nature portfolio

Peer Review File

Uncovering the Heritable Components of Multimorbidities and Disease Trajectories Using a Nationwide Cohort

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The pedigree-breaking procedure, described in the following sentence: "As the largest extended pedigree included >5 million individuals, we partitioned the 177 network into families of approximately 500 individuals" is a bit unclear. In principle, it should be possible to compute h2 without even for very large pedigrees. A clarification would help.

As the data spans several decades, it would be very valuable to identify cohort-specific effect, that is were patterns observed 30 years back were still present in similar age-individuals 10 years ago?

It would be very useful for a reader to get an intuitive explanation of differences in h2 across sexes.

The trajectory heritability is an interesting concept, but would require some justification and explanation of observed results.

Large pedigrees allow for quite high-resolution analysis of environmental factors. Potentially, event time-share environment (that is 10, 20, or 30 years ago) can be incorporated into modeling. It would be important to touch this point + authors rationale for modeling environmental factors the way they did.

Dominance heritability requires definition in the text + intuitive explanation of results.

Generally, I love the study and would support it publication after some clarifications are made.

Reviewer #2 (Remarks to the Author):

Westergaard and colleagues estimate heritability and genetic correlations of disease traits using family relationships and ICD-10 codes in the nationwide Danish registries. They have

developed new methodology to estimate multiple variance components including additive and dominance genetic components, but also shared environment between family members and spouses. It is similar to Haseman-Elston regression with multiple variance components suitable for binary traits. The rich dataset allows them to explore the relative contribution of shared genetics and shared environment, as well as disease trajectories over time (series of ICD10 codes).

Overall, I think the authors have performed extensive analyses in a very valuable and rich dataset. They provide interesting insights in these data and share them through a user-friendly web-interface.

In general, the figures and wording could be more concise, especially in the supplementary material. Also, because the authors developed a new method to estimate variance components in these data, more details on the simulations and benchmarking would be welcome (see Methods below). This would also help to reassure some interesting, but sometimes counterintuitive findings (higher h2 in broad ICD10 categories compared to narrow) .

Methods:
The authors have done an impressive job on reconstructing pedigree structures using the CPR registry. Is there a way to visualize the number and size of the pedigrees in graph or figure? This would make it easier to appreciate than the numbers in the text.

Supplementary figure 2 seems to be missing, in other supplemenatry figures labels on axes are unclear (simulation results) and captions are minimal. I think it is good practice that figures + captions can be read on themselves, this could be improved for all supplementary figures.

The authors state that the linear transformation of h2 from the observed scale to the liability scale can be biased when estimates are derived from close relatives. Indeed, the Lee et al paper describes that this can be the case because the observed scale h2 estimates from close relatives contain both the additive as nonadditive variance. The genetic correlation,

however, is a scale-free parameter since the genetic variances (h2 for both traits) and their genetic covariance are measured on the same scale. Can the authors explain why Rg estimates would still be biased? They refer to supplementary figure 2 in the text, but this displays the simulation results. I think supplementary figure 2 is missing.

In Supplementary Figure 1, x and y axes describe liability, but shouldn't this be "h2 on liability scale"? What do the percentages in the panels reflect, disease prevalence? If so, could misspecification of the prevalence (life-time risk) or ascertainment (such that % cases in sample > population prevalence) in the simulation also increase bias? What method was used to estimate heritability here? It would help if the authors described the simulation scheme for these simulations in more detail.

The section on "Estimation of the expected product of phenotypes" lacks any explanation or annotation of terms in the equations. It would help if the authors at least explained the assumptions (first 3 lines) in wording, and introduce S1, S2, f1, f2.

In the supplementary results, the authors compare estimates with published results. The comparison with Athanasiadis et al seems to be missing from supplementary table 1.

Results:
There is a high spousal component for OCD, but ref 22 (Nordsletten et al) actually describes much stronger assortative mating for ADHD, ASD and schizophrenia and OCD seems one of the weakest (fig 1). Can the authors explain this, otherwise it is surprising that they highlight this result as it doesn't seem to be in line with ref 22. Were the spousal components in the psychiatric traits higher as a group in general?

The terminology of "progressive nature of Huntington's disease" is not totally accurate. The authors mean "anticipation" where a disease gets worse in the next generation because the repeat expansion is unstable and expands even further.
In that case you would assume prevalence is higher in offspring when proband is assigned to parent. However, in fig S8 prevalence in offspring is lower when proband is assigned to parent. If I read this correctly, could this pattern also be explained by the late-onset of

disease (offspring is yet undiagnosed)? Or pre-implantation diagnostics (mutation not transmitted to offspring)?

Some ICD10 terms that have higher h2 for the broad terms than the narrow terms (fig S9), can this be caused by the fact that other narrow terms "under" the broad terms much more heritable? For example, the category "Systemic atrophies primarily affecting the central nervous system" contains many rare monogenic (heritable) diseases e.g. early-onset ataxia, hereditary spastic paraplegia, of these the late-onset cerebellar ataxia's might be the least heritable of them all explaining the results?

What do the authors mean by "order score" for the genetic correlation within chapters?

It is nice to see that DCM and HCM are not genetically correlated, but share correlated traits ("palpitations"). Is this suggestive of vertical pleiotropy, i.e. both DCM and HCM cause palpitations? Can this phenomenon even be used to detect vertical pleiotropy?

Fig 12B show the disease trajectory of Non-insulin-dependent DM to insulin-dependent DM to retinal disorders. The authors describe this as a trajectory from type 2 to type 1 diabetes. Insulin-dependent DM is not synonymous to type 1 DM however, as many patients with type 2 diabetes become insulin dependent at some point in their disease course (this does not mean that they have DM type 1, they still have DM type 2).

General

It would help if the authors defined what they mean by disease trajectories early in the paper as this is unclear especially in the introduction.

Maybe the authors can provide a concrete example of how temporally ordered disease trajectories increase our understanding of disease progression in the fourth paragraph of the introduction (now readers have to look them up the references). The "prostate example" from Jensen et al (ref 16) makes it less abstract.

In the section on "h2 and rg across the full spectrum of diseases" the authors suggest that a family history is important given that h2 > 0.1 for many traits. For which goal is this

important? The value of a positive family history on disease risk greatly relies on the disease prevalence which together with h2 determines the risk for 1st degree relatives. For example, at h2=0.1 and prevalence=0.05 the risk of disease for a first degree relative is only slightly increased to 0.06 (see: https://shiny.cnsgenomics.com/CHARRGe/).

Figure 3: a legend for the ICD10 chapters would help.

Reviewer #3 (Remarks to the Author):

The paper discusses a nationwide cohort study conducted in Denmark to investigate the heritability and genetic correlations of various disease phenotypes in a large population over a 40-year period and 6 M people. The study aims to better understand the genetic and environmental factors contributing to disease initiation, progression, and the shared genetics between different diseases. It highlights the importance of studying shared aetiology in comorbid diseases and disease trajectories over long-term, comprehensive data for accurate estimates. On the very positive side, full dataset and results will be accessible.

Key findings include:

The study estimated the additive genetic heritability (h^2) for 866 different ICD-10 codes. It found that congenital malformations, endocrine diseases, and mental disorders had the highest heritability. Specific diseases with known genetic components, such as cystic kidney disease, also showed high heritability.

The study highlighted the impact of shared environmental factors on heritability estimates. Diseases related to infection, perinatal conditions, and the endocrine system were particularly affected by environmental factors.

Heritability estimates were shown to increase with diagnosis specificity. Diseases with more specific diagnostic codes tended to have higher heritability, with some interesting exceptions. Heritability differed between men and women for some diseases. Generally, heritability was higher in women for the investigated diseases. These differences are

discussed in terms of possible to sex-linked biological factors, differences in disease prevalence, or environmental factors.

The study found that diseases did not necessarily cluster by their assigned ICD-10 chapters. Some clusters included mental disorders, infections, and other conditions. In some cases, bidirectional relationships between diagnoses were identified, corresponding potentially to a shared common genetic background.

The study explored various disease trajectories and showed a heritability for some of them. For instance, trajectories that switched between Type I and Type II diabetes diagnoses.

While the study is very interesting and add to a long series of influential papers on comorbidities based on the Danish registry, there are also some difficulties with the background factors behind the observations that would have to be clarified. It is also apparent that the strength of the conclusions decreases with the progression of the paper that goes from more supported to more anecdotal cases. Probably, there is not much that can be done since it depends of the amount of data but it will, at least, good to reflect the variation in the degrees of confidence in the text.

Major comments:

- Age is possibly the main missing factor in the analysis- known to be the main factor contributing to diseases and disease comorbidities. How is age affecting heritability values? Is there a correlation between mean age of diagnoses and heritability?
- Could the higher heritability values in women be due to the fact that women go to the doctor more often (to be confirmed in the Danish population)? Are these measures corrected by the frequency of medical visits per individual? This is a fundamental point, since greater attendance at health services increases the probability of diagnosing diseases.
- Are diseases with a higher prevalence in society more likely to have high levels of heritability?
- Is the heritability of bidirectional associations affected by assortative mating? (Two diseases co-occur in children (without significant directionality) because their parents tend

to each have one disease).

Minor comments:

- Figure 3 shows the ICD10 codes by chapter number, but the results refer to the chapter name. Figure 3 should have the name of the chapters to facilitate the understanding of the figure.
- The list of bidirectional relationships with their rg values should be provided in supplementary material and referred in the text. Do bidirectional pairs with significant genetic correlations belong to the same ICD10 chapters? Please evaluate the significance of this analyses and highlight the chapters with significant enrichments.

Overall, it could have been interesting/important to consider social factors involved in the diagnosis of diseases, for example, families with a disease case are likely to sick for medical scrutiny of that disease in their family members. Cascade testing is known to have a significant impact in medical diagnosis, particularly in rare diseases, potentially increasing the number of detected diseases and disease heritability cases.

We thank the reviewers for their comments which have substantially improved the paper. We note that all three reviewers are positive about the paper overall, with reviewer #1 saying "Generally, I love the study and would support it for publication after some clarifications are made." and reviewer #2 saying "The authors have done an impressive job on reconstructing pedigree structures using the CPR registry." Reviewer #3 notes "The paper discusses a nationwide cohort study conducted in Denmark to investigate the heritability … It highlights the importance of studying shared aetiology in comorbid diseases and disease trajectories over long-term, comprehensive data for accurate estimates. On the very positive side, full dataset and results will be accessible."

Below we list the reviewer's point one by one, indicating how we changed the text or in the rare cases explaining our rationale we took in the paper without changes to the text. We have enclosed a version with and with tracked changes, for your convenience.

Reviewer #1 (Remarks to the Author):

The pedigree-breaking procedure, described in the following sentence: "As the largest extended pedigree included >5 million individuals, we partitioned the 177 network into families of approximately 500 individuals" is a bit unclear. In principle, it should be possible to compute h2 without even for very large pedigrees. A clarification would help.

Answer: The reviewer is correct, it is possible to calculate the heritability for millions of individuals, especially when working with a kinship sparse matrix (compared to e.g. a genomic relationship matrix). In quantitative genetics, this is often done for millions of e.g. sheeps or cows using restricted maximum likelihood. However, the conversion between the observed h2 and liability h2 that is often proposed breaks down in the presence of closely related individuals. We show this in Supplementary Figure 1. One alternative would be a Bayesian probit model. We did test this approach, but found it would take at least 14 days for each phenotype even with substantial computational resources (>100 CPU, >1 TB of RAM). We did not investigate how long it would take to compute genetic correlations, but most likely it would be 30+ days due to the increased computational complexity. This is why we developed a novel method to estimate h2 on the liability scale. However, with our proposed method, it is not possible to estimate the standard error and thereby quantify the uncertainty. This is why we divided the pedigree into ten independent pedigrees, in order to estimate the uncertainty.

We expand this more in text by saying "We used the 10 different independent pedigrees to estimate the variance of the liability based h2 model", in the Statistical Analysis section.

As the data spans several decades, it would be very valuable to identify cohort-specific effect, that is were patterns observed 30 years back were still present in similar age-individuals 10 years ago?

Answer: This is a valid point. It is a given that the incidence and the diagnostics of diseases have changed over time. The principal issues with addressing this in data are two:

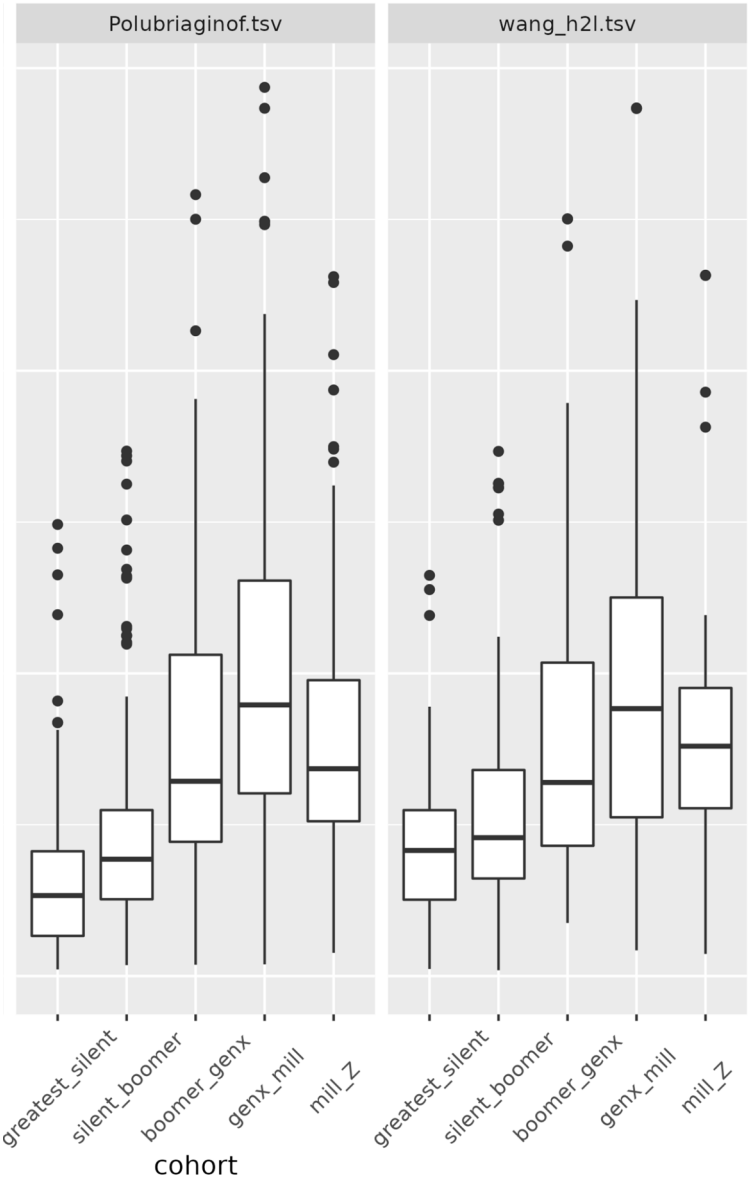
(1) We have a population followed in a fixed time window, which limits the variability of age, period and cohort: If we split the data by birth cohort, then the cohorts will have substantially different age-at-risk distributions. This complicates interpretation and/or requires further modeling or subsetting of the data. We have calculated the birth cohort specific heritability, and find XXX…

(2) Any difference between cohorts might be equally well attributed to period as age due to the collinearity between age, period and cohort. This means that a difference in the age-specific risk between two cohorts might be attributable to cohort, period or a mix of the two. This complicates interpretation.

(3) Change in screening patterns over the time period may also affect heritability, and this is difficult to account for.

(4) Family structure limits the year specific heritability

We did attempt to investigate birth cohort specific heritability. Due to using pedigrees, we had to include two birth year cohorts into each comparison. I.e., there are a total of five comparisons to be made. We noticed that the heritability increased until we estimated it in the strata where the majority of data lies (generation X - Millenium cohort). Afterwards, it decreases, most likely due to power issues.,



We have added in the discussion “Future studies might explore new ways to integrate information from extended families into environmental covariance matrices and contrasting age effects, which will need more coordination on the data types present in the extraction”. under the section “Limitations and Strengths”.

It would be very useful for a reader to get an intuitive explanation of differences in h2 across sexes.

Answer: We agree and have revised the section describing the differences in heritability for men and women. The section “Sex-specific estimates of heritability and genetic correlations” now reads “Deviations in heritability could be attributed to genetic architecture, gene-environment interactions, epigenetic modifications, and diagnosis misclassification. Men and women exhibit distinct genetic makeups, especially in relation to sex chromosomes, which can lead to variations in how diseases are genetically expressed. Furthermore, the interaction

between genetic predispositions and environmental factors can manifest differently across sexes. Epigenetic modifications may also vary between sexes. Lastly, diagnosis misclassification can arise from sex-specific genetic and epigenetic variations that lead to different disease presentations, which may not be accurately recognized by standard diagnostic criteria, resulting in potential misdiagnosis or delayed diagnosis in clinical settings." We have previously considered the sex differences in ref. 6, Nat. Comm. 10:666, 2019.

The trajectory heritability is an interesting concept, but would require some justification and explanation of observed results.

Answer: The concept of "trajectory heritability" is indeed novel and necessitates further elucidation. Trajectory heritability refers to the heritable component of the progression patterns, or trajectories, of a disease over time, rather than just the heritability of one disease's occurrence. This is also motivated by the fact that many diseases can appear as both risk factors and complications to other diseases. They occur therefore in different etiological contexts: Heritable trajectories might arise as a consequence of vertical pleiotropy where an initial event causes further events downstream, e.g. a particularly severe initial disease which manifests itself in faster progression and a higher likelihood of complications later on. Trajectories might also be caused by horizontal pleiotropy where genetic factors influence the risk of several disease phenotypes, which occur at different stages of life, e.g. over- or under regulation of some biological system, which increases the risk of several diseases.
We have added the following to the Results section on Trajectories: "Heritable trajectories might arise as a consequence of vertical pleiotropy where an initial event causes further events downstream, e.g. a particularly severe initial disease which manifests itself in faster progression and a higher likelihood of complications later on. Trajectories might also be caused by horizontal pleiotropy where genetic factors influence the risk of several disease phenotypes, which occur at different stages of life, e.g. over- or under regulation of some biological system, which increases the risk of several diseases."

Large pedigrees allow for quite high-resolution analysis of environmental factors. Potentially, event time-share environment (that is 10, 20, or 30 years ago) can be incorporated into modeling. It would be important to touch this point + authors rationale for modeling environmental factors the way they did.

Answer: Unfortunately, we do not have the data that allows us to precisely map out at which point in time two individuals shared an environment at population scale. Modeling the environment is an inherently difficult task, which is severely underdeveloped in the context of heritability. The variance components that we have included are the practical scenarios which we can track given the data. The terms covers the following:
(full) Siblings: non-additive dominance genetic effects, shared in utero environment from having the same birth mother, early life cohabitation
Couple: Assortative mating, cohabitation.
Nuclear family: Trans-generational cohabitation.

We have added the following to “Limitations and Strengths” in the discussion “our choice of environmental terms is driven largely by our ability to model different sorts of joint environments between people, namely full sibling, couple and nuclear family”

Dominance heritability requires definition in the text + intuitive explanation of results.

Answer: We have now clarified this concept in our manuscript, under the section “The confounding of shared environment and non-additive genetic effects”. Dominance heritability pertains to the variance in a trait attributable to interactions between alleles at a single locus, highlighting the importance of non-additive genetic effects. Our methodological approach specifically involves constructing a dominance covariance matrix derived from the additive kinship matrix, which is built using pedigree data. This technique allows us to distinguish between additive and dominance genetic effects, enhancing our understanding of the genetic architecture of diseases. The inclusion of this methodology in our study underscores the importance of considering complex genetic interactions in the development of personalized treatment strategies and thereby contribute to a more nuanced understanding of genetic predisposition in diseases.
We have added the following to the Results section describing Dominance patterns:
“Dominance heritability pertains to the variance in a trait attributable to interactions between alleles at a single locus, highlighting the importance of non-additive genetic effects”

Generally, I love the study and would support it publication after some clarifications are made.

Answer: Thank you! We hope it will be a great resource for many years to come.

Reviewer #2 (Remarks to the Author):

Westergaard and colleagues estimate heritability and genetic correlations of disease traits using family relationships and ICD-10 codes in the nationwide Danish registries. They have developed new methodology to estimate multiple variance components including additive and dominance genetic components, but also shared environment between family members and spouses. It is similar to Haseman-Elston regression with multiple variance components suitable for binary traits. The rich dataset allows them to explore the relative contribution of shared genetics and shared environment, as well as disease trajectories over time (series of ICD10 codes).

Overall, I think the authors have performed extensive analyses in a very valuable and rich dataset. They provide interesting insights in these data and share them through a user-friendly web-interface.

In general, the figures and wording could be more concise, especially in the supplementary material. Also, because the authors developed a new method to estimate variance components in these data, more details on the simulations and benchmarking would be welcome (see Methods below). This would also help to reassure some interesting, but sometimes counterintuitive findings (higher h2 in broad ICD10 categories compared to narrow) .

Answer: We have added further clarification to figure captions and supplementary methods. We fully agree that the higher h2 in broader ICD-10 categories, compared to more specific ICD-10 terms, is not quite intuitive. We are confident in the new method we have developed, and we have revised the Supplementary Methods with more details on the Method and Simulations.

Methods:
The authors have done an impressive job on reconstructing pedigree structures using the CPR registry. Is there a way to visualize the number and size of the pedigrees in graph or figure? This would make it easier to appreciate than the numbers in the text.

Answer: We have struggled with this visualisation. The pedigrees are hard to visualise and most look like either a mess of lines or, if laid out in a more circular way, a large hairball. This is mainly due to the scale of the data we are dealing with and the inherent interconnection of the largest pedigree in Denmark. Having looked at some options, we do not think there is an easy visualisation.

Supplementary figure 2 seems to be missing, in other supplemenatry figures labels on axes are unclear (simulation results) and captions are minimal. I think it is good practice that figures + captions can be read on themselves, this could be improved for all supplementary figures.

Answer: The figure was meant to illustrate that using a linear model, in this case Haseman-Elston (HE) regression, which estimates the observed heritability and covariance, is not suitable for binary traits. Linear models, including HE and linear mixed models, are commonly used to estimate the genetic correlation for binary traits in pedigrees, under the pretext that the conversion between observed and liability scale cancels out. As we show, it is actually biased at increasing values of genetic correlation. We have now emphasized this. We have improved labels and captions for most Supplementary Figures.

The authors state that the linear transformation of h2 from the observed scale to the liability scale can be biased when estimates are derived from close relatives. Indeed, the Lee et al paper describes that this can be the case because the observed scale h2 estimates from close relatives contain both the additive as nonadditive variance. The genetic correlation, however, is a scale-free parameter since the genetic variances (h2 for both traits) and their genetic covariance are measured on the same scale. Can the authors explain why Rg estimates would

still be biased? They refer to supplementary figure 2 in the text, but this displays the simulation results. I think supplementary figure 2 is missing.

Answer: Indeed, the genetic correlation is supposed to be a scale free parameter, as argued and shown (theoretically) in many studies. However, in our simulations, we see that estimates become biased at higher levels. As we also show, in the method that we propose, there is no such bias (Supplementary Figure 2).

In Supplementary Figure 1, x and y axes describe liability, but shouldn't this be "h2 on liability scale"? What do the percentages in the panels reflect, disease prevalence? If so, could misspecification of the prevalence (life-time risk) or ascertainment (such that % cases in sample > population prevalence) in the simulation also increase bias? What method was used to estimate heritability here? It would help if the authors described the simulation scheme for these simulations in more detail.

Answer: Many thanks for spotting this oversight, the reviewer is correct, it should be heritability on the liability scale. The percentages reflect the bias we observed at different disease prevalence. Supplementary Text and Figures have been revised to include a more detailed description of the Method and Simulations.

The section on "Estimation of the expected product of phenotypes" lacks any explanation or annotation of terms in the equations. It would help if the authors at least explained the assumptions (first 3 lines) in wording, and introduce S1, S2, f1, f2.

Answer: Many thanks for bringing this to our attention. The description of the method in Supplementary Methods have been significantly improved based on input from all reviewers.

In the supplementary results, the authors compare estimates with published results. The comparison with Athanasiadis et al seems to be missing from supplementary table 1.

Answer: We thank the reviewer for this input, and we have now added this comparison. When including only the genetic component there is a good correlation (rho=0.77). However, when adding the environmental covariance matrices, the correlation drops (rho=0.63).

Results:
There is a high spousal component for OCD, but ref 22 (Nordsletten et al) actually describes much stronger assortative mating for ADHD, ASD and schizophrenia and OCD seems one of the weakest (fig 1). Can the authors explain this, otherwise it is surprising that they highlight this result as it doesn't seem to be in line with ref 22. Were the spousal components in the psychiatric traits higher as a group in general?

Answer: While certainly weaker, there is evidence of assortative mating with respect to OCD. However, the reviewer is correct that in ref 22, the association is weaker compared to .e.g ADHD, ASD and schizophrenia. It is difficult to make a one to one comparison for these

disorders, as the phenotypic definitions using ICD10 vary. E.g. ICD10 is encompassed by F42, whereas Schizophrenia is typically defined across the range F20-F29 (see e.g. Athanasiadis, ref 11). Generally speaking, the spousal component within the psychiatric traits did not have a higher mean value, but some of the more extreme values were found there (see Supplementary Figure 7). We have added the following to "The confounding of shared environment and non-additive genetic effects" in the Results Section: "However, we did not find any evidence of assortative mating for ADHD, ASD, or schizophrenia. This may be due to variation in phenotypic definitions. In general, perinatal disorders, followed by congenital disorders and infectious diseases had the highest mean spousal component (Supplementary Figure 7)." The lack of evidence for assortative mating may be due to phenotypic inaccuracies. Which would bias the heritability estimates downwards.

The terminology of "progressive nature of Huntington's disease" is not totally accurate. The authors mean "anticipation" where a disease gets worse in the next generation because the repeat expansion is unstable and expands even further.
In that case you would assume prevalence is higher in offspring when proband is assigned to parent. However, in fig S8 prevalence in offspring is lower when proband is assigned to parent. If I read this correctly, could this pattern also be explained by the late-onset of disease (offspring is yet undiagnosed)? Or pre-implantation diagnostics (mutation not transmitted to offspring)?

Answer: The reviewer is correct. This is anticipation. In discussions with clinical colleagues we need to do more investigation around Huntington's heritability and prevalence patterns, which show interesting but non-obvious transmission. Given this we have removed this section from the paper, and added "More investigations are required for these non-obvious transmission patterns." to the Result section.

Some ICD10 terms that have higher h2 for the broad terms than the narrow terms (fig S9), can this be caused by the fact that other narrow terms "under" the broad terms much more heritable? For example, the category "Systemic atrophies primarily affecting the central nervous system" contains many rare monogenic (heritable) diseases e.g. early-onset ataxia, hereditary spastic paraplegia, of these the late-onset cerebellar ataxia's might be the least heritable of them all explaining the results?

Answer: We agree that if something on the more specific level has a higher heritability, it may deflate when using the less specific broader terms. We would expect that combining rare monogenic diseases should, in theory, give a lower estimate of heritability.

What do the authors mean by "order score" for the genetic correlation within chapters?

Answer: If the reviewer is referring to the results displayed in Figure 9 and Supplementary Figure 12, this was an attempt to dissect how often diagnoses from the same functional group (i.e. ICD-10 chapter) clustered together, based on the observed genetic correlation. Two diagnoses would cluster together if the genetic correlation to all other diagnoses are similar.

The outlier here was the mental disorder chapter, which cluster together with other mental disorders >80% of cases.

It is nice to see that DCM and HCM are not genetically correlated, but share correlated traits ("palpitations"). Is this suggestive of vertical pleiotropy, i.e. both DCM and HCM cause palpitations? Can this phenomenon even be used to detect vertical pleiotropy?

Answer: The reviewer raises a good question. We are not so sure there's talk about vertical pleiotropy, but perhaps multiple independent pathways can lead to palpitations.

Fig 12B show the disease trajectory of Non-insulin-dependent DM to insulin-dependent DM to retinal disorders. The authors describe this as a trajectory from type 2 to type 1 diabetes. Insulin-dependent DM is not synonymous to type 1 DM however, as many patients with type 2 diabetes become insulin dependent at some point in their disease course (this does not mean that they have DM type 1, they still have DM type 2).

Answer: We agree with the reviewer that there is not a one to one correspondence, and have altered this. The Result section now specifically talks about the progression from non-insulin to insulin-dependent diabetes.

General
It would help if the authors defined what they mean by disease trajectories early in the paper as this is unclear especially in the introduction.

Maybe the authors can provide a concrete example of how temporally ordered disease trajectories increase our understanding of disease progression in the fourth paragraph of the introduction (now readers have to look them up the references). The "prostate example" from Jensen et al (ref 16) makes it less abstract.

Answer: We have added the prostate example from Jensen et al. to the introduction, which now reads: "For example, hyperplasia of the prostate may develop into prostate cancer, which can then progress to metastatic disease".

In the section on "h2 and rg across the full spectrum of diseases" the authors suggest that a family history is important given that h2 > 0.1 for many traits. For which goal is this important? The value of a positive family history on disease risk greatly relies on the disease prevalence which together with h2 determines the risk for 1st degree relatives. For example, at h2=0.1 and prevalence=0.05 the risk of disease for a first degree relative is only slightly increased to 0.06 (see: https://shiny.cnsgenomics.com/CHARRGe/).

Answer: The reviewer is correct that the life-time risk of disease is increased to 0.06 in first degree siblings. However, this is for lifetime risk, which is different from prevalence in this

study. If the lifetime risk and prevalence were to be equivalent, we would have had to follow the full population from cradle to grave. The importance for accurate family history may be overstated for diagnosis, but the point was to emphasize that more than half of diagnoses have varying degrees of additive heritability.

Figure 3: a legend for the ICD10 chapters would help.

Answer: We have added a legend to Figure 3.

Reviewer #3 (Remarks to the Author):

The paper discusses a nationwide cohort study conducted in Denmark to investigate the heritability and genetic correlations of various disease phenotypes in a large population over a 40-year period and 6 M people. The study aims to better understand the genetic and environmental factors contributing to disease initiation, progression, and the shared genetics between different diseases. It highlights the importance of studying shared aetiology in comorbid diseases and disease trajectories over long-term, comprehensive data for accurate estimates. On the very positive side, full dataset and results will be accessible.

Key findings include:

The study estimated the additive genetic heritability (h^2) for 866 different ICD-10 codes. It found that congenital malformations, endocrine diseases, and mental disorders had the highest heritability. Specific diseases with known genetic components, such as cystic kidney disease, also showed high heritability.

The study highlighted the impact of shared environmental factors on heritability estimates. Diseases related to infection, perinatal conditions, and the endocrine system were particularly affected by environmental factors.

Heritability estimates were shown to increase with diagnosis specificity. Diseases with more specific diagnostic codes tended to have higher heritability, with some interesting exceptions. Heritability differed between men and women for some diseases. Generally, heritability was higher in women for the investigated diseases. These differences are discussed in terms of possible to sex-linked biological factors, differences in disease prevalence, or environmental factors.

The study found that diseases did not necessarily cluster by their assigned ICD-10 chapters. Some clusters included mental disorders, infections, and other conditions. In some cases, bidirectional relationships between diagnoses were identified, corresponding potentially to a shared common genetic background.

The study explored various disease trajectories and showed a heritability for some of them. For instance, trajectories that switched between Type I and Type II diabetes diagnoses.

While the study is very interesting and add to a long series of influential papers on comorbidities based on the Danish registry, there are also some difficulties with the background factors behind the observations that would have to be clarified. It is also apparent that the strength of the conclusions decreases with the progression of the paper that goes from more supported to more anecdotal cases. Probably, there is not much that can be done since it depends of the amount of data but it will, at least, good to reflect the variation in the degrees of confidence in the text.

Major comments:
- Age is possibly the main missing factor in the analysis- known to be the main factor contributing to diseases and disease comorbidities. How is age affecting heritability values? Is there a correlation between mean age of diagnoses and heritability?

Answer: All analyses have been adjusted for age. We have added a figure showing the relationship between mean age of diagnosis and heritability, and report the correlation, … The heritability for age-of-onset is different from what we study, and would require the development of models that incorporate time at risk and censoring. There is indeed a negative correlation between mean age of onset and heritability (rho= -0.24; p=2.86e-13). We have added this to the Results section: "We observed a negative correlation between the heritability and mean age of onset (ρ=-0.24, p=2.97e-13), as well as a small correlation to the prevalence (ρ=-0.07, p=0.054)".

- Could the higher heritability values in women be due to the fact that women go to the doctor more often (to be confirmed in the Danish population)? Are these measures corrected by the frequency of medical visits per individual? This is a fundamental point, since greater attendance at health services increases the probability of diagnosing diseases.

Answer: It is true that the number of contacts with the healthcare system is higher among females. This is particularly true for the primary care sector (see e.g. https://statbank.dk/sygk2). And while the primary care sector does not provide the ICD-codes used in this study, a majority of referrals to hospitals are done by the primary sector. Some of the difference is due to pregnancy and childbirth, some is likely due to the comparatively higher longevity of women, which leaves more time at ages where morbidity is high. However, it is unclear how this, in and of itself, might work to bias heritability. For instance, diseases like heart disease are arguably less well diagnosed in women due to a bias towards basing diagnosis on how symptoms present themselves in men. Poor diagnosis would introduce random noise into the model, which would bias estimates towards the null. It is possible that women-only within-family social 'contagion' might inflate heritability: If the propensity to seek diagnosis for certain symptoms depends on having a family member with the same symptoms, then this would increase heritability estimates. If this mechanism was only present for females, then it might explain the sex difference. To our knowledge this

phenomenon is not well examined, but results from behavioral economics suggest that there might be behavioral spillover from diagnosis between family members (https://www.aeaweb.org/articles?id=10.1257/aer.20171993)

- Are diseases with a higher prevalence in society more likely to have high levels of heritability?

Answer: See answer above.

- Is the heritability of bidirectional associations affected by assortative mating? (Two diseases co-occur in children (without significant directionality) because their parents tend to each have one disease).

Answer: That is an interesting idea. We have estimated the genetic correlations between bidirectional associations, in order to demonstrate that the concept can be used to identify new phenotypes and differentiate between diagnosis bias and the underlying biology. When calculating the genetic correlations we did include the spousal component if the variance explained in the two individual heritability estimates were greater than 5%. However, we are uncertain how assortative mating might affect the bidirectional associations.

Minor comments:
- Figure 3 shows the ICD10 codes by chapter number, but the results refer to the chapter name. Figure 3 should have the name of the chapters to facilitate the understanding of the figure.

Answer: We have adjusted Figure 3 to include a legend.

- The list of bidirectional relationships with their rg values should be provided in supplementary material and referred in the text. Do bidirectional pairs with significant genetic correlations belong to the same ICD10 chapters? Please evaluate the significance of this analyses and highlight the chapters with significant enrichments.

Answer: We have added all estimates as Supplementary Material available through the website. There is, per say, no significant genetic correlations. All estimates are derived from a Bayesian model. 89/160 (55%) of the bidirectional pairs that we investigate stem from the same chapter. Inherently there was no difference in the average genetic correlation between those pairs from the same chapter, versus those from different chapters (Supplementary Figure 13).

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

I am happy with the revision.

Reviewer #2 (Remarks to the Author):

I thank the authors for their comprehensive response. In my opinion they have substantially clarified the manuscript. I have no further comments.

Reviewer #4 (Remarks to the Author):

Dear authors,

I want to congratulate you on the great work you have done. This is a very relevant work. However, two points have not been adequately clarified/solved. On the one hand, the examples provided of possible factors that may affect the diagnosis of different diseases in women are interesting, and it is something that could be mentioned in the discussion. However, the authors still need to answer the question numerically. I would like to know if, in their data, they have information on the number of times each patient goes to the hospital. If so, one could evaluate whether this correlates with a higher number of diagnoses, thus resulting in higher heritability values. In turn, the following question, focused on the impact of disease prevalence on heritability values, has also yet to be answered statistically. The reformulated question would be: is there a correlation between disease prevalence and heritability values? This is something that can be measured by correlation.

We thank the reviewer for the clarification of questions. We have responded to the questions, point by point.

Reviewer #4 (Remarks to the Author):

Dear authors,
I want to congratulate you on the great work you have done. This is a very relevant work.

However, two points have not been adequately clarified/solved. On the one hand, the examples provided of possible factors that may affect the diagnosis of different diseases in women are interesting, and it is something that could be mentioned in the discussion.

However, the authors still need to answer the question numerically. I would like to know if, in their data, they have information on the number of times each patient goes to the hospital. If so, one could evaluate whether this correlates with a higher number of diagnoses, thus resulting in higher heritability values.

Answer: As we understand it, the reviewer is curious about how the individual healthcare seeking behavior affects heritability at the population level. We also discussed this in the previous rebuttal, on average women do have more contacts with the healthcare sector for example. However, making a robust calculation as suggested is highly complex, and it may not be helpful in relation to assessing the impact of visit frequencies, primarily because there are many visits which are irrelevant for the heritability aspect.

In the Danish setting, a diagnosis must be assigned to each patient contact. This is how the Danish healthcare model works: Hospitals are reimbursed at specific rates for the specific diagnosis. There are multiple types of patient hospital contacts, including in-patients, out-patients, and emergency room visits. Out-patients are typically of longer duration (even years), in which a contact between the person and hospital is kept open for monitoring, scheduled visits, etc. These do not result in a diagnosis every time the patient visits. Furthermore, not all hospital contacts are due to illness or disease. Danish hospitals also handle first trimester pregnancy risk assessment, second trimester risk assessment, and births. 90%+ of all pregnant women accept the offer for first- and second-trimester risk assessment. 95%+ of all births in Denmark take place in a hospital. In the majority of cases, they will only see nurses or midwives. Secondly, after being referred to a hospital, a patient may be referred for extra examinations (e.g. MRI, CT, etc) at another hospital or department. While technically the same visit, this would in the administrative registries be two different events. Consequently, defining and performing statistical analysis on healthcare seeking behavior is extremely complex and requires much more detailed considerations, beyond the scope of this work.

In turn, the following question, focused on the impact of disease prevalence on heritability values, has also yet to be answered statistically. The reformulated question would be: is there

a correlation between disease prevalence and heritability values? This is something that can be measured by correlation.

Answer: In the latest revision, we added the sentence (relevant parts in bold) to the Results,

"We observed a negative correlation between the heritability and mean age of onset (ρ=-0.24, p=2.97e-13), **as well as a small correlation to the prevalence (ρ=-0.07, p=0.054)**"

Based on these findings, the correlation between prevalence and heritability is negligibly small.