# Assessing agreement between different polygenic risk scores in the UK Biobank

**Lei Clifton[1*], Jennifer A Collister[1], Xiaonan Liu[1], Tom J Littlejohn[1], David J Hunter[1,2]**

1. Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom
2. Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, United States of America

* Corresponding Author
Email: lei.clifton@ndph.ox.ac.uk

## 1. Abstract

### 1.1 Background

Polygenic risk scores (PRS) are proposed to be used in clinical settings for risk stratification, with the public being accustomed to the concept of genetic predisposition to diseases. Much work has focused on developing PRSs for informing people about their risk of future health conditions; however, there are limited investigations on how different PRSs diverge from each other for risk prediction of individuals.

### 1.2 Methods and Findings

We compared recently published PRS for three conditions, breast cancer, hypertension and dementia, to assess the stability in practice of running these algorithms for risk prediction in a single large population. We used imputed genotyping data from the UK Biobank (UKB) prospective cohort, limited to the White British subset. We found that

1. Only 65%-79% of SNPs in the first PRS were represented in the more recent PRS for all three diseases, after having taken linkage disequilibrium (LD) into account ($R^2$>0.8).

2. Although the difference in the area under the received operator curve (AUC) obtained using the two PRS is hardly appreciable for all three diseases, there were large differences in individual risk prediction between the two PRS.

### 1.3 Conclusions

We found substantial discordance between different PRS for the same disease, indicating that individuals could receive different medical advice depending on which PRS is used to assess their genetic susceptibility for these disorders. It is desirable to resolving this uncertainty before using PRS for risk stratification in clinical settings.

## 2. Background

Genome-wide association studies (GWAS) have revealed that the inherited genetic component of most traits not due to mutations in a single gene is highly polygenic. Dozens or thousands of single nucleotide polymorphisms (SNPs) can be combined to produce a polygenic risk score (PRS) representing an individual's genetic propensity for a given trait or disease.

There is much enthusiasm for the use of PRS to inform individuals about their risk of future health conditions, either as stand-alone information, or combined with non-genetic data in integrated risk scores (Torkamani, Wineinger and Topol, 2018; Yanes *et al.*, 2020). PRS have been proposed in a wide variety of settings such as prioritizing people for disease screening, informing the prescription of preventive medicines, and even in embryo selection (Tellier *et al.*, 2021). Mutations in single genes associated with diseases are already utilised in a clinical setting, however, recent large-scale studies have found that PRS could potentially identify a greater proportion of at risk individuals (Khera *et al.*, 2018).

Early in the development of PRS, researchers (Kraft and Hunter, 2009; Chatterjee *et al.*, 2013) quantified the degree to which a PRS with a limited number of SNPs would misclassify people, when compared with future PRS with many additional variants. At the time it was thought that once dozens, or even hundreds of SNPs were included, diminishing returns would set in and the PRS would be relatively stable. This perception appeared to be supported by many large multi-study consortia, now identifying additional SNPs with very small odds ratios (OR; as low as 1.02 per allele) in which the area under the curve (AUC) for Receiver Operator Characteristic (ROC) curves of newer versions of the PRS are only minimally higher than that of previous versions.

Recent discussion on the use of PRS in the clinic has largely focused on the reporting standards for the derivation and archiving of PRS (Ding *et al.*, 2021; Wand *et al.*, 2021), the health economic value of PRS, the potential contribution of PRS to health disparities given the limited databases available for non-European ancestry populations (Lewis and Vassos, 2020a), the most appropriate way for benefiting the patients (Knowles and Ashley, 2018; Lewis and Vassos, 2020b), and the means of communicating PRS to patients or members of the public.

During the construction of PRS, there are multiple design options for deciding the number of SNPs to include and for assigning an appropriate weight to each SNP. Consequently, multiple sets of SNPs exist, resulting in multiple PRS for the same trait. For example, the 313-SNP PRS for breast cancer (Mavaddat *et al.*, 2019) was developed using hard-thresholding stepwise forward regression, whereas the 118k-SNP PRS (Fritsche *et al.*, 2020) was selected by the penalised regression "lassosum" and the highest pseudo-$R^2$. However, these PRS are typically compared at a population level using metrics such as the AUC or OR, and limited attention has been paid to how they differ from each other for risk prediction of individuals.

## 3. Methods

### 3.1 Study populations

We used the genetic data from the UK Biobank (UKB), a large-scale population-based prospective cohort study of approximately 500,000 individuals aged 40-69 years at recruitment across the United Kingdom between March 2006 and October 2010. Full details of the genotyping and imputation are described elsewhere (Bycroft *et al.*, 2018).

Our study populations for each of the three disease outcomes are defined as follows:

- The breast cancer eligible population was women who had not had breast cancer, carcinoma in situ or mastectomy prior to baseline.
- For the hypertension eligible population, we excluded individuals with missing or implausible systolic blood pressure (SBP) measurements (< 70 or > 270 mmHg) at baseline, and those with prior Major Adverse Cardiovascular Events (MACE).
- The dementia eligible population was restricted to individuals without prior diagnosis of dementia or Alzheimer's disease.

Disease ascertainment of UKB during the following-up period utilised linkage to death registry, cancer registry, and Hospital Episode Statistics (HES). Hypertension is defined as SBP>=140 at baseline; the International Classification of Diseases (ICD) code for breast cancer and dementia can be found in Supplementary Tables 13-15.

### 3.2 Calculating PRS

We computed PRS of an individual $j$ by the weighted sum of trait-associated SNPs,

$$PRS_j = \sum_{i}^{N} \beta_i * dosage_{ij}$$

where N is the total number of SNPs, $\beta_i$ is the effect size (or beta) of SNP $i$, and $dosage_{ij}$ is the number of effect alleles (usually encoded as 0, 1 or 2 in SNP $i$ for individual $j$ for the effect allele).

We used published effect size of SNPs, and applied genetic quality control (QC) pipelines for both SNPs and samples. During SNP QC, we removed ambiguous SNPs (A/T or C/G SNPs with MAF > 0.49) and rare variants with MAF < 0.005; we only retained SNPs with high imputation quality (imputation information score > 0.4) (Supplementary Table 1). During sample QC, we excluded participants who were sex-discordant, outliers for missingness or heterozygosity, or related at 3$^{rd}$ degree or higher, using UKB Data Field 22020.

For each of the three disease outcomes, we selected a pair of recently published PRS within two years of each other for each disease, where the sample sizes for PRS derivation are large. Our PRS for breast cancer and dementia are from the polygenic score (PGS) Catalogue (www.PGSCatalog.org) (Lambert *et al.*, 2021), while PRS for hypertension are from the literature. The earlier PRS is denoted as PRS-A (typically contains fewer SNPs), while the more recent one is PRS-B with more SNPs. For each pair of PRS, we compared the total number of SNPs and the overlap between scores by checking the number of SNPs in common between the scores or in high LD ($R^2$>0.8).

We then computed PRS for those within the defined study population of each condition, restricting to genetically White British individuals using UKB Data Field 22006. This yields the final size N of each study population.

- For breast cancer, our "baseline" PRS-A (313 SNPs, PGS ID: PGS000004) (Mavaddat *et al.*, 2019) has been widely validated and is included in the current implementation of the BOADICEA breast cancer risk model (Lee *et al.*, 2019; Lakeman *et al.*, 2020). For the "comparison" PRS-B, we used a score containing (118,388 SNPs, PGS ID: PGS000511) (Fritsche *et al.*, 2020) which was largely developed from the same Breast Cancer Association Consortium (BCAC) GWAS data as PRS-A (Michailidou *et al.*, 2017).

- For hypertension, we selected PRS for SBP from the literature. SBP PRS-A contains 267 SNPs (Warren *et al.*, 2017), which we compared to a later SBP PRS-B containing 884 SNPs (Evangelou *et al.*, 2018) with effect sizes from the International Consortium of Blood Pressure-Genome Wide Association Studies (ICBP), Million Veteran Program (MVP) and Estonian Genomic Centre of the University of Tartu (EGCUT).

- PRS-A for dementia originally contained 22 SNPs (PGS ID: PGS000334) (Zhang *et al.*, 2020); we subsequently removed the two APOE SNPs (rs429358 and rs7412) to avoid the APOE genotype dominating the PRS, leaving us with 20 SNPs. Our 39-SNP PRS-B (Ebenau *et al.*, 2021) (PGS ID: PGS001775) used effect sizes from the International Genomics of Alzheimer's Project (IGAP) GWAS (Kunkle *et al.*, 2019); we retained all the 39 SNPs since they do not include the two APOE SNPs.

### 3.3 Quantifying the stability of PRS

In each disease-specific study population, we calculated the correlation coefficient between each pair of PRS and the age- and sex-adjusted odds ratios (ORs) between various cut-points compared with the middle quintile of the PRS distribution. We then used predictions from a multivariable logistic regression model containing age, sex, the continuous PRS, genetic array, and first 5 PCs to compute the area under curve (AUC) for each PRS (Table 1). The continuous Net Reclassification Index (NRI) was used to compare PRS-A with PRS-B in multivariable logistic models (Table 1), whereas the categorical NRI was used in cross-classification of PRS percentile risk categories (Percentage reclassification for participants who experienced the outcome are shown in Supplementary Tables 4-6 and 10-12, for top 1% and top 5% risk categorisations, respectively).

## 4. Results

Our study populations were N=171,490 (cases=6,347) for breast cancer, N=317,581 (cases=137,649) for hypertension, and N=335,689 (cases=4,460) for dementia. For comparing different PRS, we focused on two aspects: firstly, the consistency of the selected SNPs and performance metrics. Then we assessed the correlation between each pair of PRS, and the extent to which PRS-B gave the same predictions for individuals as PRS-A.

We found that only 65%-79% of SNPs in PRS-A were represented in PRS-B for all three diseases, after having taken linkage disequilibrium (LD) into account ($R^2$>0.8). This is somewhat surprising, as one might expect a newer score (PRS-B) to incorporate most of the previously identified SNPs from PRS-A.

Table 1 presents the performance characteristics of each PRS against the corresponding disease outcome in UKB. In each case the more recent PRS-B was associated with a slightly higher OR than the earlier PRS-A. For example, the OR of breast cancer among women in the top 1% compared to those in the middle quintile was 3.41 for PRS-A and 3.94 for PRS-B. Their corresponding AUCs were only minimally different (0.638 vs 0.641), and the ROCs looked almost identical (**Error! Reference source not found.**Figure 1). Similar results were obtained for hypertension (PRS-A OR=1.83, AUC=0.69; PRS-B OR=2.18, AUC=0.70) and dementia (PRS-A OR=1.78, AUC=0.80; PRS-B OR=2.30, AUC=0.80).

Despite similar AUCs, PRS-A and PRS-B were not highly correlated for any outcome, with their Pearson correlation coefficient $r$ only in the range of 0.65 to 0.73. Compatible with these correlation coefficients, there was substantial reclassification of predicted risk according to percentiles of PRS-A and PRS-B for all three diseases, as shown for breast cancer in Table 2. For women in the top 1% of breast cancer risk by PRS-A, only 23.1% were in the top 1% risk of PRS-B. The equivalent percentage was 22.9% and 22.7% for hypertension and dementia, respectively (Supplementary Tables 2-3). We focused on the top 1% of risk because of the widely-promulgated concept that these risks approximate those of the risks for monogenic traits (Khera *et al.*, 2018).

Using a more relaxed risk category, participants in the top 5% of risk for breast cancer by PRS-A, only 35.7% were in the top 5% risk by PRS-B. The equivalent percentage was 35.8% and 40.0% for hypertension and dementia, respectively (Supplementary Tables 7-9).

*Figure 1: ROC plots obtained from predictions from multivariable logistic regression of age, sex, continuous PRS, genotyping array and first 5 PCs against disease outcome*

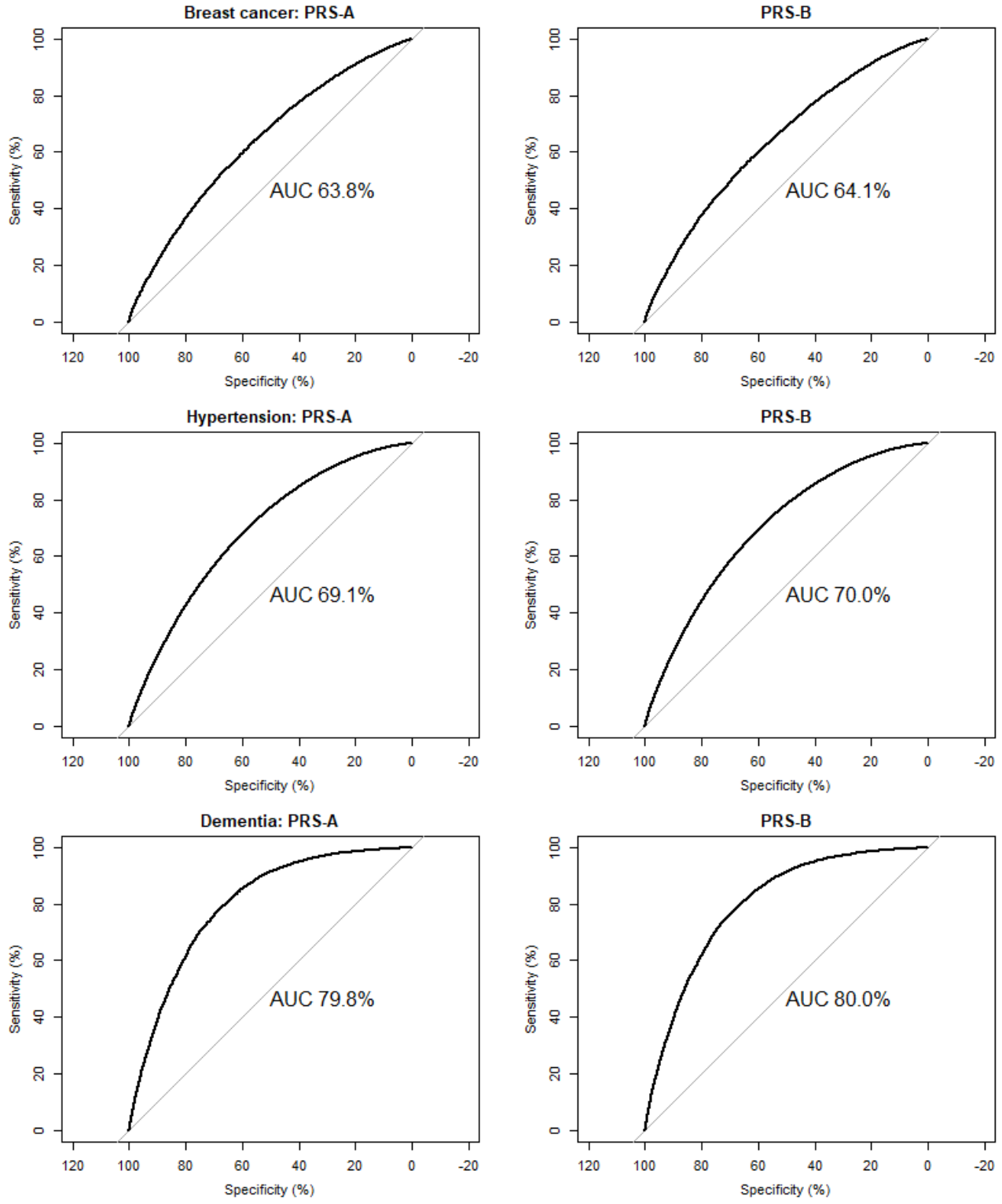

*Table 1: PRS compared for each outcome and their performance characteristics in the UK Biobank. N: number of participants whose PRS score was obtained. nSNPs: number of SNPs in PRS prior to genetic quality control. OR: odds ratio for top 1% vs middle quintile of PRS from multivariable logistic regression model adjusted for age, sex, genotyping array and first 5 PCs. AUC: area under receiver-operating curve using predicted risk from multivariable logistic regression model containing age, sex, continuous PRS, genotyping array and first 5 PCs. NRI: continuous net reclassification index using predicted risks from two multivariable logistic regression models containing age, sex, continuous PRS for this disease, genotyping array and first 5 PCs. The model containing PRS-B is considered the "updated" model. $r$: Pearson correlation coefficient between the two continuous PRS for this disease. LD: number (%) of SNPs in PRS-A which either appear in or are in linkage disequilbrium ($R^2$ > 0.8) with SNPs in PRS-B. Breast cancer models are not adjusted for sex because its population is restricted to females.*

| Disease (N) | PRS | nSNPs | OR (95% CI) | AUC (95% CI) | NRI (95% CI) | $r$ | LD: $R^2$> 0.8 (% overlap) |
|---|---|---|---|---|---|---|---|
| Breast cancer (171,490) | A | 313 | 3.41 (2.89, 4.03) | 0.64 (0.63, 0.64) | 0.03 (0.00, 0.05) | 0.65 | 225 (72%) |
| | B | 118388 | 3.94 (3.36, 4.61) | 0.64 (0.63, 0.65) | | | |
| Hypertension (317,581) | A | 267 | 1.83 (1.70, 1.98) | 0.69 (0.69, 0.69) | 0.16 (0.16, 0.17) | 0.66 | 210 (79%) |
| | B | 884 | 2.18 (2.02, 2.35) | 0.70 (0.70, 0.70) | | | |
| Dementia (335,689) | A | 20 | 1.78 (1.40, 2.26) | 0.80 (0.79, 0.80) | 0.09 (0.06, 0.12) | 0.73 | 13 (65%) |
| | B | 39 | 2.30 (1.85, 2.87) | 0.80 (0.79, 0.81) | | | |

*Table 2: Cross-classification of predicted risk of breast cancer, according to the percentiles of each PRS. Number of participants are shown as n (col%, row%, cell%). Higher percentiles of PRS indicate increased risk of breast cancer; "≥ 99%" percentile corresponds to the top 1% risk.*

| Percentiles of PRS-A | Percentiles of PRS-B | | | | | | |
|---|---|---|---|---|---|---|---|
| | < 1% | 1-20% | 20-40% | 40-60% | 60-80% | 80-99% | ≥ 99% |
| < 1% | 345 (20.1, 20.1, 0.2) | 1140 (66.5, 3.5, 0.7) | 176 (10.3, 0.5, 0.1) | 44 (2.6, 0.1, 0.0) | 9 (0.5, 0.0, 0.0) | 1 (0.1, 0.0, 0.0) | 0 (0.0, 0.0, 0.0) |
| 1-20% | 1117 (3.4, 65.1, 0.7) | 15409 (47.3, 47.3, 9.0) | 8818 (27.1, 25.7, 5.1) | 4696 (14.4, 13.7, 2.7) | 2052 (6.3, 6.0, 1.2) | 490 (1.5, 1.5, 0.3) | 1 (0.0, 0.1, 0.0) |
| 20-40% | 198 (0.6, 11.5, 0.1) | 8788 (25.6, 27.0, 5.1) | 9989 (29.1, 29.1, 5.8) | 8053 (23.5, 23.5, 4.7) | 5210 (15.2, 15.2, 3.0) | 2051 (6.0, 6.3, 1.2) | 9 (0.0, 0.5, 0.0) |
| 40-60% | 43 (0.1, 2.5, 0.0) | 4648 (13.6, 14.3, 2.7) | 7998 (23.3, 23.3, 4.7) | 8908 (26.0, 26.0, 5.2) | 8050 (23.5, 23.5, 4.7) | 4607 (13.4, 14.1, 2.7) | 44 (0.1, 2.6, 0.0) |
| 60-80% | 11 (0.0, 0.6, 0.0) | 2098 (6.1, 6.4, 1.2) | 5296 (15.4, 15.4, 3.1) | 7988 (23.3, 23.3, 4.7) | 10047 (29.3, 29.3, 5.9) | 8688 (25.3, 26.7, 5.1) | 170 (0.5, 9.9, 0.1) |
| 80-99% | 1 (0.0, 0.1, 0.0) | 497 (1.5, 1.5, 0.3) | 2006 (6.2, 5.8, 1.2) | 4574 (14.0, 13.3, 2.7) | 8736 (26.8, 25.5, 5.1) | 15674 (48.1, 48.1, 9.1) | 1095 (3.4, 63.8, 0.6) |
| ≥ 99% | 0 (0.0, 0.0, 0.0) | 3 (0.2, 0.0, 0.0) | 15 (0.9, 0.0, 0.0) | 35 (2.0, 0.1, 0.0) | 194 (11.3, 0.6, 0.1) | 1072 (62.5, 3.3, 0.6) | 396 (23.1, 23.1, 0.2) |

## 5. Discussion

The clinical utility of PRS depends on the clinical validity of the predictions. Clinical validity is not only dependent on the information PRS give on the risk of future events, but also on the stability of these estimates. Here we show that for three common conditions (breast cancer, hypertension and dementia), the risk estimates derived from different PRS would result in very different information being given to a high proportion of people.

We found that the more recent PRS-B had minimal increase in AUC compared to the older PRS-A, in line with the small improvement measured by net reclassification index (NRI). However, the PRS differed substantially in how they assigned participants into risk categories with a substantial proportion of individuals classified at high risk by one PRS, not so classified by the other PRS. This suggests a major potential problem for the use of these PRS in clinical practice, given the changes in clinical

recommendations associated with labelling a person in the same category of risk as a monogenic disorder. Our results demonstrated large differences across all percentiles of risk; although the clinical consequences at the lower percentiles may not be as extreme as at the higher percentiles, the clinical utility will still be reduced by incorrect classification.

While UKB data was used in the development of some of these PRS, since our interest is in comparing the classification of individuals by each score and not in developing a prediction model, we expect this to have little impact on the results.

We note that we have not established the reasons for the extent of misclassification between different PRS. It does not appear to be attributable solely to the number of SNPs included in the PRS – we show this for PRS comprised of over 100 thousand versus several hundred SNPs (breast cancer), for PRS composed of hundreds of SNPs (systolic blood pressure), and for PRS composed of <50 SNPs (dementia). The surprisingly small number of SNPs held in common by different PRS for the same condition published only a year apart indicates that different analytical methods used to derive the PRS may account for some of the discrepancies in classification, but understanding this phenomenon is clearly important for PRS selection in broad clinical practice. The correlations we observed between PRS for the same condition are in the same range as is seen for variation in risk predictors measured several years apart such as blood pressure and serum cholesterol, far from the "fixed" or "one-time" value at birth that is often assumed for PRS.

We are not the first to notice this phenomenon. For instance, Läll (Läll *et al.*, 2019) compared the performance of four PRS in breast cancer prediction, noted that some of the correlations between them were as low as $r$=0.3, and observed that a "metaGRS" of the PRS performed better than any of the individual PRS. However, the issue does not seem to be widely appreciated, and most publications comparing a new PRS with previous versions assert the superiority of the new PRS and do not address the issue of misclassification of risk between PRS. Our observation shows two PRS that only minimally differ in predictive performance on a population level may substantially differ in terms of individual risk classification, even among individuals with the same continental ancestry. This issue requires careful consideration before utilising PRS in real-world settings, because such an arbitrary element in health care is obviously undesirable. A person's genetic profile is generally considered fixed at birth, leading to the widely held conviction that genetic susceptibility is an immutable value; however, our findings show that such assertions may be premature in the context of PRS which are still a developing area of research. While it is reasonable to expect incremental improvements in any risk prediction algorithm over time, these results suggest there is still considerable uncertainty associated with estimates of risk derived from different PRS for the same disease. It will be important to develop guidelines on best practice in constructing PRS to minimize the extent to which people will be given contradictory information over short periods of time.

## 6. Declarations

### 6.1 Ethics approval and consent to participate

The UK Biobank study received ethical approval from the North West Multi-center Research Ethics Committee (REC reference: 11/NW/03820). All participants gave written informed consent before

enrolment in the study, which was conducted in accordance with the principles of the Declaration of Helsinki. This study has been conducted under the UK Biobank application ID 33952.

### 6.2 Patient and community involvement

The analyses presented here are based on existing data from the UK Biobank cohort study, and the authors were not involved in participant recruitment. To the best of our knowledge, no patients were explicitly engaged in the design or implementation of the UK Biobank study. No patients were asked to advise on interpretation or writing these results. Results from UK Biobank are routinely disseminated to study participants via the study website and social media outlets.

### 6.3 Consent for publication

Yes.

### 6.4 Availability of data and material

The data reported in this paper are available via application directly to the UK Biobank, *https://www.ukbiobank.ac.uk*

### 6.5 Competing interests

All authors declare no support from any organization for the submitted work, no financial relationship with any organization that might have an interest in the submitted work in the previous three years; no other relationship or activities that could appear to have influenced the submitted work.

### 6.6 Funding

The UK Biobank study was supported by the Wellcome Trust, Medical Research Council, Department of Health, Scottish government, and Northwest Regional Development Agency. It has also received funding from the Welsh Assembly government and British Heart Foundation. The analyses here were funded by the Cancer Research UK (grant no C16077/A29186), and supported by the Nuffield Department of Population Health, Oxford University.

### 6.7 Author contributions

LC and DH outlined the statistical methods and drafted the manuscripts; JC conducted the statistical analysis; XL reviewed the R scripts by JC. All authors have revised the manuscript and agreed on its contents.

### 6.8 Transparency statement

The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

### 6.9 Acknowledgements

We thank the participants of the UK Biobank study for enabling us to conduct this research.

## 7. References

Bycroft, C. *et al.* (2018) 'The UK Biobank resource with deep phenotyping and genomic data', *Nature*, 562(7726), pp. 203–209. doi: 10.1038/s41586-018-0579-z.

Chatterjee, N. *et al.* (2013) 'Projecting the performance of risk prediction based on polygenic analyses of genome-wide association studies', *Nature Genetics*. Nat Genet, 45(4), pp. 400–405. doi: 10.1038/ng.2579.

Ding, Y. *et al.* (2021) 'Large uncertainty in individual polygenic risk score estimation impacts PRS-based risk stratification', *Nature Genetics*. Nature Publishing Group, 54(1), pp. 30–39. doi: 10.1038/s41588-021-00961-5.

Ebenau, J. L. *et al.* (2021) 'Risk of dementia in APOE ε4 carriers is mitigated by a polygenic risk score', *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*. Wiley-Blackwell, 13(1). doi: 10.1002/dad2.12229.

Evangelou, E. *et al.* (2018) 'Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits', *Nature Genetics*. Nature Publishing Group, 50(10), pp. 1412–1425. doi: 10.1038/s41588-018-0205-x.

Fritsche, L. G. *et al.* (2020) 'Cancer PRSweb: An Online Repository with Polygenic Risk Scores for Major Cancer Traits and Their Evaluation in Two Independent Biobanks', *The American Journal of Human Genetics*. Cell Press, 107(5), pp. 815–836. doi: 10.1016/J.AJHG.2020.08.025.

Khera, A. V. *et al.* (2018) 'Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations', *Nature Genetics*. Nature Publishing Group, pp. 1219–1224. doi: 10.1038/s41588-018-0183-z.

Knowles, J. W. and Ashley, E. A. (2018) 'Cardiovascular disease: The rise of the genetic risk score', *PLoS*

*Medicine*. Public Library of Science, 15(3), p. e1002546. doi: 10.1371/journal.pmed.1002546.

Kraft, P. and Hunter, D. J. (2009) 'Genetic Risk Prediction — Are We There Yet?', *New England Journal of Medicine*. N Engl J Med, 360(17), pp. 1701–1703. doi: 10.1056/nejmp0810107.

Kunkle, B. W. *et al.* (2019) 'Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing', *Nature Genetics 2019 51:3*. Nature Publishing Group, 51(3), pp. 414–430. doi: 10.1038/s41588-019-0358-2.

Lakeman, I. M. M. *et al.* (2020) 'Validation of the BOADICEA model and a 313-variant polygenic risk score for breast cancer risk prediction in a Dutch prospective cohort', *Genetics in Medicine 2020 22:11*. Nature Publishing Group, 22(11), pp. 1803–1811. doi: 10.1038/s41436-020-0884-4.

Läll, K. *et al.* (2019) 'Polygenic prediction of breast cancer: Comparison of genetic predictors and implications for risk stratification', *BMC Cancer*. BMC Cancer, 19(1). doi: 10.1186/s12885-019-5783-1.

Lambert, S. A. *et al.* (2021) 'The Polygenic Score Catalog as an open database for reproducibility and systematic evaluation', *Nature Genetics*. Nature Publishing Group, pp. 420–425. doi: 10.1038/s41588-021-00783-5.

Lee, A. *et al.* (2019) 'BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors', *GENETICS in MEDICINE*. Springer US, 21(8). doi: 10.1038/s41436-018-0406-9.

Lewis, C. M. and Vassos, E. (2020a) 'Polygenic risk scores: From research tools to clinical instruments', *Genome Medicine*, 12(1), pp. 1–11. doi: 10.1186/s13073-020-00742-5.

Lewis, C. M. and Vassos, E. (2020b) 'Polygenic risk scores: From research tools to clinical instruments', *Genome Medicine*. BioMed Central. doi: 10.1186/s13073-020-00742-5.

Mavaddat, N. *et al.* (2019) 'Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes', *American Journal of Human Genetics*. Cell Press, 104(1), pp. 21–34. doi: 10.1016/j.ajhg.2018.11.002.

Michailidou, K. *et al.* (2017) 'Association analysis identifies 65 new breast cancer risk loci', *Nature 2017 551:7678*. Nature Publishing Group, 551(7678), pp. 92–94. doi: 10.1038/nature24284.

Tellier, L. C. A. M. *et al.* (2021) 'Embryo screening for polygenic disease risk: Recent advances and ethical considerations', *Genes*. MDPI AG, 12(8). doi: 10.3390/genes12081105.

Torkamani, A., Wineinger, N. E. and Topol, E. J. (2018) 'The personal and clinical utility of polygenic risk scores', *Nature Reviews Genetics*. Nat Rev Genet, pp. 581–590. doi: 10.1038/s41576-018-0018-x.

Wand, H. *et al.* (2021) 'Improving reporting standards for polygenic scores in risk prediction studies', *Nature*. Springer US, 591(April 2020). doi: 10.1038/s41586-021-03243-6.

Warren, H. R. *et al.* (2017) 'Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk', *Nature Genetics 2017 49:3*. Nature Publishing Group, 49(3), pp. 403–415. doi: 10.1038/ng.3768.

Yanes, T. *et al.* (2020) 'The emerging field of polygenic risk scores and perspective for use in clinical care', *Human Molecular Genetics*. Hum Mol Genet, pp. R165–R176. doi: 10.1093/HMG/DDAA136.

Zhang, Q. *et al.* (2020) 'Risk prediction of late-onset Alzheimer's disease implies an oligogenic architecture', *Nature Communications 2020 11:1*. Nature Publishing Group, 11(1), pp. 1–11. doi: 10.1038/s41467-020-18534-1.