

# **Transferability of European-derived Alzheimer's Disease Polygenic Risk Scores across Multi-Ancestry Populations**

Corresponding Author: Dr Jean-Charles Lambert

Version 0:

Decision Letter:

12th Jul 2023

Dear Dr. Lambert,

Thank you very much for your enquiry about submitting a manuscript "Portability of an Alzheimer's Disease Genetic Risk Score across multi-ancestry populations" to Nature Genetics. It sounds interesting and we would like to see it, although I am sure you will appreciate that it can be very difficult to judge a paper without having seen it in its entirety. We cannot guarantee that we'll send it out for review until we have had a chance to look at it.

Please note that supplementary materials must not be included with the main text and figures of your manuscript. Supplementary items should be uploaded as separate files.

At this stage you will also need to upload:

1) The Editorial Policy Checklist:

<https://www.nature.com/documents/nr-editorial-policy-checklist.pdf>

2) The Reporting Summary:

<https://www.nature.com/documents/nr-reporting-summary.pdf>

(Here you can read about the role of the Reporting Summary in reproducible science:

<https://www.nature.com/news/announcement-towards-greater-reproducibility-for-life-sciences-research-in-nature-1.22062> )

In order to submit your complete manuscript, please use the link below and follow the prompt to "Revise NG-PI62968":

Link Redacted

We look forward to receiving it.

Sincerely,

Wei

Wei Li, PhD

Senior Editor

Nature Genetics

New York, NY 10004, USA

[www.nature.com/ng](http://www.nature.com/ng)

Version 1:

Decision Letter:

21st Nov 2023

Dear Dr Lambert,

Your Letter, "Transferability of a European-derived Alzheimer's Disease Genetic Risk Score across Multi-Ancestry Populations." has now been seen by 2 referees. You will see from their comments copied below that while they find your

work of considerable potential interest, they have raised quite substantial concerns that must be addressed. In light of these comments, we cannot accept the manuscript for publication, but would be very interested in considering a substantially revised version that addresses these serious concerns.

We hope you will find the referees' comments useful as you decide how to proceed. If you wish to submit a substantially revised manuscript, please bear in mind that we will be reluctant to approach the referees again in the absence of major revisions.

To guide the scope of the revisions, the editors discuss the referee reports in detail within the team with a view to identifying key priorities that should be addressed in revision. In this case, we think both referees have provided constructive reviews aimed at strengthening the analyses and improving the presentation. Both referees have identified important limitations of the study (for example, lack of comparison of the PGS with PRS derived from a modern approach, no attempt to improve performance in global/underrepresented populations, etc). We ask that you perform additional analyses and address their technical comments as thoroughly as possible with appropriate revisions. We hope that you will find the prioritized set of referee points to be useful when revising your study.

If you choose to revise your manuscript taking into account all reviewer and editor comments, please highlight all changes in the manuscript text file. At this stage we will need you to upload a copy of the manuscript in MS Word .docx or similar editable format.

We are committed to providing a fair and constructive peer-review process. Do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

If revising your manuscript:

\*1) Include a "Response to referees" document detailing, point-by-point, how you addressed each referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the referees along with the revised manuscript.

\*2) If you have not done so already please begin to revise your manuscript so that it conforms to our Letter format instructions, available <a href="http://www.nature.com/ng/authors/article\_types/index.html">here</a>. Refer also to any guidelines provided in this letter.

\*3) Include a revised version of any required Reporting Summary: <https://www.nature.com/documents/nr-reporting-summary.pdf>

It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review.

A revised checklist is essential for re-review of the paper.

Please be aware of our <a href="https://www.nature.com/nature-research/editorial-policies/image-integrity">guidelines on digital image standards.</a>

You may use the link below to submit your revised manuscript and related files:

Link Redacted

**Note:** This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

If you wish to submit a suitably revised manuscript we would hope to receive it within 6 months. If you cannot send it within this time, please let us know. We will be happy to consider your revision so long as nothing similar has been accepted for publication at Nature Genetics or published elsewhere. Should your manuscript be substantially delayed without notifying us in advance and your article is eventually published, the received date would be that of the revised, not the original, version.

Please do not hesitate to contact me if you have any questions or would like to discuss the required revisions further.

Nature Genetics is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit please visit <a href="http://www.springernature.com/orcid">www.springernature.com/orcid</a>.

Thank you for the opportunity to review your work.

Sincerely,  
Wei

Wei Li, PhD  
Senior Editor  
Nature Genetics  
New York, NY 10004, USA  
www.nature.com/ng

#### Reviewers' Comments:

##### Reviewer #1:

Remarks to the Author:

NG-LE62968R

Transferability of a European-derived Alzheimer's disease genetic risk score across multi-ancestry populations

Nicolas et al.

In this manuscript, Nicolas and colleagues derived a genetic risk score (GRS) for Alzheimer's disease (AD) using genome-wide significant loci from the latest AD GWAS in European ancestry individuals, and assessed the associations of the PGS with AD diagnosis, age of onset and AD biomarkers across global populations. There are several strengths of this study, including a large testing sample, a comprehensive evaluation of the AD PGS across many countries, both within and outside the European continent, and the evaluation of the PGS associations with age at onset, biomarker levels and diagnosis criteria. However, the study also has a number of limitations as detailed below.

- The authors used a PGS derived from 83 independent SNPs which was first constructed and evaluated in a 2022 Nature Genetics publication. This raises two issues: (i) the novelty and overlap of this study with the 2022 publication, although new datasets have been added in the present study; (ii) it is known that for the vast majority of complex traits and diseases, PRS constructed by more sophisticated Bayesian methods substantially outperform PRS derived from genome-wide significant loci. I would love to see a comparison of the PGS with PRS derived from a modern approach such as PRS-CS and LDpred2.

- The results of the manuscript didn't provide much new insights into the transferability of PGS or AD genetics. A large number of studies have shown that for a wide range of complex traits and diseases, PGS derived from European GWAS have reduced predictive performance as the genetic distance between the training and target samples increase (btw- the authors failed to cite almost all of them). So the results in this manuscript are totally expected. There are some interesting results - for example, the authors showed that in admixed populations the PGS prediction accuracy decreased as the genomic proportion of African and Native American ancestries increased. These are in line with the recent report (Ding et al. Nature 2023). However, the authors didn't dig deeper into this interesting observation. There is also no description in the methods section on how ancestry percentage was quantified. Did the authors perform any local ancestry inference?

- I would suggest a more systematic investigation of the relationship between the PGS and APOE alleles. What are the frequencies of APOE alleles in different populations? What was the relative contribution of the PGS and APOE to AD risk? What about the joint prediction of PGS + APOE? Reporting other PGS performance metrics in addition to OR including variance explained (on the liability scale) and AUC would also be helpful.

- In the mega-analysis when pooling data across countries, it is important to assess if there is any shift of PGS distributions in different populations, and to calibrate PGS distributions before stratifying the PGS into quantiles (see e.g. the work from eMERGE consortium PMID: 37333246). The authors didn't assess the calibration of PGS in any way.

- There is a significant disconnection with the latest PRS literature. The authors almost exclusively cited AD PRS work, some of which are quite outdated, and largely ignored the latest developments in the larger PRS community including some of the most important papers in the field, from Martin et al. showing the portability issue of PRS across ancestries, to a large body of methodological work on improving the predictive performance of PRS within and across populations, and to the more recent developments and discussions on the transferability of PRS across global populations (see e.g. PMID: 37620596 and many references cited therein). The authors should better review and acknowledge existing work.

- The authors used PGS to refer to scores derived from genome-wide significant loci and used PRS to denote scores derived from genome-wide variants. This doesn't seem to be consistent with how the field uses these terms (again see PMID: 37620596).

- The manuscript was also not very well written with numerous typos. Some careful proofreading is needed.

##### Reviewer #2:

Remarks to the Author:

This manuscript provides the most comprehensive evaluation of a genetic risk score in Alzheimer's disease across global populations. The manuscript concludes that a genetic risk score relates to AD risk, with decreasing associations as the precision of the phenotype wanes and among populations that are not well-represented in the initial GWAS studies of AD.

The manuscript has numerous strengths and fills an important gap in current understanding of applications of genetic risk scores in AD, but misses an opportunity to characterize and address some of the major challenges in this sphere. In current form the manuscript seems to be incremental in its advance from previous work. At the very least, I think a clear statement about the current and future clinical utility of GRS should be included in this manuscript and a more clear roadmap for improving applications in global populations.

#### Major Concerns

(1) While the authors highlight differences in the performance of the GRS by population, they do very little to try and address this major limitation within the manuscript. There is no adjustment for linkage disequilibrium and no attempt to improve performance in populations that are not well-represented. Moreover, there is no attempt to evaluate or demonstrate that the leveraged score narrowly focused on the top set of sentinel variants within the top loci is indeed the best approach for such a score, and perhaps more importantly, whether such an approach by design hampers the ability to apply a GRS in other populations. There are many tools available to try and improve PRS performance across populations and such a goal seems within scope of this global assessment of a genetic risk score.

(2) It is not clear why variants that have population specific effects were not integrated into this manuscript. The authors make a strong argument that African admixture may modify the association for some of the variants selected for this GRS and AD, but don't attempt to integrate variants that are known to show a stronger association in populations with African ancestry.

(3) Sample overlap is a major consideration in a study like this, but is not raised. Did the authors attempt to estimate sample overlap in any of the populations in which the GRS is assessed or adjust estimates for potential biases due to sample overlap? Were the participants included in the biomarker analysis also in the original GWAS?

(4) Considerations around the sex distributions across populations should be discussed, particularly given the preponderance of women in many of the cohorts used in the initial analyses and the preponderance of men in many of the studies used to assess the GRS in this manuscript.

#### Minor Concerns

(5) The overall association for the GRS is incredibly low, with odds ratios  $<1.10$  in most populations. Can the authors speculate on why GRS continue to perform so poorly outside of the APOE region given the high estimated heritability from twin studies?

(6) Given the variable coverage across global populations, and the desire to compare estimates across populations, it would seem more appropriate to limit the comparison to a score calculated on the set of variants available in all populations. Given the small number of variants included in the overall score, the current approach is dropping more than 10% of the variants in some populations.

(7) The authors report a modest interaction between APOE and the GRS, and then state that there is no difference in stratified results and "APOE and GRS risk are independent). From the interaction result, and the stratified results, a more appropriate interpretation seems to be that the GRS has a slightly stronger association among E4 homozygotes but is equivalent in other strata.

(8) The authors raise a biological interpretation of the lack of heterogeneity in the GRS associations with tau compared to amyloid, but couldn't this just as easily be technical given differences in performance of AB42 assays, particularly when a ratio is not used?

Version 2:

Decision Letter:

6th Dec 2024

Dear Dr Lambert,

Your Letter, "Transferability of European-derived Alzheimer's Disease polygenic Risk Scores across Multi-Ancestry Populations" has now been seen by 1 referee. You will see from their comments below that while they find your work of interest, some important points are raised. We are interested in the possibility of publishing your study in Nature Genetics, but would like to consider your response to these concerns in the form of a revised manuscript before we make a final decision on publication.

We therefore invite you to revise your manuscript taking into account all reviewer and editor comments. Please highlight all changes in the manuscript text file. At this stage we will need you to upload a copy of the manuscript in MS Word .docx or similar editable format.

We are committed to providing a fair and constructive peer-review process. Do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

When revising your manuscript:

\*1) Include a "Response to referees" document detailing, point-by-point, how you addressed each referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the referees along with the revised manuscript.

\*2) If you have not done so already please begin to revise your manuscript so that it conforms to our Letter format instructions, available

[here](http://www.nature.com/ng/authors/article_types/index.html).

Refer also to any guidelines provided in this letter.

\*3) Include a revised version of any required Reporting Summary: <https://www.nature.com/documents/nr-reporting-summary.pdf>

It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review.

A revised checklist is essential for re-review of the paper.

Please be aware of our [guidelines on digital image standards](https://www.nature.com/nature-research/editorial-policies/image-integrity).

Please use the link below to submit your revised manuscript and related files:

Link Redacted

**Note:** This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

We hope to receive your revised manuscript within four to eight weeks. If you cannot send it within this time, please let us know.

Please do not hesitate to contact me if you have any questions or would like to discuss these revisions further.

Nature Genetics is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit [www.springernature.com/orcid](http://www.springernature.com/orcid).

We look forward to seeing the revised manuscript and thank you for the opportunity to review your work.

Sincerely,  
Wei

Wei Li, PhD  
Senior Editor  
Nature Genetics  
[www.nature.com/ng](http://www.nature.com/ng)

Reviewers' Comments:

Reviewer #1 (Remarks to the Author):

NG-LE62968R1

Transferability of a European-derived Alzheimer's disease genetic risk score across multi-ancestry populations  
Nicolas et al.

This manuscript has been significantly enhanced by adding several new analyses and datasets. My remaining comments largely concern the new Bayesian cross-ancestry analyses that were added to the manuscript during the revision.

- In PRS-CSx, the global shrinkage parameter does not correspond to marginal GWAS p-values so the authors should not justify the choice of  $1e-8$  in this way.

- Bayesian PRS methods try to incorporate secondary association signals in significant GWAS loci and sub-threshold

variants across the genome to improve prediction. Imposing a very strong shrinkage by setting the global shrinkage parameter to  $1e-8$  is at odds with the goal to model polygenic signals outside the APOE region.

- PRS-CSx was referred to both as a method and as a score (PRS) in the manuscript. The authors can perhaps better separate them and call the PRS built from PRS-CSx as a cross-ancestry PRS or something similar. It would also be good to distinguish "cross-ancestry" vs. "trans-ancestry" or stick to one term.

- It would be helpful to report the sample sizes of the African American, Latino American or East Asian GWAS in the main text. I cannot find this information even in the methods section. I suspect that the fact that the Bayesian PRS was not improved after integrating these GWAS was at least in part due to the very low sample sizes of these GWAS such that the cross-ancestry PRS was dominated by the European GWAS.

- Line 529: "a trans-ancestry AD-calibrated PRS-CSx" - I'm not sure what does "AD-calibrated" mean.

- Figure 6: The Nagelkerke  $R^2$  estimates were very low and did not seem to be in line with the OR and liability  $R^2$  estimates. I'm not sure this is an error but the authors should double check their calculations.

#### Other comments:

- I'd suggest more explicitly define the use of genetic risk scores (GRS), polygenic risk scores (PRS), and polygenic scores (PGS) upfront - e.g. at the beginning of the introduction such that the reader is not confused (i.e. explicitly say that you will use GRS for scores constructed from genome-wide significant loci, PRS for genome-wide scores and PRS for quantitative traits).

- The authors should look up the latest guidance on the use of genetic ancestry and race/ethnicity, and follow the best practices throughout the manuscript. The use of ancestry is currently confusing in some places. For example:

Line 378: "most of the available studies compared a European multi-ancestry population with another multi-ancestry population"

Line 466: "comparing European populations with the European-ancestry population"

I guess the authors wanted to distinguish individuals of European genetic ancestries from those who live in European countries. If so, the authors can clarify that in the text.

- Line 456: "admixed populations" --> "admixed populations"

- Line 484: I would suggest report ORs using the 40-60% stratum as the reference, which is consistent with the figure and other analyses in the manuscript. Using 0-20% as the reference inflates the effect size.

- Line 490: It's still not totally clear from the text how the percentage of African ancestry and Native American ancestry in the genome was calculated. Did you use the estimated probability from the PC-based ancestry assignment? Please clarify in the main text.

- Line 582: The lower OR in African-ancestry populations may be a result of effect size heterogeneity, differences in LD patterns, phenotyping differences, or simply low sample sizes, and thus does not necessarily suggest a different AD genetic architecture.

- Line 587: Missing bracket.

- Line 819: I agree with the other reviewer that sample overlap may be an issue. Previous studies have shown that even if the overlap is small, prediction accuracy can be significantly inflated. I totally understand that it is not practical to completely rule out sample overlap or relatedness between discovery GWAS and target datasets. Perhaps this can be discussed as a limitation.

- Line 840: "FlashPCA25" is supposed to be "FlashPCA2"?

- Line 917: I would suggest call this method an "adjustment" rather than a "calibration", as calibration often refers to producing accurate confidence intervals of the prediction.

Version 3:

Decision Letter:

Our ref: NG-LE62968R2

28th Feb 2025

Dear Dr. Lambert,

Thank you for submitting your revised manuscript "Transferability of European-derived Alzheimer's Disease Polygenic Risk Scores across Multi-Ancestry Populations" (NG-LE62968R2). It has now been seen by the original referees and their comments are below. The reviewers find that the paper has improved in revision, and therefore we'll be happy in principle to publish it in Nature Genetics, pending minor revisions to satisfy the referees' final requests and to comply with our editorial and formatting guidelines.

We are now performing detailed checks on your paper and will send you a checklist detailing our editorial and formatting requirements soon. Please do not upload the final materials and make any revisions until you receive this additional information from us.

Thank you again for your interest in Nature Genetics Please do not hesitate to contact me if you have any questions.

Sincerely,

Wei Li, PhD  
Senior Editor  
Nature Genetics  
[www.nature.com/ng](http://www.nature.com/ng)

Reviewer #1 (Remarks to the Author):

The authors have addressed my concerns.  
Note that 'coPRS' was never defined or mentioned in the main text.  
There are quite many typos throughout the manuscript.  
Careful proofreading or language editing is needed.

Version 4:

Decision Letter:

In reply please quote: NG-LE62968R3 Lambert

29th Apr 2025

Dear Dr. Lambert,

I am delighted to say that your manuscript "Transferability of European-derived Alzheimer's Disease Polygenic Risk Scores across Multi-Ancestry Populations" has been accepted for publication in an upcoming issue of Nature Genetics.

Over the next few weeks, your paper will be copyedited to ensure that it conforms to Nature Genetics style. Once your paper is typeset, you will receive an email with a link to choose the appropriate publishing options for your paper and our Author Services team will be in touch regarding any additional information that may be required.

After the grant of rights is completed, you will receive a link to your electronic proof via email with a request to make any corrections within 48 hours. If, when you receive your proof, you cannot meet this deadline, please inform us at [rjsproduction@springernature.com](mailto:rjsproduction@springernature.com) immediately.

You will not receive your proofs until the publishing agreement has been received through our system.

Due to the importance of these deadlines, we ask that you please let us know now whether you will be difficult to contact over the next month. If this is the case, we ask you provide us with the contact information (email, phone and fax) of someone who will be able to check the proofs on your behalf, and who will be available to address any last-minute problems.

Your paper will be published online after we receive your corrections and will appear in print in the next available issue. You can find out your date of online publication by contacting the Nature Press Office ([press@nature.com](mailto:press@nature.com)) after sending your e-proof corrections.

You may wish to make your media relations office aware of your accepted publication, in case they consider it appropriate to organize some internal or external publicity. Once your paper has been scheduled you will receive an email confirming the publication details. This is normally 3-4 working days in advance of publication. If you need additional notice of the date and time of publication, please let the production team know when you receive the proof of your article to ensure there is sufficient time to coordinate. Further information on our embargo policies can be found here: <https://www.nature.com/authors/policies/embargo.html>

Before your paper is published online, we shall be distributing a press release to news organizations worldwide, which may very well include details of your work. We are happy for your institution or funding agency to prepare its own press release, but it must mention the embargo date and Nature Genetics. Our Press Office may contact you closer to the time of publication, but if you or your Press Office have any enquiries in the meantime, please contact [press@nature.com](mailto:press@nature.com).

Acceptance is conditional on the data in the manuscript not being published elsewhere, or announced in the print or electronic media, until the embargo/publication date. These restrictions are not intended to deter you from presenting your data at academic meetings and conferences, but any enquiries from the media about papers not yet scheduled for publication should be referred to us.

**Authors may need to take specific actions to achieve [compliance](https://www.springernature.com/gp/open-research/funding/policy-compliance-faqs) with funder and institutional open access mandates.** If your research is supported by a funder that requires immediate open access (e.g. according to [Plan S principles](https://www.springernature.com/gp/open-research/plan-s-compliance)) then you should select the gold OA route, and we will direct you to the compliant route where possible. For authors selecting the subscription publication route, the journal's standard licensing terms will need to be accepted, including [those licensing terms](https://www.nature.com/nature-portfolio/editorial-policies/self-archiving-and-license-to-publish) will supersede any other terms that the author or any third party may assert apply to any version of the manuscript.

If you have any questions about our publishing options, costs, Open Access requirements, or our legal forms, please contact [ASJournals@springernature.com](mailto:ASJournals@springernature.com)

If you have posted a preprint on any preprint server, please ensure that the preprint details are updated with a publication reference, including the DOI and a URL to the published version of the article on the journal website.

To assist our authors in disseminating their research to the broader community, our SharedIt initiative provides you with a unique shareable link that will allow anyone (with or without a subscription) to read the published article. Recipients of the link with a subscription will also be able to download and print the PDF.

As soon as your article is published, you will receive an automated email with your shareable link.

You can now use a single sign-on for all your accounts, view the status of all your manuscript submissions and reviews, access usage statistics for your published articles and download a record of your refereeing activity for the Nature journals.

An online order form for reprints of your paper is available at <https://www.nature.com/reprints/author-reprints.html>. Please let your coauthors and your institutions' public affairs office know that they are also welcome to order reprints by this method.

If you have not already done so, we strongly recommend that you upload the step-by-step protocols used in this manuscript to protocols.io. protocols.io is an open online resource that allows researchers to share their detailed experimental know-how. All uploaded protocols are made freely available and are assigned DOIs for ease of citation. Protocols can be linked to any publications in which they are used and will be linked to from your article. You can also establish a dedicated workspace to collect all your lab Protocols. By uploading your Protocols to protocols.io, you are enabling researchers to more readily reproduce or adapt the methodology you use, as well as increasing the visibility of your protocols and papers. Upload your Protocols at <https://protocols.io>. Further information can be found at <https://www.protocols.io/help/publish-articles>.

Sincerely,  
Wei

Wei Li, PhD  
Senior Editor  
Nature Genetics

Click here if you would like to recommend Nature Genetics to your librarian  
<http://www.nature.com/subscriptions/recommend.html#forms>

\*\* Visit the Springer Nature Editorial and Publishing website at [http://editorial-jobs.springernature.com?utm\\_source=ejp\\_NGen\\_email&utm\\_medium=ejp\\_NGen\\_email&utm\\_campaign=ejp\\_NGen](http://editorial-jobs.springernature.com?utm_source=ejp_NGen_email&utm_medium=ejp_NGen_email&utm_campaign=ejp_NGen) for more information about our career opportunities. If you have any questions please click [here](mailto:editorial.publishing.jobs@springernature.com).



**Open Access** This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>



**Dr. Jean-Charles Lambert**

Tel +33 320 877 710

Fax +33 320 877 894

jean-charles.lambert@pasteur-lille.fr

**Nature Genetics**

Lille, October 31, 2024

Dear Editor, Dear Dr. Li,

Thank you very much for giving us the opportunity to revise our manuscript entitled "Transferability of European-derived Alzheimer's Disease polygenic Risk Score across Multi-Ancestry Populations."

We really want to thank the reviewers for their very constructive comments, and we are grateful to them for giving us the possibility to significantly improve the quality and relevance of our manuscript.

Both reviewers found the paper of particular interest and acknowledged that our work has many strengths including the large number of samples and datasets analyzed.

I would like to emphasize this latter point since unlike other human complex multi-factorial diseases for which it is possible to work on hundreds of thousands or even millions of cases/controls, the field of Alzheimer's disease (AD) is still limited by access to clinically diagnosed cases, even at the level of the European ancestry populations. This limited number of clinically diagnosed cases can be explained both by the continuing difficulty of access to diagnosis in developed countries and by the nature of the diagnosis itself, which is differential and requires multiple, often costly, approaches. Even it is going to change with the advent of plasma biomarkers, this still clearly limits the number of clinically diagnosed cases that can be recruited in developing countries, and therefore the number and size of case-control studies available in multi-ancestry populations.

We have thus probably succeeded in bringing together a large part of the genetic data available in the world, considering also the limitations of accessibility of raw genetic data and the constraints imposed by different national legislations. We have therefore systematically developed and shared scripts for carrying out analyses in accordance with these constraints and the mastery of analytical tools.

We would like to apologize for the time it has taken to revise the manuscript, but we have tried to do our best to respond to the various suggestions made by the reviewers, considering all the specificities described above. This requested a large amount of work in strong interaction with our collaborators to complete all the analyses requested:

- (i) We re-ran **all our analyses** based on the increase in the AD risk associated with an increment of one standard deviation in the GRS rather than as the increase in the risk of AD associated with an increment of one allele of average risk in the GRS. This was done to standardize and facilitate comparisons between the different populations studied.
- (ii) Although we had already included nearly 6,000 individuals for whom we had measurements of tau, p-tau and A $\beta$ 42 in cerebrospinal fluid, we were able **to add nearly 7,000 additional samples** in collaboration with Dr Carlos Cruchaga. This has allowed us to analyze the largest genetic database for these biomarkers to date and to perform additional stratification analyses.
- (iii) Where possible, we **systematically performed analyses with or without calibration** to correct for any shift in the GRS/PRS distributions in different populations.

- (iv) As requested by the reviewers, we included new analyses such as **sex stratification** or a **GRS (named GRS+)** that included additional SNPs at the European GWAS-defined loci, potentially associated with risk of developing AD in non-European multi-ancestry populations.
  - (v) We finally implemented **PRS-CSx, a Bayesian polygenic modelling method**, to construct a cross-ancestry polygenic risk score.
- 

### **Reviewer #1**

***The authors used PGS to refer to scores derived from genome-wide significant loci and used PRS to denote scores derived from genome-wide variants. This doesn't seem to be consistent with how the field uses these terms (again see PMID: 37620596).***

We did not use the term PGS (polygenic score, a term which covers all types of analyses based on cumulative effect scores) in our manuscript, but the terms "genetic risk score (GRS)" and "polygenic risk score (PRS)" to distinguish between models based on a limited number of SNPs in genome-wide significant associated loci or based on genome-wide approaches, respectively. As we have included a PRS-CSx approach in the revised manuscript, we believe it is useful for readers to retain the GRS and PRS designations to make the paper as comprehensive as possible.

However, as mentioned in the paper mentioned by the reviewer<sup>2</sup>, polygenic scores (PGS) have to be used for a general (quantitative) trait. Therefore, when looking at CSF concentrations or age at onset, we used the term PGS to be consistent with the field.

In summary, in the manuscript we have referred to GRS when including a limited number of SNPs in genome-wide associated loci, PRS when developing a genome-wide approach and PGS when analyzing quantitative traits. However, we will be happy to follow any nomenclature that the reviewer considers more appropriate.

***The authors used a PGS derived from 83 independent SNPs which was first constructed and evaluated in a 2022 Nature Genetics publication. This raises two issues:***

***(i) the novelty and overlap of this study with the 2022 publication, although new datasets have been added in the present study;***

The paper by Bellenguez et al. is a GWAS paper with the main objective of validating known AD genetic risk factors and of characterizing new ones in the most comprehensive way. This Bellenguez paper presents the association of the GRS derived from 83 SNPs associated with the risk of conversion to dementia in population-based or clinical prospective populations.

These longitudinal populations are not included in the current paper and all the analyses we present here are completely original, since the associations of GRSs/PGSs across Europe with AD risk, age at onset, sex and CSF biomarker concentrations were not presented in the Bellenguez et al paper<sup>1</sup>. In addition, as previously mentioned, this is the first time that a study compares to such an extent the association of GRSs/PRSs in so many different multi-ancestry populations which were not included at all in the Bellenguez et al. paper.

***(ii) it is known that for the vast majority of complex traits and diseases, PRS constructed by more sophisticated Bayesian methods substantially outperform PRS derived from genome-wide significant loci. I would love to see a comparison of the PGS with PRS derived from a modern approach such as PRS-CS and LDpred2.***

We thank the reviewer for this important suggestion. Our first aim was to keep our analyses as simple as possible to compare the known genetic component of AD between multi-ancestry populations. However, using a PRS derived from a modern approach is indeed a very good option and we have implemented a PRS-CSx as suggested. We chose the approach developed by Ge et al.<sup>3</sup>, which allows us not to use tuning populations. Indeed, as mentioned in the introductory paragraph of this letter, we are limited by the number of independent datasets to which we have access, particularly for African American ancestry

populations. We have compared PRS-CSx based on three types of summary statistics: one based solely on European multi-ancestry population, a second based solely on the multi-ancestry population of interest (either African American, Asian or Latin American), a third using combined summary statistics from European, African American, Asian and Latin American multi ancestry populations. Importantly, we systematically favored the use of summary statistics based on the largest number of individuals in each multi-ancestry population. In addition, we first defined a global shrinkage parameter at  $10^{-8}$  (to facilitate comparison with GRS). The results are described in line 518-554 of the manuscript and in Fig. 6 and extended Fig. 3 (and supplementary Table 10 and 11). In summary, when the APOE region was excluded to generate the PRS-CSx results, we did not detect any improvement in the score's level of performance, relative to the GRS whatever the multi-ancestry population studied. We discussed these results as well as potential limitations in line 598-606.

However, when including the APOE region when generating the PRS-CSx, we observed an increase in the strength of the association with the AD risk and better liability performance for PRS-CSx. This observation suggests that the APOE locus contains additional genetic information (along with the  $\epsilon 2/\epsilon 3/\epsilon 4$  alleles) in our various multi-ancestry populations and we discussed these results (line 607-620).

Of note, we also tested for additional global shrinkage parameter values, and we observed an increase in the AD risk or predictive performance in the MVP Latino American population when using a global shrinkage parameter at  $10^{-6}$  (Extended Fig. 3 and Supplementary Table 11). However, we cannot exclude the possibility that this improvement is due to overfitting.

***The results of the manuscript didn't provide much new insights into the transferability of PGS or AD genetics. A large number of studies have shown that for a wide range of complex traits and diseases, PGS derived from European GWAS have reduced predictive performance as the genetic distance between the training and target samples increase (btw- the authors failed to cite almost all of them). So the results in this manuscript are totally expected.***

AD presents a particular high genetic component (estimated between 60 and 80% of the attributable risk in twin studies<sup>4</sup>) and to the best of our knowledge, most of the AD studies seeking to compare this genetic component in multi-ancestry populations only included comparison between European and another multi-ancestry population in limited datasets. None of them had the possibility to perform the complete overview we present in this current work. So even if such works has been already performed in other human pathologies, this does not preclude the relevance of our work in the field of neurodegenerative disease and dementia. Since AD represents 50-70% of all types of dementia, our work is thus highly informative in this field.

It should be noted that in the Introduction and Discussion sections, we now cite several recent papers relating to other human pathologies that compare multi-ancestry pathologies using large samples.

In the discussion, we also highlighted the many messages that are important in the field of neurodegenerative diseases but could also be of great interest to a wider audience (lines 630-643).

***There are some interesting results - for example, the authors showed that in admixed populations the PGS prediction accuracy decreased as the genomic proportion of African and Native American ancestries increased. These are in line with the recent report (Ding et al. Nature 2023). However, the authors didn't dig deeper into this interesting observation. There is also no description in the methods section on how ancestry percentage was quantified. Did the authors perform any local ancestry inference?***

We apologize that this information was not enough visible (in the supplementary materials) and this is now added in the online materials (Line 833-840).

We did not perform any local ancestry inference analyses even if we are fully aware of the fact that such information is important to better understand the impact of genetic background in multiple phenotypes. For instance, local ancestry has been described to modulate APOE association with AD neuropathology and cognitive outcomes in an African American admixed population<sup>5</sup>. However, we considered that this interesting question deserves a specific analysis and publication to disentangle such a complexity at

every locus associated with AD risk in multi-ancestry populations. However, we have taken this point into account by analyzing whether genetic information specifically localized within the 75 loci associated with the AD risk and specific to multi-ethnic populations could improve the association of GRS with the risk of developing AD.

To this end, we selected SNPs associated with the risk of developing AD using summary statistics generated by Lake et al<sup>6</sup> and Shigemizu et al<sup>7</sup>. This led us to develop an extended GRS (called GRS+) containing 30, 13 and 47 additional variants to the initial 83 ones for Latino American, East Asian and African American ancestries respectively ( $p < 1 \times 10^{-3}$  and  $r^2 < 0.1$  with sentinel variants from loci observed in European multi-ancestry populations<sup>1</sup>). The approach is described in the online method in line 871-898

It is important to note that this GRS+ has not been analyzed in the Japanese population and in the ADGC populations (Latin American and Afro-American), since the summary statistics of these populations have been used to extract the  $\beta$  for the calculation of the GRS+. This approach showed no improvement in the association of GRS83+ with the risk of developing AD compared with GRS as described in lines 504-516 and in Supplementary Table 9).

***I would suggest a more systematic investigation of the relationship between the PGS and APOE alleles. What are the frequencies of APOE alleles in different populations?***

These data were already partly presented. To make them more visible, we included them in Figure 4. We also added a short description in lines 461-464.

***What was the relative contribution of the PGS and APOE to AD risk? What about the joint prediction of PGS + APOE? Reporting other PGS performance metrics in addition to OR including variance explained (on the liability scale) and AUC would also be helpful.***

We did not include AUC analyses, mainly because this tool is very sensitive to the structure of the case-control studies and, among other things, to the difference in age and sex distribution between cases and controls. Many of the case-control studies in this paper are not perfectly well age- or gender-matched, which means that the AUC tool mainly captures such demographic differences. This makes comparisons within and between populations difficult.

However, we systematically reported Nagelkerke pseudo- $R^2$  and liability scale pseudo- $R^2$ . We also pointed out some limitations related to the use of liability scale in the context of AD in the online material (lines 948-952).

Of note, we systematically assessed the impact of the APOE region in our PRS-CSx analyses,

***In the mega-analysis when pooling data across countries, it is important to assess if there is any shift of PGS distributions in different populations, and to calibrate PGS distributions before stratifying the PGS into quantiles (see e.g. the work from eMERGE consortium PMID: 3733246). The authors didn't assess the calibration of PGS in any way.***

We performed a mega-analysis to take full advantage of the statistical power of our European case/control studies by performing analyses by highly contrasted quantile, e.g. 98-100%. We performed this merge analysis because we did not find any heterogeneity between these European populations regarding the association of GRS with the risk of developing AD (Fig. 1).

However, to assess the important comment of the reviewer, we have first compared the GRS obtained in the meta- and mega-analysis and they are highly similar. As suggested by the reviewer, we have second carried out a meta-analysis after calibration (Supplementary Figure 5). We did not find any difference in the observed results, suggesting that the mega-analysis is not biased by a potential shift in the distribution of GRS between the different European populations.

Of note, we performed calibration for the GRS association in sub-Saharan Africa, Korean, and all MVP populations (see Figure 5 and supplementary table 10 for a comparison) and did not find any difference between non-calibrated and calibrated GRS<sup>ALZAdjAPOE</sup>. Unfortunately, we were not able to calibrate all the populations due to the unavailability of certain raw genetic data and the difficulty to implement the

necessary methods in some centers. Importantly, we have also carried out this calibration work for all the analyses based on the PRS-CSx approach.

***There is a significant disconnection with the latest PRS literature. The authors almost exclusively cited AD PRS work, some of which are quite outdated, and largely ignored the latest developments in the larger PRS community including some of the most important papers in the field, from Martin et al. showing the portability issue of PRS across ancestries, to a large body of methodological work on improving the predictive performance of PRS within and across populations, and to the more recent developments and discussions on the transferability of PRS across global populations (see e.g. PMID: 37620596 and many references cited therein). The authors should better review and acknowledge existing work.***

We acknowledged existing work as requested by the reviewer in both introduction and discussion.

***The manuscript was also not very well written with numerous typos. Some careful proofreading is needed.***

We read the document carefully to correct any typos and we asked an English-speaking writer to improve the English.

---

## **Reviewer #2:**

### ***Major Concern***

***While the authors highlight differences in the performance of the GRS by population, they do very little to try and address this major limitation within the manuscript. There is no adjustment for linkage disequilibrium and no attempt to improve performance in populations that are not well-represented. Moreover, there is no attempt to evaluate or demonstrate that the leveraged score narrowly focused on the top set of sentient variants within the top loci is indeed the best approach for such a score, and perhaps more importantly, whether such an approach by design hampers the ability to apply a GRS in other populations. There are many tools available to try and improve PRS performance across populations and such a goal seems within scope of this global assessment of a genetic risk score.***

We want to thank the reviewer for this suggestion that is of clear importance (as also suggested by reviewer 1). We have implemented a PRS-CSx approach developed by Ge et al.<sup>3</sup>, which allows us to not use tuning populations. Indeed, as mentioned in the introductory paragraph of this letter, we are limited by the number of independent datasets to which we have access, particularly for African American multi-ancestry populations. Importantly, for the first step of this approach, we combined summary statistics from European, African American, Asian and Latin multi ancestry populations and we systematically favored the use of the summary statistics based on the largest number of individuals in each multi-ancestry populations. The results are described in line 518-554 of the manuscript and in Fig. 6 and extended Fig. 3 (as well as in supplementary Table 10 and 11). In summary, when the APOE region was excluded to generate the PRS-CSx results, we did not detect any improvement in the score's level of performance, relative to the GRS whatever the multi-ancestry population studied. We discussed these results as well potential limitations in line 598-606.

However, when including the APOE region when generating the PRS-CSx, we observed an increase in the strength of the association with the AD risk and better liability performance for PRS-CSx. This observation suggests that the APOE locus contains additional genetic information (along with the  $\epsilon 2/\epsilon 3/\epsilon 4$  alleles) in our various multi-ancestry populations and we discussed these results (line 607-620).

Of note, we also tested for additional global shrinkage parameter values, and we observed an increase in the AD risk or predictive performance in the MVP Latino American population when using a global

shrinkage parameter at  $10^{-6}$  (Extended Fig. 3 and Supplementary Table 11). However, we cannot exclude the possibility that this improvement is due to overfitting.

***It is not clear why variants that have population specific effects were not integrated into this manuscript. The authors make a strong argument that African admixture may modify the association for some of the variants selected for this GRS and AD, but don't attempt to integrate variants that are known to show a stronger association in populations with African ancestry.***

This is an important point and we have taken this point into account by analysing whether genetic information specifically localized to the 75 loci associated with the risk of developing AD and specific to multi-ethnic populations could improve the association of GRS with the risk of developing Alzheimer's disease. To this end, we selected SNPs associated with the risk of developing AD using summary statistics generated by Lake et al<sup>6</sup> and Shigemizu et al<sup>7</sup>. This led us to develop an extended GRS (we called GRS+) containing 30, 13 and 47 additional variants to the initial GRS variants for Latino American, East Asian and African-American ancestries, respectively ( $p < 1 \times 10^{-5}$  and  $r^2 < 0.1$  with sentinel variants from loci observed in European multi-ancestry populations<sup>1</sup>). The approach is described in the online method in line 871-898.

It is important to note that this GRS+ has not been analyzed in the Japanese population and in the ADGC populations (Latin American and Afro-American), since the summary statistics of these populations have been used to extract the  $\beta$  for the calculation of the GRS+. This approach showed no improvement in the association of GRS+ with the risk of developing AD compared with GRS as described in line 504-516 and in Supplementary Table 9).

***Sample overlap is a major consideration in a study like this, but is not raised. Did the authors attempt to estimate sample overlap in any of the populations in which the GRS is assessed or adjust estimates for potential biases due to sample overlap? Were the participants included in the biomarker analysis also in the original GWAS?***

Sample overlap was systematically assessed and there was no sample overlap between any of the non-US studies analysed. Overlap between ADSP and MVP is likely to be negligible.

For biomarker analysis, the individuals included in the first submission are part of the EADB and the EADB samples were not included in the summary statistics used to generate the  $\beta$  incorporated for the PGS calculation. On the other hand, the samples we included for this revision are from C. Cruchaga's laboratory and there is overlap of 460 samples between the Cruchaga's ones and the ADGC (which is included in the summary statistics we used to generate the  $\beta$  for the GRS). However, this overlap is limited (less than 2.5%) and in addition, we only analyzed in Cruchaga's samples the association of GRS with quantitative traits ( $p$ -tau, tau and  $A\beta_{42}$  CSF concentrations).

***Considerations around the sex distributions across populations should be discussed, particularly given the preponderance of women in many of the cohorts used in the initial analyses and the preponderance of men in many of the studies used to assess the GRS in this manuscript.***

We of course fully agree with the reviewer that sex is an important variable in AD. To address this important issue, we performed a stratified analysis by sex in the European EADB populations and found no difference of the the GRS association with AD risk between men and women (line 399-401 and Figure 1).

### **Minor Concerns**

***The overall association for the GRS is incredibly low, with odds ratios <1.10 in most populations. Can the authors speculate on why GRS continue to perform so poorly outside of the APOE region given the high estimated heritability from twin studies?***

We presented the increase in AD risk associated with each additional allele of average risk in the GRS. This explains why the ORs are so low. We have now re-run all our analyses, and we present ORs as the

increase in risk of AD associated with each standard deviation of the GRS. The ORs have increased significantly.

***Given the variable coverage across global populations, and the desire to compare estimates across populations, it would seem more appropriate to limit the comparison to a score calculated on the set of variants available in all populations. Given the small number of variants included in the overall score, the current approach is dropping more than 10% of the variants in some populations.***

This is a point we discussed internally before deciding to use the number of SNPs available in each multi-ancestry sample. In the end, we chose a GRS that was as close as possible to the one generated by Bellenguez et al. in each of the populations studied. This meant that the number of variants could be lower than the 83 SNPs if they were not present or incorrectly imputed in the GWAS data for these populations. In addition, we found these data interesting because they also allowed us to highlight potential inter-population differences for loci associated with AD risk in European multi-ancestry populations. However, in line with the reviewer's comment, the presentation of our data as the increase in AD risk associated with each additional allele of average risk in the GRS is not appropriate, and this is one of the reasons that led us to re-run all our analyses and present the ORs as the increase in AD risk associated with each standard deviation in the GRS.

***The authors report a modest interaction between APOE and the GRS, and then state that there is no difference in stratified results and "APOE and GRS risk are independent). From the interaction result, and the stratified results, a more appropriate interpretation seems to be that the GRS has a slightly stronger association among E4 homozygotes but is equivalent in other strata.***

Given our results, it is very difficult to say what this weak interaction implies. The reviewer's interpretation is possible given the observed data. However, we cannot objectively demonstrate this, especially as the  $\epsilon 4\epsilon 4$  population is the smallest and therefore has the greatest measurement variability. In order to take into account the reviewer's comment, we modified the document accordingly:

*"Remarkably,  $GRS^{ALZ}$  appeared to be similarly associated with the AD risk in all the strata; this indicates that the APOE and  $GRS^{ALZ}$  risks may be independent, even if a stronger association might be present among  $\epsilon 4\epsilon 4$  carriers (Fig. 2B and Supplementary Table 4).*

***The authors raise a biological interpretation of the lack of heterogeneity in the GRS associations with tau compared to amyloid, but couldn't this just as easily be technical given differences in performance of AB42 assays, particularly when a ratio is not used?***

We fully agree with the reviewer, and we have mentioned this possibility in the manuscript (line 441-442).

---

We hope that you will be satisfied with the way in which we have responded to all the mainly methodological comments made by the two reviewers, whom we thank once again for helping us to improve our manuscript.

Sincerely,

Jean-Charles Lambert, PhD  
Inserm Research Director  
PI, European Alzheimer's & Dementia Biobank (EADB)





## References

1. Bellenguez, C. et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet* 54, 412–436 (2022).
2. Kachuri, L. et al. Principles and methods for transferring polygenic risk scores across global populations. *Nat Rev Genet* 25, 8–25 (2024).
3. Ge, T. et al. Development and validation of a trans-ancestry polygenic risk score for type 2 diabetes in diverse populations. *Genome Med* 14, 70 (2022)
4. Gatz, M. et al. Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry* 63, 168–74 (2006)
5. Naslavsky, M. S. et al. Global and local ancestry modulate APOE association with Alzheimer's neuropathology and cognitive outcomes in an admixed sample. *Mol Psychiatry* 27, 4800–4808 (2022).
6. Lake, J. et al. Multi-ancestry meta-analysis and fine-mapping in Alzheimer's disease. *Mol Psychiatry* 28, 3121–3132 (2023).
7. Shigemizu, D. et al. Ethnic and trans-ethnic genome-wide association studies identify new loci influencing Japanese Alzheimer's disease risk. *Transl Psychiatry* 11, 151 (2021).



**Dr. Jean-Charles Lambert**

Tel +33 320 877 710

Fax +33 320 877 894

jean-charles.lambert@pasteur-lille.fr

**Nature Genetics**

Lille, January 16, 2024

Dear Editor,

Thank you very much for giving us the opportunity to revise again our manuscript entitled "Transferability of European-derived Alzheimer's Disease polygenic Risk Score across Multi-Ancestry Populations". We thank the reviewer for acknowledging that our manuscript has been clearly enhanced. We again appreciated the comments and the effort of the reviewer to read the paper carefully. We have taken these new constructive comments into account.

- 
- 1. In PRS-CSx, the global shrinkage parameter does not correspond to marginal GWAS p-values so the authors should not justify the choice of  $1e-8$  in this way.**
  - 2. Bayesian PRS methods try to incorporate secondary association signals in significant GWAS loci and sub-threshold variants across the genome to improve prediction. Imposing a very strong shrinkage by setting the global shrinkage parameter to  $1e-8$  is at odds with the goal to model polygenic signals outside the APOE region.**

These two points are related, and we agree with the reviewer. In the previous version of the manuscript, we already showed the PRS results for a sparseness parameter of  $10^{-6}$ . To address these two points, we have decided to show the results for all the sparseness parameters we tested (at  $10^{-8}$ ,  $10^{-7}$ ,  $10^{-6}$ ,  $10^{-5}$ ,  $10^{-4}$ ,  $10^{-2}$  and 1) in the revised figure 5 and 6 (excluding or including the *APOE* region) as well as in supplementary revised figure 14 (EPIDEMCA study).

- 3. PRS-CSx was referred to both as a method and as a score (PRS) in the manuscript. The authors can perhaps better separate them and call the PRS built from PRS-CSx as a cross-ancestry PRS or something similar. It would also be good to distinguish "cross-ancestry" vs. "trans-ancestry" or stick to one term.**

We thank the reviewer for this comment. To avoid potential confusion, we have now only used PRS-CSx when referring to the method/software and have systematized the use of (cross-ancestry) PRS as suggested.

- 4. It would be helpful to report the sample sizes of the African American, Latino American or East Asian GWAS in the main text. I cannot find this information even in the methods section. I suspect that the fact that the Bayesian PRS was not improved after integrating these GWAS was at least in part due to the very low sample sizes of these GWAS such that the cross-ancestry PRS was dominated by the European GWAS.**

This information was available in the supplementary Figure 13. However, to make this information easier to find, we have included it in the main manuscript as suggested (Lines 459-461).

We also fully agree with the reviewer that it is probable that the Bayesian PRS did not systematically improve AD risk increase and predictive performance in diverse ancestry populations due to the lack of statistical power associated with the small size of the multi-ancestry GWASs. This is a limitation that we mentioned in the previous manuscript. Although we had to significantly reduce the size of the paper to fit the letter format, we have retained this important comment in the discussion of the revised document (lines 499-501).

- 5. Line 529: "a trans-ancestry AD-calibrated PRS-CSx" - I'm not sure what does "AD-calibrated" mean.**

This was a typo and has been corrected.

- 6. Figure 6: The Nagelkerke R<sup>2</sup> estimates were very low and did not seem to be in line with the OR and liability R<sup>2</sup> estimates. I'm not sure this is an error but the authors should double check their calculations.**

We presented an adjusted Nagelkerke's Pseudo-R<sup>2</sup>, which is the difference between a Pseudo-R<sup>2Full</sup> and a Pseudo-R<sup>2Null</sup>. The Nagelkerke's Pseudo-R<sup>2Full</sup> includes both the polygenic score and the covariates, and the Nagelkerke's Pseudo-R<sup>2Null</sup> includes only the covariates (see Online Methods). We used a well-established package, rcompanion in R, and found no errors after carefully double checking.

Unfortunately, there are a few studies in AD that used the adjusted Nagelkerke's pseudo-R<sup>2</sup>. In Kikuchi et al<sup>1</sup>, this pseudo-R<sup>2</sup> was calculated in a Japanese population using a PRS method that was not the same as the one we used. In our Korean population, which is probably the closest population in terms of ancestry, we observed a similar range of values for the Nagelkerke's pseudo-R<sup>2</sup>, suggesting that our results are consistent with previous works.

To note, it is not clear for us why the reviewer mentioned that the adjusted Nagelkerke's pseudo-R<sup>2</sup> is not consistent with the OR and liability R<sup>2</sup>. All values follow the same evolution according to the different PGS/PRS calculated in the same ancestry.

- 7. I'd suggest more explicitly define the use of genetic risk scores (GRS), polygenic risk scores (PRS), and polygenic scores (PGS) upfront - e.g. at the beginning of the introduction such that the reader is not confused (i.e. explicitly say that you will use GRS for scores constructed from genome-wide significant loci, PRS for genome-wide scores and PRS for quantitative traits.**

To make the paper more readable and to save space in the letter format as required by the editor, we finally decided to follow the advice of reviewer 1 at the first revision step.

We ended up using the term polygenic score (PGS), a term that covers all types of analyses based on cumulative effect scores, for scores constructed from genome-wide significant loci (previously GRS) and for quantitative traits (already PGS). In the manuscript, we systematically use the term PGS<sup>ALZ</sup>.

As previously advised (see comment 3), we have also systematized the use of cross ancestry PRS (PRS) when referring to the results generated by PRS-CSx.

- 8. The authors should look up the latest guidance on the use of genetic ancestry and race/ethnicity, and follow the best practices throughout the manuscript. The use of ancestry is currently confusing in some places. For example: Line 378: "most of the available studies compared a European multi-ancestry population with another multi-ancestry population" Line 466: "comparing European populations with the European-ancestry population". I guess the**

**authors wanted to distinguish individuals of European genetic ancestries from those who live in European countries. If so, the authors can clarify that in the text.**

We did our best to follow the best practices and corrected the manuscript accordingly. The reviewer was right in the way he/she understood the two sentences, and we have made it clearer.

**9. Line 456: "admixed populations" --> "admixed populations"**

This typo has been corrected.

**10. Line 484: I would suggest report ORs using the 40-60% stratum as the reference, which is consistent with the figure and other analyses in the manuscript. Using 0-20% as the reference inflates the effect size.**

We have removed this paragraph describing results using the 0-20% strata as a reference.

**11. Line 490: It's still not totally clear from the text how the percentage of African ancestry and Native American ancestry in the genome was calculated. Did you use the estimated probability from the PC-based ancestry assignment? Please clarify in the main text.**

We used a classical approach based on SNPweights, and this is described in the Methods online (lines 678-686), in the supplementary materials (ADSP section) as well as in<sup>2</sup>. Because of the format letter, it is difficult to include such technical information in the main text.

**12. Line 582: The lower OR in African-ancestry populations may be a result of effect size heterogeneity, differences in LD patterns, phenotyping differences, or simply low sample sizes, and thus does not necessarily suggest a different AD genetic architecture.**

We agree with the reviewer and the paragraph has been removed.

**13. Line 819: I agree with the other reviewer that sample overlap may be an issue. Previous studies have shown that even if the overlap is small, prediction accuracy can be significantly inflated. I totally understand that it is not practical to completely rule out sample overlap or relatedness between discovery GWAS and target datasets. Perhaps this can be discussed as a limitation.**

In terms of relatedness, this has been controlled within all GWAS/sequencing datasets. In addition, we also controlled this between all European or South American GWAS datasets. It is very unlikely that we have undetected relatedness between the Chinese, Korean and Japanese populations. We cannot rule out the possibility of undetected relatedness between MVP and ADSP samples, but this is likely to affect very few samples.

We have systematically checked for overlap between samples. We have added a paragraph in the Methods online to describe this (lines 660-668).

**14. Line 840: "FlashPCA25" is supposed to be "FlashPCA2"?**

This was a typo and has been corrected.

**15. Line 917: I would suggest call this method an "adjustment" rather than a "calibration", as calibration often refers to producing accurate confidence intervals of the prediction.**

We developed the approach described in Hao et al<sup>3</sup>, and the authors indeed preferred the term adjustment rather than calibration. As suggested by the reviewer, we have now indicated that we performed an adjustment for the difference in PGS<sup>ALZ</sup>/PRS distribution between populations. We have replaced the term "calibration" throughout the manuscript accordingly.

---

In addition, we have reduced the text from 3824 to 2234 words (including introducing paragraph and main text). We hope this is acceptable.

To comply with the letter format, we have now 6 figures in the main document (180 mm width and at least at least 300 dpi) and 4 extended figures.

We hope that you will be satisfied with the way in which we have responded to all comments made by the reviewer, whom we thank once again for helping us to improve our manuscript.

Sincerely,



Jean-Charles Lambert, PhD  
Inserm Research Director  
coordinator, European Alzheimer's & Dementia Biobank (EADB)

## References

1. Kikuchi M, et al. Alzheimer's Disease Neuroimaging Initiative; Japanese Alzheimer's Disease Neuroimaging Initiative. Polygenic effects on the risk of Alzheimer's disease in the Japanese population. *Alzheimers Res Ther.* 2024 Feb 27;16(1):45.
2. Le Guen Y, et al. Multiancestry analysis of the HLA locus in Alzheimer's and Parkinson's diseases uncovers a shared adaptive immune response mediated by HLA-DRB1\*04 subtypes. *Proc Natl Acad Sci U S A.* 2023 Sep 5;120(36):e2302720120.
3. Hao, L. et al. Development of a clinical polygenic risk score assay and reporting workflow. *Nat Med* 28, 1006–1013, (2022).