

**From:** Robert Kaplan robert.kaplan@einsteinmed.edu  
**Subject:** Re: Height WGS Paper (Exeter) for your review (TOPMed as Replication)  
**Date:** October 4, 2023 at 11:40  
**To:** Li, Xihao xihaoli@unc.edu  
**Cc:** Li, Zilin li@hsph.harvard.edu, North, Kari E kari\_north@unc.edu, Justice, Anne E. aejustice1@geisinger.edu, Ching-Ti Liu ctliau@bu.edu, Xihong Lin xlin@hsph.harvard.edu, Jennifer Brody jeco@uw.edu



Dear Xihao,

First of all I apologize for overlooking your email from September. I also want to thank you for the careful consideration of the concern we raised.

The solution is good, insofar as you are describing accurately the method for inferring that the participant group had a predominant proportion of European ancestry. The term White is a racial designation, which can only be self report and that was why we objected to that term.

We do have published work using a method for inferring group membership down to the "national background", so this might be something to consider in future work. [Genetic Diversity and Association Studies in US Hispanic/Latino Populations: Applications in the Hispanic Community Health Study/Study of Latinos - PMC \(nih.gov\)](#)

Thank you and you can consider this manuscript approved by HCHS SOL now.

Robert

Robert Kaplan, PhD (he/him)  
Albert Einstein College of Medicine  
718-430-4076 (New York)  
206-667-6913 (Seattle)

**From:** Li, Xihao <xihaoli@unc.edu>  
**Sent:** Wednesday, October 4, 2023 8:13 AM  
**To:** Robert Kaplan <robert.kaplan@einsteinmed.edu>  
**Cc:** Li, Zilin <li@hsph.harvard.edu>; North, Kari E <kari\_north@unc.edu>; Justice, Anne E. <aejustice1@geisinger.edu>; Ching-Ti Liu <ctliau@bu.edu>; Xihong Lin <xlin@hsph.harvard.edu>; Jennifer Brody <jeco@uw.edu>  
**Subject:** Re: Height WGS Paper (Exeter) for your review (TOPMed as Replication)

**CAUTION: This email comes from an external source, the attachments and/or links may compromise our secure environment. Do not open or click on suspicious emails. Please click on the "Phish Alert" button on the top right of the Outlook dashboard to report any suspicious emails.**

Dear Dr. Kaplan,

Good morning and hope you are doing well. We are writing to follow up and kindly ask if the revised manuscript and supplementary table are adequate and appropriate for the study-specific approval from HCHS/SOL?

Thank you very much for your time and support. We look forward to hearing from you.

Best,  
Xihao & Zilin

On Sep 26, 2023, at 13:17, Li, Xihao <xihaoli@unc.edu> wrote:

Dear Dr. Kaplan,

Thank you very much for taking your time to review the manuscript and for bringing your questions to us. We went back from beginning and double checked the study demographics and phenotype/covariate harmonization. Specifically, the population group definitions and anthropometric traits harmonization (including height, BMI, etc.) in TOPMed were preprocessed following the same procedure. Specifically, the population groups were created using the HARE method from work of DCC, Jen (cc'ed) and us downloaded these HARE groups the DCC created from the Exchange Area. If the ancestry group was "Missing" or "Other", HARE (<https://pubmed.ncbi.nlm.nih.gov/31564439/>) used genetic information to assign the sample to a group. In addition, the TOPMed Anthropometry-Adiposity Working Group has recently posted a [report on medRxiv](#), detailing the methods used for the study demographics for freeze 8 BMI analysis.

To remedy the confusion and make it consistent with what we have actually performed and with the BMI paper, we have

(1) revised the supplementary table 1 (ST1 of Supplementary Tables\_Final\_v22.xlsx) to make it consistent with the BMI demographics (Supplementary Table 1 in page 49 of TOPMed-F8-BMI-medRxiv-Supp.pdf) as follows:

HCHS_SOL_CentralAmerican	HCHS_SOL_Cuban	HCHS_SOL_Dominican	HCHS_SOL_Mexican	HCHS_SOL_PuertoRican	HCHS_SOL_SouthAmerican	HCHS_SOL_Euro
509 (0.58%)	1901 (2.17%)	1126 (1.28%)	1732 (1.98%)	2105 (2.40%)	297 (0.34%)	33 (0.04%)
308 (60.51%)	989 (52.03%)	739 (65.63%)	1049 (60.57%)	1207 (57.34%)	181 (60.94%)	15 (45.45%)
44.73	49.12	45.50	44.06	47.41	46.49	50.70
13.66	13.08	14.25	13.97	14.27	13.09	15.21
160.48	164.04	162.11	161.73	163.22	162.11	166.42
8.84	9.18	8.89	9.49	9.46	8.87	9.00

(2) included a paragraph in the height manuscript to explain our procedure for TOPMed in the main text (Height\_WGS\_Paper\_200k\_v22\_untracked.docx) as follows:

"The WGS was performed at a target depth of >30x using DNA extracted from blood. We analyzed 87,652 multi-population samples from 33 studies in the freeze 8 TOPMed (ST1). Population group was defined by self-reported information from participant questionnaires in each study. For individuals who had unreported or non-specific population memberships (e.g., "Multiple" or "Other"), we applied the Harmonized Ancestry and Race/Ethnicity (HARE) method (Fang et al. 2019; Zhang et al. 2023) to infer their group memberships using genetic data. The population groups were thus labeled by their self-identified or primary inferred population group. Among the 87,652 participants, 52,519 (60%) were female and 44,846 (51%) were non-European. Additional descriptive tables of the participants are presented in ST1."

We hope that the revised manuscript and supplementary table is adequate and appropriate for the study-specific approval from HCHS/SOL. Please let us know if you have any questions or comments.

Thank you very much for your support. We look forward to hearing from you.

Best,  
Xihao & Zilin

On Sep 20, 2023, at 10:55, Robert Kaplan <robert.kaplan@einsteinmed.edu> wrote:

Dear Xihao,

Thanks for the follow up.

Unfortunately I cannot approve this. The information in ST1 looks unusual for HCHS SOL.

Below I pasted the HCHS SOL groups and the Ns. We do not have a "White" group, and the paper does not explain what this means. It does not sound like the correct label.

HCHS_SOL_CentralAmerican	HCHS_SOL_Cuban	HCHS_SOL_Dominican	HCHS_SOL_Mexican	HCHS_SOL_PuertoRican	HCHS_SOL_SouthAmerican	HCHS_SOL_White
509	1901	1126	1732	2105	297	33

Robert

Robert Kaplan, PhD (he/him)  
Albert Einstein College of Medicine  
718-430-4076 (New York)

206-667-6913 (Seattle)

**From:** Li, Xihao <xihao@unc.edu>  
**Sent:** Wednesday, September 20, 2023 7:45 AM  
**To:** Robert Kaplan <robert.kaplan@einsteinmed.edu>  
**Cc:** Li, Zilin <li@hsph.harvard.edu>; North, Kari E <kari\_north@unc.edu>  
**Subject:** Re: Height WGS Paper (Exeter) for your review (TOPMed as Replication)

**CAUTION: This email comes from an external source; the attachments and/or links may compromise our secure environment. Do not open or click on suspicious emails. Please click on the "Phish Alert" button on the top right of the Outlook dashboard to report any suspicious emails.**

Dear Dr. Kaplan and Dr. North,

Thank you very much for your comments and feedback. Attached please find a copy of revised manuscript for the Height WGS Paper. Could we kindly ask if the revised manuscript can be reviewed and approved by HCHS/SOL?

Please let us know if you have any additional comments and suggestions.

Thank you very much for your support. We look forward to hearing from you soon.

Best,  
Xihao & Zilin

On Sep 11, 2023, at 14:00, Robert Kaplan <robert.kaplan@einsteinmed.edu> wrote:

Dear Xihao,

OK I understand now.

One thing I'm looking at is the funding acknowledgement, and that all looks correct.

The part that we need to see is the additional information that you mention, which will be added. To describe the cohorts including HCHS SOL in terms of who was included (N) and any relevant data descriptions. When you have that ready, please recontact me and we can finalize the SOL publications committee review.

Thanks,

Robert

-----  
Robert Kaplan, PhD (he/him)  
Albert Einstein College of Medicine  
718-430-4076 (New York)  
206-667-6913 (Seattle)

**From:** Li, Xihao <xihao@unc.edu>  
**Sent:** Thursday, September 7, 2023 8:32 AM  
**To:** Robert Kaplan <robert.kaplan@einsteinmed.edu>  
**Cc:** Li, Zilin <li@hsph.harvard.edu>  
**Subject:** Re: Height WGS Paper (Exeter) for your review (TOPMed as Replication)

**CAUTION: This email comes from an external source; the attachments and/or links may compromise our secure environment. Do not open or click on suspicious emails. Please click on the "Phish Alert" button on the top right of the Outlook dashboard to report any suspicious emails.**

Dear Dr. Kaplan,

Thank you very much for taking time to read through the manuscript. In this manuscript, the UK Biobank data was used for discovery of height, and TOPMed data was only used for replication study of the significant rare variant associations. We confirm that HCHS/SOL data was included for TOPMed replication. Specifically, we did two analysis: (1) using all TOPMed Freeze 8 data, including HCHS/SOL; (2) using individuals of "European" background only.

Sure, we will communicate with the Exeter team to include supplemental table describing the populations included: UKB, AoU and TOPMed cohorts. May we kindly ask if you would like to revisit the manuscript and supplementary materials again before granting the approval from HCHS/SOL?

Many thanks for your support.

Best,  
Xihao & Zilin

On Sep 5, 2023, at 20:50, Robert Kaplan <robert.kaplan@einsteinmed.edu> wrote:

You don't often get email from robert.kaplan@einsteinmed.edu. [Learn why this is important](#)

Dear Xihao,

Well done manuscript.

With regard to my TOPMed affiliation, which is HCHS/SOL, I am unclear about whether our cohort was part of this analysis.

Related to this, the manuscript states that individuals of "European" background were included. In our Hispanic HCHS/SOL cohort, of course, the great majority of people have some European heritage (among other continental ancestries), and for some people the proportion of their ancestors who were "European" is quite high indeed. The same goes for some self identified "African Americans" for that matter, although less so than for US Hispanics. So what is your criterion for what persons count as "European" and which do not?

Also the paper needs a supplemental table describing the populations included: UKB, AoU, and the constituent TOPMed cohorts.

Robert

-----  
Robert Kaplan, PhD (he/him)  
Albert Einstein College of Medicine  
718-430-4076 (New York)  
206-667-6913 (Seattle)

**From:** Li, Xihao <xihao@unc.edu>  
**Sent:** Tuesday, September 5, 2023 7:18 AM  
**To:** Mitchell, Braxton <bmitchel@som.umaryland.edu>; Boerwinkle, Eric <Eric.Boerwinkle@uth.tmc.edu>; Kathleen.barnes@cuanschutz.edu <Kathleen.barnes@cuanschutz.edu>; ruth.loos@mssm.edu <ruth.loos@mssm.edu>; Myriam.Fornage@uth.tmc.edu <Myriam.Fornage@uth.tmc.edu>; chungm@ccf.org <chungm@ccf.org>; sredline1@rics.bwh.harvard.edu <sredline1@rics.bwh.harvard.edu>; murabito@bu.edu <murabito@bu.edu>; rmathia2@jhmi.edu <rmathia2@jhmi.edu>; Jhe <Jhe@tulane.edu>; Robert Kaplan <robert.kaplan@einsteinmed.edu>; April P. Carson <apcarson@umc.edu>; Ashrani.Aneel@mayo.edu <Ashrani.Aneel@mayo.edu>; ellinor@mgh.harvard.edu <ellinor@mgh.harvard.edu>; Allison Ashley-Koch, Ph.D. <allison.ashleykoch@duke.edu>; slubitz@mgh.harvard.edu <slubitz@mgh.harvard.edu>; Joanne Curran <joanne.curran@utrgv.edu>; rminster

<[rminster@pitt.edu](mailto:rminster@pitt.edu)>; [moore.b.shoemaker@vumc.org](mailto:moore.b.shoemaker@vumc.org) <[moore.b.shoemaker@vumc.org](mailto:moore.b.shoemaker@vumc.org)>; [darbar@uic.edu](mailto:darbar@uic.edu) <[darbar@uic.edu](mailto:darbar@uic.edu)>; [vgordeuk@uic.edu](mailto:vgordeuk@uic.edu) <[vgordeuk@uic.edu](mailto:vgordeuk@uic.edu)>; [Christine.Albert@cshs.org](mailto:Christine.Albert@cshs.org); [Christine.Albert@cshs.org](mailto:Christine.Albert@cshs.org); Chami, Nathalie <[nathalie.chami@mssm.edu](mailto:nathalie.chami@mssm.edu)>; HCM(胡啟民主任) <[chhwu@vghtpe.gov.tw](mailto:chhwu@vghtpe.gov.tw)>; Zhou, Hufeng <[hufengzhou@g.harvard.edu](mailto:hufengzhou@g.harvard.edu)>  
Cc: Li, Zilin <[li@hsph.harvard.edu](mailto:li@hsph.harvard.edu)>; Xihong Lin <[xlin@hsph.harvard.edu](mailto:xlin@hsph.harvard.edu)>; Ching-Ti Liu <[ctliu@bu.edu](mailto:ctliu@bu.edu)>; North, Kari E <[kari\\_north@unc.edu](mailto:kari_north@unc.edu)>; Justice, Anne E. <[aejustice1@geisinger.edu](mailto:aejustice1@geisinger.edu)>  
Subject: Re: Height WGS Paper (Exeter) for your review (TOPMed as Replication)

**CAUTION: This email comes from an external source; the attachments and/or links may compromise our secure environment. Do not open or click on suspicious emails. Please click on the "Phish Alert" button on the top right of the Outlook dashboard to report any suspicious emails.**

Dear all,

Hope you are doing well. We sent out this paper **Whole genome association testing in 333,100 individuals across three Biobanks identifies rare non-coding single variant and genomic aggregate associations with height** earlier for your review. Please find the latest version of height manuscript and supplementary materials led by the team from University of Exeter (UK) based on the comments and feedbacks we have collected, as well as a copy of the approval TOPMed paper proposal for this manuscript (proposal id #15383) attached.

Manuscript: [https://www.dropbox.com/scl/fi/pywhy410ittgk88bw5en/Height\\_WGS\\_Paper\\_200k\\_v19\\_Untracked.docx?rlkey=p4s8u5nxl9j6cadlrpdvmhrki&dl=0](https://www.dropbox.com/scl/fi/pywhy410ittgk88bw5en/Height_WGS_Paper_200k_v19_Untracked.docx?rlkey=p4s8u5nxl9j6cadlrpdvmhrki&dl=0)

Study design: [https://www.dropbox.com/scl/fi/fugbsht29up71yy1soi4t/Study\\_Design\\_Figure.pdf?rlkey=z3jfg8vid5ifm1iqg5tov4vkw&dl=0](https://www.dropbox.com/scl/fi/fugbsht29up71yy1soi4t/Study_Design_Figure.pdf?rlkey=z3jfg8vid5ifm1iqg5tov4vkw&dl=0)

Supplementary Tables: [https://www.dropbox.com/scl/fi/8mgszfh52d2zuvtdf9g00/Supplementary\\_Tables\\_Final.xlsx?rlkey=srlrabewc1e1v5uch1s91e4v3s&dl=0](https://www.dropbox.com/scl/fi/8mgszfh52d2zuvtdf9g00/Supplementary_Tables_Final.xlsx?rlkey=srlrabewc1e1v5uch1s91e4v3s&dl=0)

Supplementary Note (Acknowledgement): [https://www.dropbox.com/scl/fi/60k93q03krtogim04zx9b/Height\\_WGS\\_Paper\\_200k\\_TOPMED\\_Acknowledgements\\_08222023.docx?rlkey=pllx7vc3msl3etyppo2xjorm&dl=0](https://www.dropbox.com/scl/fi/60k93q03krtogim04zx9b/Height_WGS_Paper_200k_TOPMED_Acknowledgements_08222023.docx?rlkey=pllx7vc3msl3etyppo2xjorm&dl=0)

Could you please let us know any additional comments/suggestions on the content of manuscript and supplementary materials?

We are also writing to kindly follow up regarding the study-specific approvals of this manuscript. Specifically, TOPMed ACC have informed us that, unless there is a specific need for the P&P committee approval, we can assume our studies are automatically approved if we do not receive any feedback within two weeks. Based on this guidance, we would like to seek your approval for each study by **September 19 (Tuesday)**. Your timely feedback would be greatly appreciated, ensuring we maintain compliance and align this work with the guidelines set forth. Please let us know if there are any additional procedures or requirements we should be aware of.

Thank you very much for your collaboration and support.

Best,  
Xihao & Zilin

On Jul 31, 2023, at 15:21, Li, Xihao <[xihaoli@hsph.harvard.edu](mailto:xihaoli@hsph.harvard.edu)> wrote:

Dear all,

Hope you are doing well. We're happy to share the draft of the Height UKBB WGS Paper for your review/feedback. This project is led by the team from the University of Exeter, UK. The manuscript analyzes the UKBB Height WGS data and uses the TOPMed Height WGS data as the replication. The TOPMed proposal for this manuscript has been approved (proposal id #15383).

Manuscript: [https://www.dropbox.com/scl/fi/3y5wnf622d56cp1a3l9x/Height\\_WGS\\_Paper\\_200k\\_TOPMED\\_Approval\\_July2023.docx?rlkey=kv9tk5kyyps2of2zen0xfj03&dl=0](https://www.dropbox.com/scl/fi/3y5wnf622d56cp1a3l9x/Height_WGS_Paper_200k_TOPMED_Approval_July2023.docx?rlkey=kv9tk5kyyps2of2zen0xfj03&dl=0)

Supplementary Note (Acknowledgement): [https://www.dropbox.com/scl/fi/hxgmpfpxv4demb9e24u2/Height\\_WGS\\_Paper\\_200k\\_TOPMED\\_Acknowledgements.docx?rlkey=etcxp7r56r56ld2nwu34jp0l2&dl=0](https://www.dropbox.com/scl/fi/hxgmpfpxv4demb9e24u2/Height_WGS_Paper_200k_TOPMED_Acknowledgements.docx?rlkey=etcxp7r56r56ld2nwu34jp0l2&dl=0)

TOPMed proposal (Approved): <https://www.dropbox.com/scl/fi/1vyqj53l3ryzbe8xa0sh/Revised-Replication-of-rare-non-coding-anthropometric-trait-associations-from-whole-genome-sequencing-downloaded2023-07-10.pdf?rlkey=8bx85xq23gh0wr92l3cz8kok&dl=0>

At convenience, could you please also send the manuscript to your corresponding P&Ps (or let us know if you'd like to do this & which email address to send it to) for study-specific approval of this manuscript? As discussed with the TOPMed Anthropometry Working Group, one author representing each study has currently been included for this TOPMed replication paper. Please kindly let us know if you have any additional comments/suggestions regarding this.

We recognize that many of you are on or about to go on vacation. We are hoping we get feedback from everyone by **August 13 (Sunday)**; please let us know if you need more time.

Please do not hesitate to reach out with any questions/comments. Many thanks for your collaboration and we look forward to your constructive feedback (as well as the administrative guidance – P&Ps and approval!)

Best wishes,  
Xihao & Zilin on behalf of the TOPMed Anthropometry WG

<-Height\_WGS\_Paper\_200k\_v22\_untracked.docx><-Height\_WGS\_Paper\_200k\_v22.docx><-Supplementary\_Tables\_Final\_v22.xlsx><-TOPMed-F8-BMI-medRxiv.pdf><-TOPMed-F8-BMI-medRxiv-Supp.pdf>

