



Open Access This 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 the cases where the authors are anonymous, such as is the case for the reports of anonymous peer reviewers, author attribution should be to 'Anonymous Referee' followed by a clear attribution to the source work. The images or other third party material in this 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 <http://creativecommons.org/licenses/by/4.0/>.

Editorial Note: This manuscript has been previously reviewed at another journal that is not operating a transparent peer review scheme. This document only contains reviewer comments and rebuttal letters for versions considered at Nature Communications.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

I appreciate the effort that the authors have put into the revisions. Most of my previous comments have been adequately addressed. The addition of a proper replication cohort instead of the UKBB family history GWAS is an important improvement. I have no suggestions for further improvement.

Reviewer #2 (Remarks to the Author):

I have previously reviewed this paper and they have addressed all my comments. The paper is greatly improved, especially with the addition of a more powerful replication set.

As a purely optional suggestion, the new section on the replication – lines 202 – 235 – would be strengthened by an overall graphical display of the new data vs the replication set. E.g. an effect size plot as done in sup figure 3d, or even something like using IVW two sample MR to show that the instrument (lead SNPs from your discovery GWAS) are associated with lung cancer risk in aggregate in the replication set. This can be clearer than a list of significant loci or SNPs. Actual content is fine.