**ARTICLE DETAILS**

<table>
<thead>
<tr>
<th>TITLE (PROVISIONAL)</th>
<th>Familial aggregation of multimorbidity in Sweden: a national explorative family study</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHORS</td>
<td>Zöller, Bengt; Pirouzifard, MirNabi; Holmquist, Björn; Sundquist, Jan; Halling, Anders; Sundquist, Kristina</td>
</tr>
</tbody>
</table>

**GENERAL COMMENTS**

These are terrific data, which is a tough read, but which potentially generate some interesting hypotheses. The core strength of the paper in the way there is an a priori ordering of the familial context. More problematic is the arbitrary definition of multimorbidity and conditions in which the genetic determinants are either unknown or pleiotropic. Whilst associations can be shown, this study does not distinguish the gradient of genetic-environment interaction, so the call for more genetic studies rather than social determinants, is somewhat tenuous. One might contentiously hypothesise that closer relations talk more about the shared conditions they experience.

So to the multimorbidity list and definition. The list was based on a previous paper and then another 6 added, with a score then of up to 45. However the score is dependent on the prevalence of the problem and recording biases where some conditions will be coded more than others. Especially over the time frame from 1964 to 2015. This list is then reduced to 2 or more. The cohort is born from 1932 but on the ascertainment of multimorbidity is in an 18 year time-period (1997-2015), with only 16.63% with multimorbidity of two are more. Median age at study end is only 35 years (IQR 24-46), so young which means they were recently recruited. The prevalence of the conditions is also age-dependent, which means the list of 45 will be skewed. Much more critical thought into the data accrual over time and change with age should be included in the analyses.

The most interesting part of the paper is the method of principal component analysis. One wishes there had been an a priori list of genetic attribution, so factor 1 (CVD) and factor 2 (mental illness) seem obvious but in design terms (the temporal window), these could be for reasons other than genetic. If this method was applied to a more fully measured cohort, then the application and inference would be much more plausible.
The research question is important and the authors can be commended for their attempt, but the timeframe and limitation of multimorbidity definition, makes any inference highly cautious.

---

**GENERAL COMMENTS**

**SUMMARY:**
Thank you for the opportunity to review this paper on an important and complex topic. In the study, the researchers aimed to determine the familial aggregation of multimorbidity using a large multigenerational national registry from Sweden. They further test multimorbidity inheritance across genetic resemblance and disease clusters.

The study’s strengths include extensive analyses testing various aspects of the association between a composite multimorbidity outcome and family relations. The study is limited in the factors that are adjusted for, particularly those related to lifestyle, which makes it difficult to rule out the family environmental influences vs. genetic influences.

This article addresses a valuable question and has attempted to address it from various angles; however, I have some concerns about the premise of the multimorbidity outcome definition and other comments for clarification.

**MAJOR COMMENTS:**

- The study uses a simplistic composite definition of multimorbidity that may be too crude an outcome measure to answer the question the researchers are trying to answer. It is not entirely clear what the details were of the multimorbidity scoring, and would be helpful if it were elaborated on. However, it seems as though the researchers took any individual with 2 or more chronic conditions as the definition for the dichotomous multimorbidity outcome and compared the risk for a relative to have any 2+ conditions (which could be a completely different set of conditions from those their sibling, cousin, etc.). Several considerations that are contributing to this incongruence:

  o Some of the 45 diseases that were examined have known heritable characteristics, while others have less or no known heritable traits, and it may not be realistic to combine all these conditions into a single composite outcome measure, due to their heterogeneity.
  o Furthermore, I am not sure that using a count of diseases can really answer the question of genetic resemblance, because for example, even when counting, those who have 3 specific diseases may comprise a very different outcome than someone who has 3 completely different diseases.
  o It seems to me that a more nuanced characterization of multimorbidity may be needed to answer the question of whether there is familial aggregation.

- In grouping MZ and DZ twins together in the model analyses, the researchers may not be disentangling the effect of genetic...
resemblance. Maybe the researchers want to consider separating this group into two groups or regroup the DZ under 'siblings' and study just the MZ twins separately?

If the authors decide to keep the groupings as they currently are, I would suggest adding the rationale for this grouping, despite the difference in genetic resemblance.

SPECIFIC COMMENTS:

Abstract:
The abstract can use some clarification.
There is no study design indicated in the abstract; the authors should add this.
The main outcome measure is the multimorbidity score and seems to fit better under the outcomes section.
Can the authors please clarify what was meant by the following sentence on p., lines 54-56 starting "There was a strong graded association..."?

Introduction:
A relevant study that I would suggest adding to the background on pleiotropy influencing multimorbidity is an article by Amell et al. 2018 (DOI:10.1038/s41598-018-34361-3). The authors can consider including this article's findings as part of the background and even potentially as part of the discussion (in how the PCA results compare to Amell et al.'s findings) of their study.

It would help in clarifying the different parts of the study if the authors specified the objectives in more detail related to each of the main analyses conducted at the end of the introduction (e.g. determine association of multimorbidity in relatives, examine the trend of risk for multimorbidity across the genetic resemblance spectrum.

Methods:
I did not see a study design indicated in the methods and would guess that this is a retrospective cohort study. It would be helpful to include the study design to provide the overall framing, intention and limitations.

In the statistical analysis description on p.9, it would be helpful if the authors elaborated on the use of proband relative pairs related to the odds ratio regression analyses.

Odds for multimorbidity – The set of covariates used in model adjustment seems to be fairly minimal. Have the authors considered additional covariates such as ethnicity/race (evidence as seen in an article by Kalgotra et al. 2020, https://doi.org/10.1038/s41598-020-70470-8), and possibly other surrogates of lifestyle factors, if they are not available in the existing registry database?

PCA – I am most familiar with principal component analysis as a method used to narrow down large sets of risk factors. Therefore, I am curious to better understand why the researchers chose to use PCA for the outcome being studied. Maybe the authors can elaborate on the rationale for using this analysis in the specific context?
Results:
Table 2 – It is not clear what the size of each group (N) is across the models for familial associations (e.g. twins, siblings, etc.). Can the authors please add Ns to the table?

Can the authors add a bit more to the results observations based on the results presented in Table 5?

Discussion and limitations:
Some of the points in the discussion are repeated, e.g. third-degree relatives do not usually share a household environment. Consider removing repetition.

The authors do not dedicate much of the discussion to comparing results from their study to previous studies' findings. It would be meaningful for the researchers to compare their findings to prior studies’ findings in order to provide more context on the value of their results. For example, how do the groupings of diseases found in the PCA compare to other groupings of related chronic diseases and pleiotropy in multimorbidity?

p.16, line 58 – There seems to be a type-o.

The relevance of the discussion on p.17 lines 39-58 about self-reported data vs. hospital diagnoses data was not entirely clear to me.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Recommendation:

Comments:
These are terrific data, which is a tough read, but which potentially generate some interesting hypotheses. The core strength of the paper in the way there is an a priori ordering of the familial context. More problematic is the arbitrary definition of multimorbidity and conditions in which the genetic determinants are either unknown or pleiotropic. Whilst associations can be shown, this study does not distinguish the gradient of genetic-environment interaction, so the call for more genetic studies rather than social determinants, is somewhat tenuous. One might contentiously hypothesise that closer relations talk more about the shared conditions they experience.

So to the multimorbidity list and definition. The list was based on a previous paper and then another 6 added, with a score then of up to 45. However the score is dependent on the prevalence of the problem and recording biases where some conditions will be coded more than others. Especially over the time frame from 1964 to 2015. This list is then reduced to 2 or more. The cohort is born from 1932 but on the ascertainment of multimorbidity is in an 18 year time-period (1997-2015), with only 16.63% with multimorbidity of two are more. Median age at study end is only 35 years (IQR 24-46), so young which means they were recently recruited. The prevalence of the conditions is also age-dependent, which means the list of 45 will be skewed. Much more critical thought into the data accrual over time and change with age should be included in the analyses.

The most interesting part of the paper is the method of principal component analysis. One wishes there had been an a priori list of genetic attribution, so factor 1 (CVD) and factor 2 (mental illness)
seem obvious but in design terms (the temporal window), these could be for reasons other than genetic. If this method was applied to a more fully measured cohort, then the application and inference would be much more plausible.

The research question is important and the authors can be commended for their attempt, but the timeframe and limitation of multimorbidity definition, makes any inference highly cautious.

Author response: Thanks for helpful comments. We now realize that we have not explained matters clearly. We have now explained our findings much better thanks to reviewer comments. The PCA (principal component analysis) clearly shows that the clustering of disease in multimorbidity is specific and not random. Within the groups identified with PCA, there are strong correlations. Thus, the following reviewer comment is not correct "More problematic is the arbitrary definition of multimorbidity and conditions in which the genetic determinants are either unknown or pleiotropic." Using the Barnett et al definition of comorbidity published in Lancet, we performed a PCA analysis as an agnostic method showing that disease clustering in multimorbidity is not random. Instead, disease clustering in multimorbidity is disease-specific. Moreover, these disease clusters are inherited in families. Thus, the clustering of disease in multimorbidity is disease-specific and not a random event. Having more than one related disease gives even high odds ratios for inheritance suggesting an increased risk for a genetic contribution when several related diseases are clustering in a family. We therefore have changed the discussion, abstract and title to clarify this.

The critics of diseases included in the multimorbidity score is not appropriate either. Our definition is based on a modified version by Barnett et al published in Lancet, which we have cited. In fact, Barnett et al specifically sought to include morbidities recommended as core for any multimorbidity measure by a systematic review, diseases in the quality and outcomes framework (QOF) of the UK general practice contract, and long-term disorders identified as important by NHS Scotland. Thus, our definition of multimorbidity is not arbitrary. Most importantly, using the PCA analysis, we show that disease clustering is not random - it is specific.

We have now added two references based on genomic data that indicates that multimorbidity has a genetic contribution. However, our paper is the first to show this using classical tools in genetic epidemiology such as study of the familial relative risk, here expressed as odds ratios in first-, second- and third-degree relatives. We have now clarified even more in order to better explain the importance of our findings also for researchers not familiar with genetics epidemiology. We also add more in the discussion about shared environment as suggested by the reviewer. However, the present high familial odds ratios are very unlikely to be explained by shared environment only. We have now added a reference for this. Moreover, the strong association between genetic resemblance and odds ratio indicates that a genetic component is likely. This is especially true for increased risk among third degree relatives that do usually not share a household. In genetics, our findings are a strong indication of important genetic determinants.

Regarding the study period, we have carefully chosen the study period taking into account the limitations of the registers. An advantage is to include a time-period only using ICD-10. Regarding age dependence, we have adjusted for age in all models. Regarding the age, it is
well-known that the genetic determinants for complex traits are stronger in young subjects, and it could therefore be an advantage to study young subjects to find a genetic contribution. It should be emphasized that the Swedish and Danish multigeneration registers have been shown in hundreds of studies to be an invaluable source of information about inheritance.

Whether or not the genetic determinants are unknown or pleiotropic is not relevant. All genetic determinants for complex traits remain to be determined. Moreover, genetic pleiotropy, i.e. when one gene influences two or more seemingly unrelated phenotypic traits, could be an important cause of the genetic basis for multimorbidity. Moreover, our PCA analysis shows that multimorbidity is disease-specific and that certain diseases cluster together in families and this it is not a random event which diseases that cluster together.

Reviewer: 2
Recommendation:
Comments:
SUMMARY:
Thank you for the opportunity to review this paper on an important and complex topic. In the study, the researchers aimed to determine the familial aggregation of multimorbidity using a large multigenerational national registry from Sweden. They further test multimorbidity inheritance across genetic resemblance and disease clusters.

The study’s strengths include extensive analyses testing various aspects of the association between a composite multimorbidity outcome and family relations. The study is limited in the factors that are adjusted for, particularly those related to lifestyle, which makes it difficult to rule out the family environmental influences vs. genetic influences.

This article addresses a valuable question and has attempted to address it from various angles; however, I have some concerns about the premise of the multimorbidity outcome definition and other comments for clarification.

MAJOR COMMENTS:
- The study uses a simplistic composite definition of multimorbidity that may be too crude an outcome measure to answer the question the researchers are trying to answer. It is not entirely clear what the details were of the multimorbidity scoring, and would be helpful if it were elaborated on. However, it seems as though the researchers took any individual with 2 or more chronic conditions as the definition for the dichotomous multimorbidity outcome and compared the risk for a relative to have any 2+ conditions (which could be a completely different set of conditions from those their sibling, cousin, etc.). Several considerations that are contributing to this incongruence:
  - Some of the 45 diseases that were examined have known heritable characteristics, while others have less or no known heritable traits, and it may not be realistic to combine all these conditions into a single composite outcome measure, due to their heterogeneity.
  - Furthermore, I am not sure that using a count of diseases can really answer the question of genetic resemblance, because for example, even when counting, those who have 3 specific diseases may comprise a very different outcome than someone who has 3 completely different diseases.
  - It seems to me that a more nuanced characterization of multimorbidity may be needed to answer the question of whether there is familial aggregation.
Author response: Thanks for the helpful comments. However, our response to these comments is the same as the response to reviewer 1 above: We now realize that we have not explained matters clearly. We have now explained our findings much better thanks to reviewer comments. The PCA (principal component analysis) clearly shows that the clustering of disease in multimorbidity is specific and not random. Within the groups identified with PCA there are strong correlations. Thus, the following reviewer comment is not correct: “More problematic is the arbitrary definition of multimorbidity and conditions in which the genetic determinants are either unknown or pleiotropic.” Using the Barnett et al definition of comorbidity published in Lancet, we performed a PCA analysis as an agnostic method showing that disease clustering in multimorbidity is not random. Instead, disease clustering in multimorbidity is disease-specific. Moreover, these disease clusters are inherited in families. Thus, the clustering of disease in multimorbidity is disease-specific and not a random event. Having more than one related disease gives even high odds ratios for inheritance suggesting an increased risk for a genetic contribution when several related diseases are clustering in a family. We have therefore changed the discussion, abstract and title to clarify this.

The critics of diseases included in the multimorbidity score is not appropriate either. Our definition is based on a modified version by Barnett et al published in Lancet, which we have cited. In fact, Barnett et al specifically sought to include morbidities recommended as core for any multimorbidity measure by a systematic review, diseases in the quality and outcomes framework (QOF) of the UK general practice contract, and long-term disorders identified as important by NHS Scotland. Thus, our definition of multimorbidity is not arbitrary. Most importantly, using the PCA analysis, we show that disease clustering is not random - it is specific.

We have now added two references based on genomic data that indicates that multimorbidity has a genetic contribution. However, our paper is the first to show this using classical tools in genetic epidemiology such as study of the familial relative risk, here expressed as odds ratios in first-, second- and third-degree relatives. We have now clarified further in order to better explain the importance of our findings also for researchers not familiar with genetics epidemiology. We have also added more in the discussion about shared environment as suggested by the reviewer. However, the present high familial odds ratios are very unlikely to be explained by shared environment only. We have now added a reference for this. Moreover, the strong association between genetic resemblance and odds ratio indicates that a genetic component is likely. This is especially true for increased risk among third degree relatives that usually do not share a household. Thus, our findings are a strong indication of important genetic determinants.

Regarding the study period, we have carefully chosen the study period taking into account the limitations of the registers. An advantage is to include a time-period only using ICD-10. Regarding age dependence, we have adjusted for age in all models. Regarding the age, it is well-known that the genetic determinants for complex traits are stronger in young subjects, and it could therefore be an advantage to study young subjects to find a genetic contribution. It should be emphasized that the Swedish and Danish multigeneration registers have been shown in hundreds of studies to be an invaluable source of information about inheritance.
Whether or not the genetic determinants are unknown or pleiotropic is not relevant. The genetic determinants for complex traits remain to be fully determined. Moreover, genetic pleiotropy, i.e. when one gene influences two or more seemingly unrelated phenotypic traits, could be an important cause of the genetic basis for multimorbidity. Moreover, our PCA analysis shows that multimorbidity is disease-specific and that certain diseases cluster together in families and this it is not a random event which diseases that cluster together.

- In grouping MZ and DZ twins together in the model analyses, the researchers may not be disentangling the effect of genetic resemblance. Maybe the researchers want to consider separating this group into two groups or regroup the DZ under ‘siblings’ and study just the MZ twins separately? o If the authors decide to keep the groupings as they currently are, I would suggest adding the rationale for this grouping, despite the difference in genetic resemblance.

Author response: We did not have access to information about zygosity. We have written that now explicitly in the manuscript. However, it should be pointed out that we used Weinberg’s differential method for determining the mean genetic resemblance of included twins, which is stated in the manuscript: “The average genetic resemblance for twins, including both MZ and DZ twins, was determined to be 0.66 using Weinberg’s differential method. Each relative pair was assigned their genetic resemblance (i.e. 0.66 for twin pairs, 0.5 for sibling-pairs, 0.25 for half-sibling pairs, and 0.125 for cousin pairs).”

SPECIFIC COMMENTS:

Abstract:
The abstract can use some clarification.
There is no study design indicated in the abstract; the authors should add this.
The main outcome measure is the multimorbidity score and seems to fit better under the outcomes section.
Can the authors please clarify what was meant by the following sentence on p., lines 54-56 starting “There was a strong graded association….”?

Author response: Abstract is clarified and study design is included. Title is also changed.

Introduction:
A relevant study that I would suggest adding to the background on pleiotropy influencing multimorbidity is an article by Amell et al. 2018 (DOI:10.1038/s41598-018-34361-3). The authors can consider including this article’s findings as part of the background and even potentially as part of the discussion (in how the PCA results compare to Amell et al.’s findings) of their study.

Author response: Thanks. We have added this reference and a recent reference indicating a genetic contribution to multimorbidity at the genomic level.

It would help in clarifying the different parts of the study if the authors specified the objectives in more detail related to each of the main analyses conducted at the end of the introduction (e.g. determine association of multimorbidity in relatives, examine the trend of risk for multimorbidity across the genetic resemblance spectrum.

Author response. We have now clarified the aims.

Methods:
I did not see a study design indicated in the methods and would guess that this is a retrospective cohort study. It would be helpful to include the study design to provide the overall framing, intention and limitations.

**Author response: Study design is now indicated in the method also: nationwide retrospective cohort family study**

In the statistical analysis description on p.9, it would be helpful if the authors elaborated on the use of proband relative pairs related to the odds ratio regression analyses.

**Author response: We have further clarified this although this is a standard procedure in genetics epidemiology and a reference for that is given.**

Odds for multimorbidity – The set of covariates used in model adjustment seems to be fairly minimal. Have the authors considered additional covariates such as ethnicity/race (evidence as seen in an article by Kalgotra et al. 2020, https://doi.org/10.1038/s41598-020-70470-8), and possibly other surrogates of lifestyle factors, if they are not available in the existing registry database?

**Author response: In Sweden, we have no information about ethnicity. However, we have stated in the method section that all study subjects were Swedish-born and their parents were also Swedish-born.**

PCA – I am most familiar with principal component analysis as a method used to narrow down large sets of risk factors. Therefore, I am curious to better understand why the researchers chose to use PCA for the outcome being studied. Maybe the authors can elaborate on the rationale for using this analysis in the specific context?

**Author response. PCA is a standard statistical procedure just like odds ratio and Pearson and tetrachoric correlation that we have explained and referred to if the readers want to learn more. However, we now have added the following sentence: Thus, using PCA we aimed to determine if disease clustering in multimorbidity is disease-specific or not.**

Results:
Table 2 – It is not clear what the size of each group (N) is across the models for familial associations (e.g. twins, siblings, etc.). Can the authors please add Ns to the table?

**Author response: We have now added these numbers.**

Can the authors add a bit more to the results observations based on the results presented in Table 5?

**Author response: We have now explained the results in more detail.**

Discussion and limitations:
Some of the points in the discussion are repeated, e.g. third-degree relatives do not usually share a household environment. Consider removing repetition.

**Author response: repetitive statements are deleted.**

The authors do not dedicate much of the discussion to comparing results from their study to previous studies’ findings. It would be meaningful for the researchers to compare their findings to prior studies’ findings in order to provide more context on the value of their results. For example, how do the groupings of diseases found in the PCA compare to other groupings of related chronic diseases and pleiotropy in multimorbidity?
Author response: Thanks for this comment. We have now compared with more studies. Our study agrees with other studies identifying mental health conditions and cardio-metabolic conditions as the two most replicable groups of multimorbidity. We also replicate the grouping of chronic obstructive pulmonary disease with asthma. However, disease grouping is, of course, related to included diseases and selection of included study participants. Still, it is reassuring and a proof of concept that our two most important groups in the PCA (groups F1 and F2 in Tables 4 and 5) are the most two replicated groups in other multimorbidity studies, i.e. mental health conditions and cardio-metabolic condition.

p.16, line 58 – There seems to be a type-o.
Author response: Sentences are changed as a result adding more studies for comparison.

The relevance of the discussion on p.17 lines 39-58 about self-reported data vs. hospital diagnoses data was not entirely clear to me.

Author response: This is a strong advantage compared with family studies based on self-report.

VERSION 2 – REVIEW

REVIEWER 1 Cohen-Stavi, Chandra. Competing Interest: None
REVIEW RETURNED 28-Dec-2021

GENERAL COMMENTS SUMMARY:
Thank you for the opportunity to review the revised paper on this important and complex topic. The study aims do not seem to have changed from the previous version, which include determining the familial aggregation of multimorbidity using a large multigenerational national registry. The researchers further tested multimorbidity inheritance across genetic resemblance and disease clusters.

The study's strengths include a large population of family relations and extensive analyses testing various aspects of the association between multimorbidity and family relations. The study is limited in the multimorbidity definition and the factors that are adjusted for, particularly those related to lifestyle, which makes it difficult to rule out the family environmental influences vs. genetic influences.

Based on comparison with the earlier version of the paper and my initial review, it seems that the authors have made some revisions; however, the updates are not consistent. From the highlights at the beginning of the manuscript, the conclusions in the abstract, and the discussion, it seems that the researchers have tried to shift the focus toward the analyses examining disease-specific clusters and implications for genetic influence/familial aggregation. I expected from reading the end of the abstract and highlights that the main analyses would aim to evaluate the association between specific disease clusters and the extent that they are related to familial relationships. Yet, from my reading, I do not see the shift in the focus of the main analyses, which still evaluate multimorbidity defined crudely as a composite count of conditions. Yet, in the discussion, the main conclusions seem to draw largely from the PCA on disease clusters. It would seem more consistent with the updates to the manuscript if the researchers consider making the PCA analyses the main analyses and use the other analyses as...
supporting evidence that suggests correlations between familial aggregation and a more general measure of multimorbidity.

I believe that at least some of my previous comments on the article are still relevant, as outlined below.

MAJOR COMMENTS:

- The study uses a simplistic composite definition of multimorbidity that may be too crude an outcome measure to answer the question the researchers are trying to answer. The analyses examined individuals with 2 or more chronic conditions as the definition for the dichotomous multimorbidity outcome and compared the risk (ORs) for a relative to have any 2+ conditions (which could be a completely different set of conditions from those their sibling, cousin, etc.). Several considerations that are contributing to this incongruence:
  
  o Some of the 45 diseases that were examined have known heritable characteristics, while others have less or no known heritable traits, and it may not be realistic to combine all these conditions into a single composite outcome measure, due to their heterogeneity.
  
  o Furthermore, I am not sure that using a count of diseases can really answer the question of genetic resemblance, because for example, even when counting, those who have 3 specific diseases may comprise a very different outcome than someone who has 3 completely different diseases.

SPECIFIC COMMENTS:

Introduction:

It would help in clarifying the different parts of the study if the authors specified the objectives in more detail related to each of the main analyses conducted at the end of the introduction (e.g. determine association of multimorbidity in relatives, examine the trend of risk for multimorbidity across the genetic resemblance spectrum.

Methods:

Covariates – The set of covariates used in model adjustment seems to be fairly minimal. Have the authors considered additional covariates such as ethnicity/race (evidence as seen in an article by Kalgotra et al. 2020, https://doi.org/10.1038/s41598-020-70470-8)? The authors mention that lifestyle related factors were included but do not specify which lifestyle variables. Can the authors please specify which lifestyle factors were included?

PCA – I am most familiar with principal component analysis as a method used to narrow down large sets of risk factors. Therefore, I am curious to better understand why the researchers chose to use PCA for the outcome being studied. Maybe the authors can elaborate on the rationale for using this analysis in the specific context?
Results:

Table 4 – It’s not clear what the uniqueness column represents. Can the authors please explain the meaning of the uniqueness values in the methods or the results, like they do for the loading values?

Discussion and limitations:

It is a bit difficult to follow the discussion points, as they are not well organized. Specifically, the first paragraph of the discussion addresses many different points. It think it would be helpful to break up the first paragraph and clarify the organization of the topics based on the different analyses and related insights. Additionally, in the discussion of the disease clusters (PCA results), the authors note that their findings confirm previous multimorbidity studies’ findings on diseases that cluster together, however, it’s not clear what new knowledge is derived from the PCA clustered disease results. Maybe it is just a matter of making the implications about the connection between the disease clusters and the heredity/familial aggregation aspects clearer and more concrete, i.e. what does this mean for a practicing physician?

Some of the points in the discussion are repeated, e.g. cannot rule out environmental and lifestyle factors. I would suggest removing repetition.

The limitation mentioned on p.20 about not having information on DZ and MZ twins is a substantial limitation in being able to disentangle the effect of genetic resemblance. The authors should provide a bit more discussion about this limitation and the conclusions that can be drawn about genetic resemblance from their results.

p.17, lines 55-7 – I’m not sure I follow the conclusion about the age-stratified results indicating a genetic basis. Can the authors please elaborate and clarify this point?

p.19, line 46 – I’m not sure what is meant by the word pedigree. Consider a different word, maybe?

REVIEWER 2

Perera, Rafael; University of Oxford, Primary Care Health Sciences Competing Interest: I am interested in the area of multimorbidity and potential clustering of conditions. I have received funding from the NIHR (UK) to research this field which has covered part of my salary as well as that of members of my team.

REVIEW RETURNED 14-Jan-2022

GENERAL COMMENTS

Brief note made as part of screening: This is an interesting manuscript addressing the importance of genetics on multimorbidity (defined as having more than 2 chronic conditions, based on data from Sweden. There is consistency in their findings related to kinship (twins, siblings cousins) and the association with multimorbidity (MM). The dataset seems large and with relatively good follow up. It is likely to add to the discussion of ‘nature vs nurture’ for MM, the fact that they do not have any lifestyle factors
is clearly the biggest limitation but not sure they can do much about this. There are some clarifications required (e.g. when do they define a ‘case’ in their follow-up?) and their principal component analysis does not really add much. Nevertheless, most of these issues can be addressed as potential limitations.

Instead of doing a full stats report, I would like to recommend Professor Bennett to do this as he is an expert in MM, epidemiology, public health and genetics. He is based in the department of Population Health in the University of Oxford. I believe that his views would be ideal to shape this paper. If he doesn’t agree to do a full review, I would be happy to do a full stats review instead.

**REVIEWER 3**

Zhao, Xing-Ming; Zhangjiang Fudan International Innovation Center. Competing Interest: None

**REVIEW RETURNED**

27-Jan-2022

**GENERAL COMMENTS**

The authors investigated familial aggregation of multimorbidity, and gave some interesting conclusions. My main concerns with the paper can be found below.

1. As for the definition of multimorbidity, the authors may consider the time window that a patient has multiple diseases. For example, a patient may have a disease after another one, but the patient’s EHR will label the patient with two diseases. For me, this is not multimorbidity. In addition, the authors only use ICD-10 codes to define multimorbidity, where a patient has two or more ICD-10 codes is defined to suffer from multimorbidity. What will happen if a disease will be mapped to several ICD-10 codes?

2. The details about disease clusters should be given. I can’t understand how the clusters are achieved. I also don’t understand how to get the clusters with PCA. What will happen if you change to another clustering method, e.g. k-means?

3. The authors mentioned that familial aggregation of multimorbidity in Sweden is disease-specific. Although they show some disease clusters composed of multiple diseases with PCA, I don’t see what the authors want to deliver by “disease specific”. Do they mean families with specific diseases tend to suffer from multimorbidity?

**REVIEWER 4**

Bennett, Derrick. Competing Interest: None

**REVIEW RETURNED**

02-Feb-2022

**GENERAL COMMENTS**

1. It is not clear to me what the rationale was for computing tetrachoric and Pearson correlations between relatives. The authors should make it clear what the purpose of these are.

2. Do the logistic regression to compute the familial odds ratios take into account the relatedness (or correlation) in the analyses?

3. How good a proxy is educational achievement for lifestyle factors? The lack of adjustment for lifestyle factors such as smoking, alcohol and adiposity is a serious limitation of these analyses. Did the authors consider assessing the potential impact of residual confounding by exploring quantitative bias analysis
techniques such as those described by Lash (https://doi.org/10.1093/ije/dyu149)

4. On page 10 (lines 42-51) the authors state “Each relative pair was assigned their genetic resemblance (i.e. 0.66 for twin pairs, 0.5 for sibling pairs, 0.25 for half-sibling pairs, and 0.125 for cousin pairs). The same logistic regression analysis was conducted as described above, but with the inclusion of an interaction term between the genetic resemblance and multimorbidity in relatives.” How was the importance of the interaction term assessed?

5. On page 11 the details of how the PCA was performed, details of choice of type of rotation, and how the final PCs were selected should be reported in the main text.

6. The rationale for the use of PCA by the authors as opposed to other approaches should be described. Did the authors consider assessing the robustness of the PCA results by using an alternative method for MM such as Latent Class Analysis?

7. In Table 3 why have the authors reported the coefficients for the confounders? These are not really of interest in my view as the reader is only interested in the impact on the main exposure (ie. multimorbidity score).

8. Table 4 should give a description of the type of rotation used in the PCA in the footnote.

9. Table 5 should provide details of the number of participants that contributed to each of these clusters.

10. The authors state “The present study is the first one that aims to determine the familial aggregation of multimorbidity using clinical hospital diagnoses”. It is preferred that authors do not assume primacy of their research.

11. Page 13 line 11 the authors report a “96% CI”.

12. Typographical error Table 2 “Odds ratio for multimorbidity (≥ 2 score) according to multimorbidity score” the table shows scores up to five so “multimorbidity (≥ 2 score)” is incorrect.

REVIEWER 5
Riggare, Sara. Competing Interest: None

REVIEW RETURNED
14-Feb-2022

GENERAL COMMENTS
The manuscript describes an interesting study exploring potential hereditary components of multimorbidity. This review is focused on the patient perspective of the study and some comments and suggestions are given below.

General comments: This article is likely to be of interest to patients and the general public. I would encourage the authors to take a look at the manuscript once more with that in mind and try to clarify/simplify where possible. A few specific suggestions are mentioned below.
1. The summary boxes: I would suggest that the summary boxes are written in lay language, alternatively that boxes in lay language are provided in addition to the existing ones.
2. Abstract: Would it be possible to elaborate a bit on the seven identified disease clusters? If needed for word count reasons, maybe the methods & analysis be shortened.
3. Introduction: Maybe my understanding is limited but the use of the term “predisposition” to me implies causality. Is that intentional?
4. Patient and public involvement: Please elaborate on why there was no involvement.

**VERSION 2 – AUTHOR RESPONSE**

Reviewer: 1

Comments to the Author

**SUMMARY:**

Thank you for the opportunity to review the revised paper on this important and complex topic. The study aims do not seem to have changed from the previous version, which include determining the familial aggregation of multimorbidity using a large multigenerational national registry. The researchers further tested multimorbidity inheritance across genetic resemblance and disease clusters.

The study’s strengths include a large population of family relations and extensive analyses testing various aspects of the association between multimorbidity and family relations. The study is limited in the multimorbidity definition and the factors that are adjusted for, particularly those related to lifestyle, which makes it difficult to rule out the family environmental influences vs. genetic influences.

Based on comparison with the earlier version of the paper and my initial review, it seems that the authors have made some revisions; however, the updates are not consistent. From the highlights at the beginning of the manuscript, the conclusions in the abstract, and the discussion, it seems that the researchers have tried to shift the focus toward the analyses examining disease-specific clusters and implications for genetic influence/familial aggregation. I expected from reading the end of the abstract and highlights that the main analyses would aim to evaluate the association between specific disease clusters and the extent that they are related to familial relationships. Yet, from my reading, I do not see the shift in the focus of the main analyses, which still evaluate multimorbidity defined crudely as a composite count of conditions. Yet, in the discussion, the main conclusions seem to draw largely from the PCA on disease clusters. It would seem more consistent with the updates to the manuscript if the researchers consider making the PCA analyses the main analyses and use the other analyses as supporting evidence that suggests correlations between familial aggregation and a more general measure of multimorbidity.

I believe that at least some of my previous comments on the article are still relevant, as outlined below.

**Author response:** Thanks for the comments. We have commented and revised according to previous comments before, and we therefore will concentrate on the new comments. This is an explorative study which now is written explicitly to make this clear. Due to the explorative nature of the manuscript, we think that the present order is logic. First, we show that multimorbidity defined by a slightly modified method according to Barnett et al (published in Lancet) is inherited. After this we do an agnostic and hypothesis free PCA followed by factor analysis at the individual level and applicate these results on the family level in order to study family clustering i.e. inheritance at the family level for the most important groups. If we had found no inheritance of multimorbidity according to the slightly modified method by Barnett et al published in Lancet, we of course had not pursued any further analysis. As this is an explorative study, we used the definition by Barnett et al in Lancet, though slightly modified.
Other definitions of multimorbidity and included diseases could have been chosen but is not essential. The important issue is that we have a rather large number of common chronic (at least to a major part) non-communicable diseases (NCDs) as in the modified Barnett multicont method used and that the PCA followed by factor analysis method for studying clustering is hypothesis free and agnostic. Indeed, the combination of the modified multicont method by Barnett et al and PCA followed by factor analysis method resulted in that we identified the two most important multimorbidity diseases group identified by most other researchers independent of used methods at the individual level: i.e. the cardiometabolic cluster and the psychiatric cluster. We also identified another common group described by other i.e. the allergic cluster. This is a strong proof of concept.

We do not answer previous questions in detail because we have done that before. However, Barnett et al that published the paper in Lancet and many other researchers would not agree with that the reviewer 1 writes a “simplistic definition of multimorbidity”. To cite the original article by Barnett et al published in the Lancet: “No standard approach for the measurement of multimorbidity exists, and selection and definition of morbidities to include is inevitably partly subjective and dependent on the data available. We specifically sought to include morbidities recommended as core for any multimorbidity measure by a systematic review,” diseases in the quality and outcomes framework (QOF) of the UK general practice contract, and long-term disorders identified as important by NHS Scotland.

Thus, as Barnett et al write in their Lancet article there is no standard definition of multimorbidity. However, we are satisfied with the method used by Barnett et al in Lancet because this method based on a systematic review, i.e. best possible scientific evidence. Of course, other definitions of multimorbidity could be used but we preferred to use a method based on empiric experience already published in Lancet by Barnett et al. We modified it slightly in order to include only non-infectious disorders and ICD10 codes used in Swedish national registers. Moreover, we now also cite a systematic review that has shown that cardiovascular and metabolic diseases, mental health problems, and allergic diseases were identified as important disease clusters whatever method used for studying disease clustering and multimorbidity patterns.

MAJOR COMMENTS:

Author response: We have answered to these questions before. We do not answer previous questions in detail because we have done that before.

- The study uses a simplistic composite definition of multimorbidity that may be too crude an outcome measure to answer the question the researchers are trying to answer. The analyses examined individuals with 2 or more chronic conditions as the definition for the dichotomous multimorbidity outcome and compared the risk (ORs) for a relative to have any 2+ conditions (which could be a completely different set of conditions from those their sibling, cousin, etc.). Several considerations that are contributing to this incongruence:

Author response: It is not possible to use all diseases in the ICD catalog because we have not access to all. We must prioritize the most important and common non-infectious chronic disorders, which already have been done by Barnett et al in an excellent article published in Lancet. We believe this is well enough for use in an explorative study and we believe it is not a too simplistic view of multimorbidity that Barnett et al published in Lancet. Moreover, in table 2 (based on all 45 diseases) and supplementary tables S12-S20 (based on PCA followed by factor analysis that identified F1-F9 disease clusters), we have presented results that show that the number of diseases in the proband affect the risk for multimorbidity in relatives in a dose-response fashion. The famous epidemiologist Sir Bradford Hill probably also should be happy for these results. One of Bradfords Hill’s criteria for causation is Biological gradient (dose-response relationship): Greater exposure should generally lead to greater incidence of the effect. In fact, here we have two biological gradients in the same tables! One gradient for number of diseases in proband relative and one for genetic resemblance! For researchers who still think Bradfords Hill’s criteria are important this is strong indication for a genetic cause of multimorbidity. If we continue with Bradfords Hill’s criteria it is worth also to mention strength (effect size): A small association does not mean that there is not a causal effect, though the
larger the association, the more likely that it is causal. In the present study we have high odds ratios i.e. high effect size. The second criterium consistency (reproducibility) is also very important. Here we have consistent finding with regards to genetic resemblance.

Some of the 45 diseases that were examined have known heritable characteristics, while others have less or no known heritable traits, and it may not be realistic to combine all these conditions into a single composite outcome measure, due to their heterogeneity.

Author response: It is not possible to include only disorders with a known specific degree of heritability. Moreover, even if a disorder has a certain degree of heritability, it is possible that its contribution to inheritance of multimorbidity is different. In different families different diseases could cluster together. That is way we did the PCA followed with factors analysis in order to study familial clustering of different disease groups.

Furthermore, I am not sure that using a count of diseases can really answer the question of genetic resemblance, because for example, even when counting, those who have 3 specific diseases may comprise a very different outcome than someone who has 3 completely different diseases.

Author response: We have answered this question before. We do not try to answer a question of genetic resemblance. The degree of genetic resemblance among first-, second-, and third-degree relatives is a well-established general knowledge in genetics and medicine: siblings (50%), half-siblings (25%) and cousins (12.5%) exhibits. The average genetic resemblance for twins, including both MZ and DZ twins, was determined to be 0.66 using Weinberg’s differential method. This is described in the manuscript. The second question is answered by applying our PCA followed by factor analysis on the individual level on the familial risk. It is of course important what diseases that clusters together, which we have shown (Table 5 and supplementary tables S12-S20).

SPECIFIC COMMENTS:

Introduction:

It would help in clarifying the different parts of the study if the authors specified the objectives in more detail related to each of the main analyses conducted at the end of the introduction (e.g. determine association of multimorbidity in relatives, examine the trend of risk for multimorbidity across the genetic resemblance spectrum.

Author response: This is already done in the previous revision.

Methods:

Covariates – The set of covariates used in model adjustment seems to be fairly minimal. Have the authors considered additional covariates such as ethnicity/race (evidence as seen in an article by Kalgotra et al. 2020, https://doi.org/10.1038/s41598-020-70470-8)? The authors mention that lifestyle related factors were included but do not specify which lifestyle variables. Can the authors please specify which lifestyle factors were included?

Author response: Only families with Swedish born parents were included. Thus, different ethnicities were not included. This is well described in the method section and is already stated in the manuscript. We specify in the method and discussion as a limitation that lifestyle factors were not available. Instead, we adjusted for the level of education achievement as a marker for socioeconomic status and lifestyle related factors. We have recognized this limitation.

PCA – I am most familiar with principal component analysis as a method used to narrow down large sets of risk factors. Therefore, I am curious to better understand why the researchers chose to use PCA for the outcome being studied. Maybe the authors can elaborate on the rationale for using this analysis in the specific context?
Author response: We have now explained the PCA followed by factor analysis in more detail in the abstract and in extensively in the method section.

Results:

Table 4 – It’s not clear what the uniqueness column represents. Can the authors please explain the meaning of the uniqueness values in the methods or the results, like they do for the loading values?

Author response: This is explained in footnote: Uniqueness is equal to 1 – communality. The method is now explained in more detail.

Discussion and limitations:

It is a bit difficult to follow the discussion points, as they are not well organized. Specifically, the first paragraph of the discussion addresses many different points. I think it would be helpful to break up the first paragraph and clarify the organization of the topics based on the different analyses and related insights. Additionally, in the discussion of the disease clusters (PCA results), the authors note that their findings confirm previous multimorbidity studies’ findings on diseases that cluster together, however, it’s not clear what new knowledge is derived from the PCA clustered disease results. Maybe it is just a matter of making the implications about the connection between the disease clusters and the heredity/familial aggregation aspects clearer and more concrete, i.e. what does this mean for a practicing physician?

Author response. The section has already previously been rewritten.

Some of the points in the discussion are repeated, e.g. cannot rule out environmental and lifestyle factors. I would suggest removing repetition.

Author response: Repetitions are already removed.

The limitation mentioned on p.20 about not having information on DZ and MZ twins is a substantial limitation in being able to disentangle the effect of genetic resemblance. The authors should provide a bit more discussion about this limitation and the conclusions that can be drawn about genetic resemblance from their results.

Author response: Reviewer is right that lack of information about DZ and MZ limits the interpretation of twin data. However, the average genetic resemblance for twins, including both MZ and DZ twins, was determined to be 0.66 using Weinberg’s differential method. Moreover, studying a disease in first-, second-, and third-degree relatives is a classic and strong study design in genetic epidemiology. Cousins do not share household but 12.5% of genes, i.e. 12.5% genetic resemblance. In fact, the present study design is one of several useful study designs used in genetic epidemiology to disentangle nurture from nature.

p.17, lines 55-7 – I’m not sure I follow the conclusion about the age-stratified results indicating a genetic basis. Can the authors please elaborate and clarify this point?

Author response: It is well known in genetic epidemiology that genetic factors often are more important at younger age for complex traits. Higher familial risks among young individuals could be an indication of a genetic cause. We now cite the articles in Science by Lander from 1994 and the article by Burton et al. in Lancet that are excellent reviews of complex traits and genetic epidemiology.

p.19, line 46 – I’m not sure what is meant by the word pedigree. Consider a different word, maybe?

Author response. Pedigree is a common word in genetics meaning family tree. However, we now do not use the word pedigree more in the manuscript. Instead, we use the more lay word family instead that is easier to understand.
Reviewer: 2

Comments to the Author

Brief note made as part of screening: This is an interesting manuscript addressing the importance of genetics on multimorbidity (defined as having more than 2 chronic conditions, based on data from Sweden). There is consistency in their findings related to kinship (twins, siblings, cousins) and the association with multimorbidity (MM). The dataset seems large and with relatively good follow up. It is likely to add to the discussion of ‘nature vs nurture’ for MM. The fact that they do not have any lifestyle factors is clearly the biggest limitation but not sure they can do much about this. There are some clarifications required (e.g. when do they define a ‘case’ in their follow-up?) and their principal component analysis does not really add much. Nevertheless, most of these issues can be addressed as potential limitations.

Instead of doing a full stats report, I would like to recommend Professor Bennett to do this as he is an expert in MM, epidemiology, public health and genetics. He is based in the department of Population Health in the University of Oxford. I believe that his views would be ideal to shape this paper. If he doesn’t agree to do a full review, I would be happy to do a full stats review instead.

Author response: Thanks for the positive response and suggestions. The PCA followed by factor analysis confirms other studies but the most important reason for doing the PCA followed by factor analysis was that we wanted to explore why multimorbidity clusters in families. Using the PCA followed by factor analysis at the individual level and applying these groups at the family level we demonstrate that certain multimorbidity diseases clusters in families. Regarding lifestyle factors we bypassed this problem with the inclusion of cousins and studying relatives of different degree of genetic resemblance. For instance, if the disease would not be genetic at all there should be no strong correlation between genetic resemblance and familial odds ratio. Moreover, cousins in Sweden do very seldom share household. Moreover, as pointed out by Khoury et al., even in the face of a complete correlation in exposure among first-degree relatives, environmental risk factors with relative risks of <10 yield modest familial relative risks (1 to 2) and low recurrence risks, which suggest that the high risk of disease-specific multimorbidity has an important genetic contribution. All this is discussed in the discussion section. Moreover, as suggested by reviewer we now include E-values with lower 95% CI limits in Tables 2, 3 and 5. To determine whether the observed odds ratios could be explained by unobserved confounders rather than causal effects, E-values were calculated according to VanderWeele and Ding (2017). In the method section E-values are explained: "To determine whether the observed odds ratios could in effect be explained by unobserved confounders rather than causal effects, E-values were calculated. E-value is defined as the minimum strength of association, that an unmeasured confounder would need to have with both the predictor and the outcome to fully explain a specific predictor-outcome association. A large E-value indicates that considerable unmeasured confounding would be needed to explain an effect estimate. A small E-value indicates little unmeasured confounding would be needed to explain an effect estimate. The lower 95% confidence interval (CI) limit for the E-values were also calculated."

Reviewer: 3

Comments to the Author

The authors investigated familial aggregation of multimorbidity, and gave some interesting conclusions. My main concerns with the paper can be found below.

(1) As for the definition of multimorbidity, the authors may consider the time window that a patient has multiple diseases. For example, a patient may have a disease after another one, but the patient’s EHR will label the patient with two diseases. For me, this is not multimorbidity. In addition, the authors only use ICD-10 codes to define multimorbidity, where a patient has two or more ICD-10 codes is defined to suffer from multimorbidity. What will happen if a disease will be mapped to several ICD-10 codes?
Author response: This is of course a limitation but most of the conditions are chronic disorders that the patient must live with for a long time or the rest of the life. We now mention this possible limitation in the discussion. Multimorbidity is commonly defined as two or more diseases and this definition of two or more diseases is used for instance in the paper by Karen Barnett published in Lancet that we cite. However, in Table 2 and Supplementary Tables S12-S20 we illustrate that the risk of two or more diseases is dependent on the number of diseases in the proband relatives. The risk of multimorbidity was in a dose dependent way directly related to risk of multimorbidity among twins, siblings, half-siblings, and cousins. Here we must cite Sir Bradford Hill. One of Bradfords Hill's criteria for causation is biological gradient (dose-response relationship): Greater exposure should generally lead to greater incidence of the effect. In fact, here we have two biological gradients in the same table. One gradient for number of diseases in probands and one for genetic resemblance. For researchers who still think Bradfords Hill's criteria are important this is a strong indication for a genetic cause of multimorbidity. If we continue with Bradfords Hill's criteria it is worth to mention strength (effect size): A small association does not mean that there is not a causal effect, though the larger the association, the more likely that it is causal. In the present study we have high odds ratios i.e. high effect size. The second criterium consistency (reproducibility) is also very important. Here we have consistent finding with regards to genetic resemblance. Regarding temporality this is not so important for genetic disorders because we are born with our genes. We have therefore not studied the temporal order of the occurrence of diseases, though of course this would be an interesting and extensive study. However, we recognize this limitation in the discussion. Regarding ICD codes in Sweden 85-95% are correct for most studied diagnoses according to many validation studies of the Swedish registers (reference 15).

(2) The details about disease clusters should be given. I can’t understand how the clusters are achieved. I also don’t understand how to get the clusters with PCA. What will happen if you change to another clustering method, e.g. k-means?

Author response: We now in deeper detail have explained the PCA followed by factor analysis method used. The strength of the PCA followed by factor analysis method is that this is an agnostic and hypothesis free method. Using the PCA followed by factor analysis method we identified different diseases clustering at the individual levels. Here we confirm many other studies of disease clustering at the individual level, which is a strength of our study especially if we think of the great epidemiologist Sir Bradford Hill. Consistency with other studies is important. We have now applied the k-means method and we have obtained certain similarities. However, there is a built in limitations in the k-means method. Repeated analysis gives different clustering unlike the PCA followed by factor analysis method that gives the same results. We therefore have chosen not to present the results of k-means method. Another method is the latent class analysis [LCA]. Due to the large test material we only managed to do the analysis in a subset of the cohort, and though certain similarities were obtained, we do not present these results that is not valid for the entire cohort. However, we feel that PCA followed by factor analysis fulfills several important issues for an explorative study. The method is an agnostic and robust method also in a large nationwide sample like this, with high reproducibility.

(3) The authors mentioned that familial aggregation of multimorbidity in Sweden is disease-specific. Although they show some disease clusters composed of multiple diseases with PCA, I don’t see what the authors want to deliver by “disease specific”. Do they mean families with specific diseases tend to suffer from multimorbidity?

Author response: We take away disease specific from the title as suggested by reviewer. Of course, it is true certain diseases clusters more often together, which was shown at the individual level first (Table 4) and confirmed at the family level thereafter (Table 5 and supplementary tables S12-S20).

Reviewer: 4

Comments to the Author
1. It is not clear to me what the rationale was for computing tetrachoric and Pearson correlations between relatives. The authors should make it clear what the purpose of these are.

**Author response:** Thanks for this comment. We delete these correlations. This makes it easier for the reader.

2. Do the logistic regression to compute the familial odds ratios take into account the relatedness (or correlation) in the analyses?

**Author response:** No relatedness was considered. We wanted to see if the OR varies according to degree of genetic resemblance. However, in a logistic regression analysis with inclusion of all different proband-relative pairs we included an interaction term between the genetic resemblance and multimorbidity in relatives. This term was highly significant.

3. How good a proxy is educational achievement for lifestyle factors? The lack of adjustment for lifestyle factors such as smoking, alcohol and adiposity is a serious limitation of these analyses. Did the authors consider assessing the potential impact of residual confounding by exploring quantitative bias analysis techniques such as those described by Lash (https://doi.org/10.1093/ije/dyu149)

**Author response:** Thanks for this comment. We have recognized this limitation. However, regarding shared familial lifestyle factors we bypassed this problem with the inclusion of cousins and studying relatives of different degree of genetic resemblance. For instance, if the disease would not be genetic at all there should be no strong correlation between genetic resemblance and familial odds ratio. Moreover, cousins in Sweden do very seldom share household. Moreover, as pointed out by Khoury et al,26 even in the face of a complete correlation in exposure among first-degree relatives, environmental risk factors with relative risks of <10 yield modest familial relative risks (1 to 2) and low recurrence risks, which suggest that the high risk of disease-specific multimorbidity has an important genetic contribution. All this is discussed in the discussion section. As suggested by reviewer we now include E-values with lower 95% CI limits in Tables 2, 3 and 5.

4. On page 10 (lines 42-51) the authors state “Each relative pair was assigned their genetic resemblance (i.e. 0.66 for twin pairs, 0.5 for sibling-pairs, 0.25 for half-sibling pairs, and 0.125 for cousin pairs). The same logistic regression analysis was conducted as described above, but with the inclusion of an interaction term between the genetic resemblance and multimorbidity in relatives.” How was the importance of the interaction term assessed?

**Author response:** As written in the end of the statistical section in the methods we used the 0.05 level for significance: Statistical significance was set at p<0.05. As written in the result section the interaction term was highly significant: OR=3.43; p <0.001.

5. On page 11 the details of how the PCA was performed, details of choice of type of rotation, and how the final PCs were selected should be reported in the main text.

**Author response:** More detail of the PCA followed by factor analysis is now presented in the abstract and method section.

6. The rationale for the use of PCA by the authors as opposed to other approaches should be described. Did the authors consider assessing the robustness of the PCA results by using an alternative method for MM such as Latent Class Analysis?
Another study has shown that the two most replicable multimorbidity profiles were mental health conditions and cardio-metabolic conditions.31 With PCA followed by factor analysis we identified the same main groups, and we therefore did not consider other methods. However, we now recognize this limitation in the discussion section. If judged to be necessary we of course could try other methods, though this is not the primary focus of our study. The study aim was to explore the inheritance of multimorbidity. We now have applied the k-means method and we obtained certain similarities. However, there is a build in limitations in the k-means method. Repeated analysis gives different clustering unlike the PCA followed by factor analysis method that gives the same results. We therefore have chosen not to present the results of the k-means method. Another method is the latent class analysis [LCA] as suggested by reviewer. Due to the large test material we only managed to do the analysis in a subset of the cohort, and though certain similarities were obtained, we do not present these results as they are not valid for the whole cohort. However, we feel that PCA followed the by factor analysis fulfills several important issues for an explorative study: an agnostic, robust method also in a large nationwide sample like this with high reproducibility. Moreover, a systematic review has shown that cardiovascular and metabolic diseases, mental health problems, and allergic diseases were identified as important disease clusters whatever method used for studying disease clustering and multimorbidity patterns (reference 42).

7. In Table 3 why have the authors reported the coefficients for the confounders? These are not really of interest in my view as the reader is only interested in the impact on the main exposure (ie. multimorbidity score).

Author response: We reported these variables for transparency, but reviewer is right. This is not necessary. We therefore delete the results of these variables.

8. Table 4 should give a description of the type of rotation used in the PCA in the footnote.

Author response: The type of rotation used in the PCA is now described in the footnote: “An oblique promax rotation was used on the correlation matrix of tetra-choric correlations between the 45 diagnoses in patients to identify disease clusters.”

9. Table 5 should provide details of the number of participants that contributed to each of these clusters.

Author response. Details of the numbers of the participants that contributed to each of these groups are now presented in Supplementary Tables S22-S24. We add this information in the result section under additional information.

10. The authors state “The present study is the first one that aims to determine the familial aggregation of multimorbidity using clinical hospital diagnoses”. It is preferred that authors do not assume primacy of their research.

Author response. Thanks for this comment. We delete “the first”. This is for other to judge.

11. Page 13 line 11 the authors report a “96% CI”.

Author response: Sorry for the typo. This is now corrected to 95% CI.

12. Typographical error Table 2 “Odds ratio for multimorbidity (≥ 2 score) according to multimorbidity score” the table shows scores up to five so “multimorbidity (≥ 2 score)” is incorrect.
Author response: Sorry for being unclear. It should be odds ratio for multimorbidity (≥ 2 score) in relatives according to number of diseases in proband. Table text is corrected.

Reviewer: 5

Comments to the Author
The manuscript describes an interesting study exploring potential hereditary components of multimorbidity. This review is focused on the patient perspective of the study and some comments and suggestions are given below.

General comments: This article is likely to be of interest to patients and the general public. I would encourage the authors to take a look at the manuscript once more with that in mind and try to clarify/simplify where possible. A few specific suggestions are mentioned below.

Author response: Thanks for the appreciation of the study.

1. The summary boxes: I would suggest that the summary boxes are written in lay language, alternatively that boxes in lay language are provided in addition to the existing ones.

Author response: Thanks for this opinion. We have explained the word multimorbidity now and avoided other non-lay words.

2. Abstract: Would it be possible to elaborate a bit on the seven identified disease clusters? If needed for word count reasons, maybe the methods & analysis be shortened.

Author response: We now elaborate this more in the abstract and mention the first three disease clusters F1-F3 in the abstract.

3. Introduction: Maybe my understanding is limited but the use of the term "predisposition" to me implies causality. Is that intentional?

Author response: Predisposed is changed to associated: "Except for old age, female sex and low socioeconomic status appear to be associated with multimorbidity. Other suggested associated factors for multimorbidity are smoking, physical inactivity, being overweight, as well as hypertension and low level of education in men."

4. Patient and public involvement: Please elaborate on why there was no involvement.

Author response: The following sentence is added: "Statistics Sweden and the National Board of Health and Welfare provided all data."

VERSION 3 – REVIEW

| REVIEWER 2 | Perera, Rafael; University of Oxford, Primary Care Health Sciences. Competing Interest: I have worked in the area of multimorbidity for the last 5 years and I am interested in clustering of conditions. Have received funding for my salary and that of members of my group from the NIHR (UK). |
| REVIEW RETURNED | 11-Oct-2022 |

GENERAL COMMENTS
The authors present an analysis of the hereditary impact of multimorbidity. To do this, they use data from 'young' people (average age at final follow-up 32 years) in Sweden and their family. They model the odds of having 2 or more conditions (out of a list of 45 possible) based on the identification of 0, 1, 2, 3, 4, 5+ conditions in their family (separate analyse for twins, siblings, half-
siblings, cousins). They quantify that there is a hereditary element to this and then go on to carry out a factor analysis which they then use as definitions for nine different clusters of diseases. They conclude that in this group multimorbidity is hereditary and particularly so in some of these clusters of conditions. There are some critical issues in their analysis that need to be corrected before we can define if their results are valid.

Major issues:

1. The critical problem in their analysis is that in their logistic regression models, the assumption of independence does not hold. This is true for all results presented and therefore we are unable to determine the validity of the results.

They authors are creating individual logistic regression models for each one of the familial associations. They are using an individual more than once (by definition) in each one of these datasets. In some scenarios (siblings and cousins) an individual and its relationship will appear multiple times. This violates one of the main assumptions of logistic regression and the impact on the estimates is likely to be to provide artificially high precision (they have less data than actually presented). I am unclear how this might also affect the point estimates of each of these analyses. This modelling approach is used throughout the paper which means that all results arising from this will need re-doing. Related to this, their Pearson correlations for degree of genetic resemblance are also calculated using a similar approach violating the independence assumption of the datapoints. Generally, if they decide to use simple logistic regression, the units of analyses need to be independent (change to the data structure). Alternatively, they would need to think of a more complex/alternative models if they want to use this dataset/data-structure.

2. In their response to the reviewers’ comments, they have stated that alternative clustering strategies that they tried generated different clusters. Given this, it is important that they include in their manuscript that as part of a sensitivity analysis they tried several other methods to determine their clusters and that, although there were some similarities, in general the clusters identified were not consistent across the methods.

Given the lack of consistency across the clusters identified, the ‘what this study adds’ first bullet-point - “The present study shows that disease-specific clusters of multimorbidity aggregates in families depending on degree of relatedness.” is removed (as they have already removed this from their title). It is unclear that this is really showing link to disease-specific clusters. Same thing with the last bulletpoint: “The disease clustering in multimorbidity is disease-specific and not random.” This is not shown as it is not robust to the methods used.

3. The authors need to explain how they decided on the number of Factors in their Factor analysis. This is critical as Factor analysis based on different number of factors could have come up with completely different groupings. Related to this, all conditions have neatly been allocated to one of the clusters (with no overlap). Please clarify if this is part of the method used and if so, how appropriate it is to the clinical context.
4. The population used in this analysis is likely to overestimate the effect of genetic predisposition to MM. This is because the average age at end of study was 32 with a prevalence of MM of ~16%. This has actually been shown by the authors in their stratified analysis. It is very likely that this approach in older cohorts would show a significant reduction in hereditary for MM. In fact, it might be that they are looking at only one type of MM, the type that can be inherited as opposed to the one that is defined more by lifestyle. Their Discussion in this topic could be more robust regarding this critical limitation (maybe even include as part of the title?).

**VERSION 3 – AUTHOR RESPONSE**

Reviewer: 2

Comments to the Author

bmjmed-2021-000070.R1 Familial aggregation of multimorbidity in Sweden: a national explorative family study

The authors present an analysis of the hereditary impact of multimorbidity. To do this, they use data from ‘young’ people (average age at final follow-up 32 years) in Sweden and their family. They model the odds of having 2 or more conditions (out of a list of 45 possible) based on the identification of 0, 1, 2, 3, 4, 5+ conditions in their family (separate analyse for twins, siblings, half-siblings, cousins). They quantify that there is a hereditary element to this and then go on to carry out a factor analysis which they then use as definitions for nine different clusters of diseases. They conclude that in this group multimorbidity is hereditary and particularly so in some of these clusters of conditions.

There are some critical issues in their analysis that need to be corrected before we can define if their results are valid.

Major issues:

1. The critical problem in their analysis is that in their logistic regression models, the assumption of independence does not hold. This is true for all results presented and therefore we are unable to determine the validity of the results.

They authors are creating individual logistic regression models for each one of the familial associations. They are using an individual more than once (by definition) in each one of these datasets. In some scenarios (siblings and cousins) an individual and its relationship will appear multiple times. This violates one of the main assumptions of logistic regression and the impact on the estimates is likely to be to provide artificially high precision (they have less data than actually presented). I am unclear how this might also affect the point estimates of each of these analyses. This modelling approach is used throughout the paper which means that all results arising from this will need re-doing. Related to this, their Pearson correlations for degree of genetic resemblance are also calculated using a similar approach violating the independence assumption of the datapoints. Generally, if they decide to use simple logistic regression, the units of analyses need to be independent (change to the data structure). Alternatively, they would need to think of a more complex/alternative models if they want to use this dataset/data-structure.

Author response: Thanks for this comment. We have deleted all the Pearson correlations as these are not necessary for the paper. The double-entry approach is a common procedure in genetics and new references are added (references 22 and 23 were added). The point
estimates are not affected by double-entry. We have added the following sentence and references: "The double entry approach is a common procedure in genetics."

The reviewer is right in that confidence intervals are biased. However, in large materials like this, the differences are minor. However, we now use the "variance covariance (vce) cluster" method in STATA to adjust for non-independence, i.e. the intra-class correlation between (the two) entries belonging to the same class (i.e. family-identifier). This gives slightly wider confidence intervals, but the results are not changed to any major degree. The results obtained with this technique are also (in this case) very much consistent with the results obtained with the standard technique to double the variances obtained from the double entry approach in order to correct for the dependence between the two reversed-order entries. These latter results are thus not reported. Thus, all regression models are recalculated using the "variance covariance (vce) cluster" method in the Tables and Supplementary Material. In the statistical section, we added the following: "The ambiguity in unselected samples as to which twin’s, sibling’s, half-sibling’s or cousin’s trait should be used as the dependent, and which as the independent variable, is frequently resolved by using double entry. Each twin, sibling, half-sibling, or cousin is entered twice in the data, and each member of a twin, sibling, half-sibling, or cousin pair provides once the dependent and once the explanatory variable. While the consistency of the regression estimates for heritability and environmental influences is not affected by double-entry, the standard errors of the coefficients are biased and need to be adjusted. In the present study in the logistic regression models, we used the "variance covariance (vce) cluster" method using STATA, which calculated robust standard errors using families as clusters. The "variance covariance (vce) cluster" method specifies that the standard errors allow for intragroup correlation, relaxing the usual requirement that the observations be independent. The results obtained with this technique are also (in this case) very much consistent with the results obtained with the standard technique to double the variances obtained from the double entry approach in order to correct for the dependence between the two reversed-order entries. These latter results are thus not reported."

2. In their response to the reviewers’ comments, they have stated that alternative clustering strategies that they tried generated different clusters. Given this, it is important that they include in their manuscript that as part of a sensitivity analysis they tried several other methods to determine their clusters and that, although there were some similarities, in general the clusters identified were not consistent across the methods.

Author response: We now mention and discuss this in the limitations section. We have added the following to the discussion: "We believe the use of PCA for studying disease clusters is an advantage over using K-means. The k-means algorithm requires some initialization of the centroid positions. For most algorithms, these centroids are randomly initialized with some method such as the Forgy method or random partitioning, which means that repeated iterations of the algorithm can converge to vastly different results. We observed this also and therefore we did not use k-means because the clustering was not reproducible with k-means, while PCA gave identical results. Nevertheless, there were some similarities with our grouping results between k-means and PCA. We also tried Latent class analysis (LCA) for cluster analysis but due to the large number, the results did not converge for the LCA analysis. It is a known limitation of LCA for cluster analysis that it is computationally expensive, which might be inconvenient with very large datasets. Moreover, selecting the number of clusters used is a challenging task involving the inevitable subjectivity of analytical choices. However, with the PCA method we reproduced the two most commonly reported clusters of multimorbidity (mental health group and cardiometabolic disease), which shows the strength of the used PCA method."
Given the lack of consistency across the clusters identified, the ‘what this study adds’ first bullet-point - “The present study shows that disease-specific clusters of multimorbidity aggregates in families depending on degree of relatedness.” is removed (as they have already removed this from their title). It is unclear that this is really showing link to disease-specific clusters. Same thing with the last bulletpoint: “The disease clustering in multimorbidity is disease-specific and not random.” This is not shown as it is not robust to the methods used.

Author response: We do not agree completely with this comment. We have consistency of the cluster identified using the PCA method. We get the same result even with repeated analysis unlike when using the k-means method. Moreover, as mentioned above we have identified the two most identified multimorbidity clusters in the literature according to the cited systematic review on the topic (reference 36). However, we have changed the three sentences as suggested by the reviewer in the bullet points:

- The present study shows that multimorbidity aggregates in families depending on the degree of relatedness.

- A genetic component of multimorbidity is indicated, although familial lifestyle factors might contribute.

- Certain disease clusters are inherited in families.

3. The authors need to explain how they decided on the number of Factors in their Factor analysis. This is critical as Factor analysis based on different number of factors could have come up with completely different groupings. Related to this, all conditions have neatly been allocated to one of the clusters (with no overlap). Please clarify if this is part of the method used and if so, how appropriate it is to the clinical context.

Author response: We used visual inspection. In the result section this is already described under the headings Principal-component analysis (PCA) followed by factor analysis but we now clarify this: "Number of diseases could be reduced into nine different disease clusters (Supplementary Table S10 S9 and Supplementary Figure S1). Visual inspection was used for finding a point at which the amount of variance explained by subsequent principal component drops off (Supplementary Figure S1). After nine factors, the eigenvalues were levelled off (elbow) and nine factors were therefore extracted."

This is part of the method used. We now state: "As part of the method all conditions were allocated to one of the nine clusters."

This is indeed relevant for the clinical context. Certain diseases occur more often together also in the clinic. One should not forget that all data are from all hospitals in Sweden and reflect the clinical reality and are therefore of the utmost clinical relevance. This is a
A nationwide population study covering all hospitals in and outpatient treatment in Sweden for the included individuals.

4. The population used in this analysis is likely to overestimate the effect of genetic predisposition to MM. This is because the average age at end of study was 32 with a prevalence of MM of ~16%. This has actually been shown by the authors in their stratified analysis. It is very likely that this approach in older cohorts would show a significant reduction in hereditary for MM. In fact, it might be that they are looking at only one type of MM, the type that can be inherited as opposed to the one that is defined more by lifestyle. Their Discussion in this topic could be more robust regarding this critical limitation (maybe even include as part of the title?).

Author response: Thanks for this comment. Age dependency for familial odds ratios is an indication of a genetic cause. The reviewer is correct that the mean age is young, and inheritance is age dependent indicating that multimorbidity is a complex trait (i.e. polygenic disorder) (Supplementary Tables S5-S8). This is an important signature of a polygenic condition (reference 31 in Science). Complex traits or polygenic diseases have higher heritability at younger than older ages. This is a well-known and typical finding for complex traits. We have done age-stratified analysis and shown this age dependence (Supplementary Tables S5-S8), which further strengthens that genetic factors influence the inheritance of multimorbidity. To study genetic effects, as in this explorative study, it is therefore an advantage to study younger individuals (reference 31). We now mention this age dependency in the abstract and the discussion sections. We discuss that as older age genetic effects are likely to become weaker and acquired factors more important for complex traits in general.

The oldest individuals are aged 68 so it is hard in the title to state that the study is about only young and middle-aged individuals though the majority are young or middle-aged but if the statistical reviewer thinks this is important we can of course do that.

In the discussion in the limitation section we have already written the following: "However, heritability in complex traits is age dependent and young age is an advantage in an explorative study of heredity,31 including multimorbidity inheritance." We now add: "The observed age dependence in the present study of the familial odds ratio indicates a genetic cause and that multimorbidity is a complex trait.31 At older ages it is expected that acquired factors of multimorbidity will be relatively more important."

In the abstract we have added this in the result section: "The familial ORs were highest at a young age." 

In the conclusions in the abstract is now written the following: In Sweden, certain diseases cluster together in families especially in younger individuals, which indicates a genetic component.
**GENERAL COMMENTS**

This is a very complex issue to study. The data they have is unique and although their analysis could still be subject to some other interpretations, the authors have done a very good and thorough job. I believe that they have adequately answered all my relevant points and questions. Given this, I believe that the methods, results, and general interpretation presented in this paper is sound and therefore that it should be recommended for publication in the BMJ Medicine Journal.

---

**VERSION 4 – AUTHOR RESPONSE**

Reviewer: 2

Comments to the Author
This is a very complex issue to study. The data they have is unique and although their analysis could still be subject to some other interpretations, the authors have done a very good and thorough job. I believe that they have adequately answered all my relevant points and questions. Given this, I believe that the methods, results, and general interpretation presented in this paper is sound and therefore that it should be recommended for publication in the BMJ Medicine Journal.

Author response: Thank you for the positive feedback and useful comments on the manuscript.