the world, more associations were conducted, typically between a single key phenotype (e.g., periodontal disease) and many diseases simultaneously extracted from EHRs [10]. Later work involved investigating the relationship between birth month/season and all diseases extracted from the EHR having at least 1000 patients[11]. These studies follow the framework of ‘disease association studies’ where a large number of diseases are simultaneously investigated for their relationship or association with a key exposure or other outcome of interest.

However, neither PheWAS studies nor many EHR based association studies investigate the effects of visit type on results. The importance of visit type in PheWAS studies has been discussed as future work in several studies [12, 13]. However, no available methods have been developed to investigate the effect of visit type on PheWAS studies.”

We have also expanded the detail in the purpose paragraph on page 4, which now states,

“The purpose of this study is to investigate the relationship between visit type and disease associations revealed from the EHR. We call this method ‘VisitWAS’ to signify Visit-Wide Association Study. Instead of investigating the relationship between one key disease or outcome’s relationship with other diseases in the EHR as in traditional ‘Wide Association Studies’, we are studying the effect of the visit type on the diseases in the EHR. We focus on two groups that are experience a large number of different types of pain diagnoses (i.e., ‘high pain’) groups – those visiting Ob/Gyn and those visiting a cancer clinic. We will also perform a sub-analysis on pain diagnoses to investigate pain diagnoses that are associated with Ob/Gyn visits vs. cancer visits. We chose to investigate pain diagnoses because many pain diagnoses, e.g., ‘pain in abdomen’ are vague and can mean very different things depending on the context of the code (i.e., was it a cancer visit or an Ob/Gyn visit where this code was used). Our work sheds light on the importance of visit type in PheWAS studies.”
Reviewer 2, Comment 3: Line 51 in the introduction part, it will be helpful to provide reference on why and how to define the "high pain" groups. Additionally, totally how many clinics were in your data (what are they)? Why were only Ob/Gyn and cancer clinics chosen? Why not other clinics (some may be highly concerned such as heart disease clinic)? Some more rationale should be addressed.

Author Response: We thank the reviewer for this comment. We have added some additional details regarding the definition of high pain and also why the clinics and comparisons were chosen.

In the introduction on page 3, more clarity on the high pain and also the types of clinics now states,

“We focus on two groups that are experience a large number of different types of pain diagnoses (i.e., ‘high pain’) groups – those visiting Ob/Gyn and those visiting a cancer clinic. We will also perform a sub-analysis on pain diagnoses to investigate pain diagnoses that are associated with Ob/Gyn visits vs. cancer visits. We chose to investigate pain diagnoses because many pain diagnoses, e.g., ‘pain in abdomen’ are vague and can mean very different things depending on the context of the code (i.e., was it a cancer visit or an Ob/Gyn visit where this code was used). Our work sheds light on the importance of visit type in PheWAS studies.”

We also add clarity on the types of clinics and the dataset in general. On page 4 in section 2.1 on the Dataset, it now states,

“This study was conducted using out-patient data obtained from the Hospital of the University of Pennsylvania (called hereafter Penn). Penn contains data from the Philadelphia Metropolitan Area. This includes outpatient data from Southern New Jersey, Philadelphia, parts of Delaware and in the Pennsylvania suburbs of Philadelphia. These data are all structured data collected during routine clinical care between the period of 2006 and 2017. We used women only to demonstrate the effect of visit type on disease associations and also to compare pain diagnoses across cohorts (Figure 1). Our ‘all clinics’ as shown in Figure 1 represent all clinics at UPenn where our cohort of women were treated. This includes clinics for cardiology, rheumatology, immunology to name a few. We only investigate women in our analyses because we compare those being treated Ob/Gyn visits versus cancer visits and other visits. Since Ob/Gyn visits generally are made by females it would not make sense to compare against males. In addition many diagnoses, including pain, vary by sex. The demographics of women included in this study are found in Table 1.”
Reviewer 2, Comment 4: Line 11-14 in the 2.1 Dataset. More background details of data source are required. For instance, from which year to which year were the out-patient data obtained? How were these out-patient data obtained by the Hospital of UPenn (self-report? Lab examination? Follow-up call?)? Where is this hospital and patients in which area was covered?

Author Response: We thank the reviewer for this comment. We have added some additional details in the dataset section of the paper in this revision. On page 4 it now states,

“This study was conducted using out-patient data obtained from the Hospital of the University of Pennsylvania (called hereafter Penn). Penn contains data from the Philadelphia Metropolitan Area. This includes outpatient data from Southern New Jersey, Philadelphia, parts of Delaware and in the Pennsylvania suburbs of Philadelphia. These data are all structured data collected during routine clinical care between the period of 2006 and 2017.”

Reviewer 2, Comment 5: Table 1, I am curious about the age distribution of participants. Please provide std together with mean or IQR together with median of their age. If possible, stratifying age into categories (e.g.: <20 yrs, 20-50 yrs, 50+ yrs) and providing the number of participants in each category will be helpful. So that the readers can have a rough sense of how many women are at teenage, reproductive age, and post-menopause.

Author Response: This is a very important point. We have added the mean + standard deviation information to table 1. We also included some additional information on page 5 it now states,

“Because women have multiple visits that can span over several years, we calculated each woman’s average age across their record and then we calculated the overall average and standard deviation shown in Table 1.”
Reviewer 2, Comment 6: I can understand that the author used Fisher's exact test because of the small sample size. But in order to provide a clear logic chain on why Fisher's was chosen, it may be better to show some power calculation and prove that the sample size isn't big enough to give acceptable power for a chi-square test or something else.

Author Response: This is a good point. We included some additional justification for our choice of Fisher’s exact test in the methods section 5.2. In the revised version of the manuscript, on page 10-11 in section 5.2, it now states,

“We performed association analysis using Fisher’s exact test for the association between outpatient diagnosis codes and certain visit types, following the PheWAS framework [2]. We investigated the association between each diagnosis code and visit type provided that the diagnosis code appeared in at least 50 patients at Penn. We chose Fisher’s exact test because it is robust for small sample sizes. We restricted our analysis to only include those diagnoses with at least 50 patients overall. However, the sample size for a given disease at a particular visit type could be less then 5 patients and therefore Fisher’s exact test is more appropriate then Chi-square.”

Reviewer 2, Comment 7: Figure 1 is kind of weird. I cannot find N=81, N=8 and N=882 from any part of the main text. The numbers are quite different from what Table 1 provided. Figure 1 or Table 1, one of them must be wrong. Also, the pictures are confusing. It will be much better to explain totally how many clinics were included in "all clinics" and what are they.

Author Response: We thank the reviewer for this comment, we investigated the issue and realized that the counts in Figure 1 were wrong – Table 1 was correct. We have clarified that in this revised version. There are 1,048 clinics in the ‘all clinics’ and they represent any type of clinic (e.g., cardiology) and so forth. We have added some clarity on the types of clinics to page 4. On page 4 it now states,

“We used women only to demonstrate the effect of visit type on disease associations and also to compare pain diagnoses across cohorts (Figure 1). Our ‘all clinics’ as shown in Figure 1 represent all clinics at UPenn where our cohort of women were treated.”
Reviewer 3

Reviewer 3, Comment 1: Overall, the impact of this manuscript seems quite limited. The authors seem to put forward that PheWAS studies would be more accurate if they included visit type, which seems sensible. The logical test would be the conduct an EHR-wide PheWAS with and without visit type. However, the study as conducted was limited to outpatient records, female patients, and two clinics. I don't find the results very compelling. Perhaps I'm missing something.

Author Response: This study illustrates how diagnoses can be associated with visit type. We use two high-level visit types – Cancer and Ob/Gyn visits. This is not just two clinics, in fact there are 1,048 clinics at UPenn with 89 distinct Ob/Gyn clinics and 104 distinct cancer clinics. We believe it is important to demonstrate that diagnoses can vary based on clinic type – this may be obvious for specific cancer diagnoses and specific birth and delivery-related codes. However, certain infections can be more or less common in either cancer visits or Ob/Gyn visits. In addition, in the pain space this can also be an issue as many pain diagnosis codes are vague, such as ‘pain in abdomen’ and therefore could mean very different things depending on the context of that diagnosis. Part of the ‘context’ of the diagnosis code includes the visit type. We hope our work clarifies this and we have revised the paper throughout to ensure that the contribution of this work appears more transparent to readers. In addition, on page 3 in the Introduction section we have added some additional text to clarify,

“We chose to investigate pain diagnoses because many pain diagnoses, e.g., ‘pain in abdomen’ are vague and can mean very different things depending on the context of the code (i.e., was it a cancer visit or an Ob/Gyn visit where this code was used). Our work sheds light on the importance of visit type in PheWAS studies.”

Reviewer 3, Comment 2: Introduction, line 50: It is not immediately clear that "high pain" in this case means that the patients experience pain often. Rephrasing would help the reader.

Author Response: We thank the reviewer for this comment. We have revised this statement to clarify this. On page 3 it now states,

“We focus on two groups that are experience a large number of different types of pain diagnoses (i.e., ‘high pain’) groups – those visiting Ob/Gyn and those visiting a cancer clinic.”
Reviewer 3, Comment 3: Results, line 18. The authors restricted to include only women, they did not stratify by gender.

Author Response: We thank the reviewer for spotting this issue. We have clarified the text regarding our inclusion criteria. On page 4 it now states,

“This study was conducted using out-patient data obtained from the Hospital of the University of Pennsylvania (called hereafter Penn). We used women only to demonstrate the effect of visit type on disease associations and also to compare pain diagnoses across cohorts (Figure 1). Our ‘all clinics’ as shown in Figure 1 represent all clinics at UPenn where our cohort of women were treated. This includes clinics for cardiology, rheumatology, immunology to name a few. We only investigate women in our analyses because we compare those being treated Ob/Gyn visits versus cancer visits and other visits. Since Ob/Gyn visits generally are made by females it would not make sense to compare against males. In addition many diagnoses, including pain, vary by sex. The demographics of women included in this study are found in Table 1.”