#81 Improvement of a clinical colorectal cancer risk prediction model integrating polygenic risk

Erika Spaeth[1], Aviv Gafni[2], Richard Allman[2], Gillian Dite[2]
1-Phenogen Sciences, Inc; 2-Genetic Technologies, Ltd

Background

Improving colorectal cancer risk prediction and stratification is pivotal for implementing better screening and prevention programs in public health and for enabling a personalized approach to assessing patients' colorectal cancer risk.
Current US colorectal cancer screening recommendations have been changed to a start age of 45 in an attempt to address increasing incidence of the disease in young adults. Despite the effectiveness of screening, compliance remains a challenge.
Currently, when clinicians engage in joint-decision-making-discussions with their patients about screening options, they risk-assess their patients' risk by looking at family history, age, sex and ethnicity—most often as stand-alone risk factors.
We and others have show the value of combining epidemiological risk factors into a single model for improved risk assessment. Herein we show further improvement on a previous iteration of risk assessment that incorporates age, sex, ethnicity, first-degree family history, and polygenic risk with the replacement of a 45-SNP PRS with a 140-SNP PRS.

Methods

We used the UK Biobank to compare the performance of a risk prediction model incorporating two different polygenic risk scores – one comprising 45 SNPs[1] and the other comprising 140 SNPs.[2,3]
The clinical component of the risk prediction model included a simple measure of first-degree family history. We used age- and sex-specific population incidence rates to calculate 10-year and full-lifetime risks as previously described.[4]
The UK Biobank comprises 500,000 volunteers aged 40–69 years, who were recruited between 2006–2010 from England, Scotland and Wales. The UK Biobank has Research Tissue Bank approval (REC #16/NW/0274) that covers analysis of data by approved researchers. All participants provided written informed consent to the UK Biobank before data collection began. This research has been conducted using the UK Biobank resource under Application Number 47401. The eligibility criteria for this study was as previously described,[4] but in this study we included non-White British participants with ancestry defined using the UK Biobank's principal component scores analysis. PRS were centered by PCA-derived ancestry.[5]

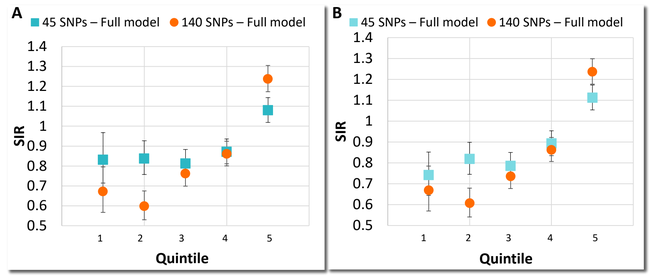

Figure 1. Standard incidence ratio (SIR) comparison between models based on 10-year risk scores. (A) The 140-SNP based model improved 10-year risk scores for adults compared to the 45-SNP model in the highest quintile and the lowest two quintiles. These participants were identified as UK genetic ancestry by PCA (n=440,518). (B) The second SIR shows all participants in the UKB (n=470,368), across all genetic ancestries.

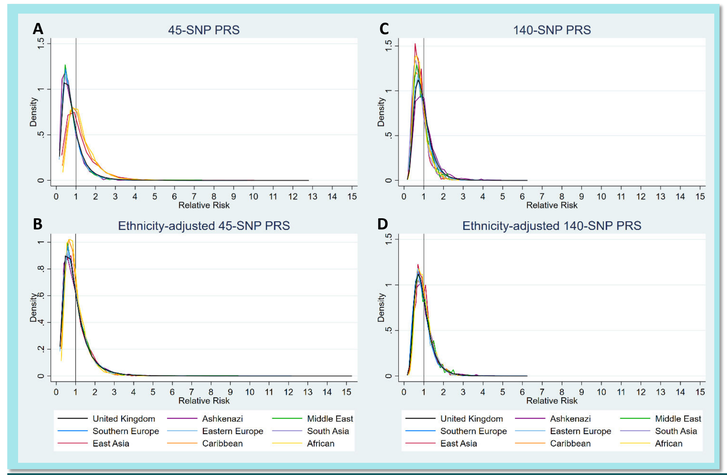

Figure 2. Distribution of PRS by UKB-PCA categories before and after centering. (A)45 SNP PRS before and (B) after PCA-based centering compared to the (C) 140 SNP PRS before and (D) after PCA-based centering.

Ancestry	Participants (n)
Ashkenazi	2468
Caribbean	2622
China	1836
India	6629
Iran	1222
Italy	6746
Nigeria	4052
Poland	4275
United Kingdom	440518
Total	470368

Table 1. Participant sample size by PCA.

Updated risk prediction score with a 140–SNP PRS and a simple measure of first-degree family history, age, sex and ethnicity, improves risk prediction and risk stratification in the general population compared with a similar model with a 45-SNP PRS.

Improved stratification observed over 10 years compared to previous model:
- Across major ethnicities and
- At the top classification of risk, where higher incidence cases observed
- At the bottom classification of risk, where lower number of incident cases observed

Risk stratification in healthy, asymptomatic adults will ultimately assist in colorectal cancer screening and risk-reduction efforts by aiding in compliance, risk-based-follow-up/prioritization and increasing patient self-awareness.

Contact: Erika.Spaeth@gtglabs.com

Results

The model using the 140-SNP PRS showed an improvement in discrimination, calibration and risk stratification over the model using the 45-SNP PRS for full-lifetime risk: discrimination was 0.706 (95% CI 0.697–0.715) and 0.674 (95% CI 0.664–0.683), respectively, and the P for difference was < 0.001. The 140-SNP model was well calibrated and showed a small overestimation of risk 0.951 (95% CI 0.918–0.986). Standardized incidence ratios (SIR) compared to population incidence rates showed that, for the 140-SNP model, the top quintile of risk shows a 27% improvement compared to the 45-SNP model. Importantly, we were able to show the improved performance of the 140-SNP PRS across multiple genetic ancestries. Using the UK Biobank, we were able to adjust the PRS to each of the PCA-defined genetic ancestry groups available to us in the UK Biobank. Finally, using 10-year risk thresholds of ≤1%, 1-4% and ≥4% to represent adults at average, moderate and increased risk of developing colorectal cancer (based on ≤1×, 1–3× and ≥3× relative risk, respectively), we showed the SIR based on the models ability to categorize participants into the appropriate risk levels. The 140-SNP model showed much higher SIR compared to the 45-SNP model for participants with 10-year risk scores ≥4%.

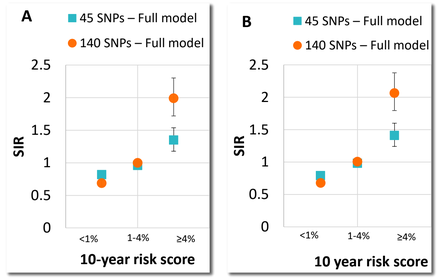

Figure 3. Standard incidence ratios by risk thresholds of ≤1%, 1-4% and ≥4% show improved classification of participants at increased risk of developing colorectal cancer (≥4%). These participants had a higher incidence rate of CRC over a 10-year period when classified by 140-SNP full model compared to the 45-SNP full model in (A) white-northern Europeans (UK-PCA) and in (B) all ancestries combined.

References

1. Jenkins MA, Win AK, Dowty JG, et al. Ability of known susceptibility SNPs to predict colorectal cancer risk for persons with and without a family history. Fam Cancer. 2019;18(4):389-397.
2. Thomas M, Sakoda LC, Hoffmeister M, Rosenthal EA, Lee JK, van Duijnhoven FJB, et al. Genome-wide Modeling of Polygenic Risk Score in Colorectal Cancer Risk. Am J Hum Genet. 2020;107(3):432-44.
3. Huyghe J.R., Bien S.A., Harrison T.A., Kang H.M., Chen S., Schmit S.L., Conti D.V., Qu C., Jeon J., Edlund C.K. Discovery of common and rare genetic risk variants for colorectal cancer. Nat. Genet. 2019;51:76–87
4. Gafni A, Dite GS, Spaeth Tuff E, Allman R, Hopper JL. Ability of known colorectal cancer susceptibility SNPs to predict colorectal cancer risk: A cohort study within the UK Biobank. PLoS One. 2021;16(9):e0251469.
5. Prive F, Aschard H, Carmi S, Folkersen L, Hoggart C, O'Reilly PF et al. (2022). Portability of 245 polygenic scores when derived from the UK Biobank and applied to 9 ancestry groups from the same cohort. Am J Hum Genet 109(1):12-23.