

# CanRisk-Prostate: a comprehensive, externally validated risk model for the prediction of future prostate cancer

Tommy Nyberg <sup>a,b\*</sup> <https://orcid.org/0000-0002-9436-0626>

Mark N. Brook <sup>c</sup> <https://orcid.org/0000-0002-8969-2378>

Lorenzo Ficorella <sup>a</sup>

Andrew Lee <sup>a</sup> <https://orcid.org/0000-0003-0677-0252>

Joe Dennis <sup>a</sup> <https://orcid.org/0000-0003-4591-1214>

Xin Yang <sup>a</sup> <https://orcid.org/0000-0003-0037-3790>

Naomi Wilcox <sup>a</sup> <https://orcid.org/0000-0002-7709-7029>

Tokhir Dadaev <sup>c</sup> <https://orcid.org/0000-0002-8268-0438>

Koveela Govindasami <sup>c</sup> <https://orcid.org/0000-0003-3047-7574>

Michael Lush <sup>a</sup>

Goska Leslie <sup>a</sup> <https://orcid.org/0000-0001-5756-6222>

Artitaya Lophatananon <sup>d</sup> <http://orcid.org/0000-0003-0550-4657>

Kenneth Muir <sup>d</sup> <https://orcid.org/0000-0001-6429-988X>

UK Genetic Prostate Cancer Study Collaborators

Elizabeth Bancroft <sup>c,e</sup> <https://orcid.org/0000-0002-5482-5660>

Douglas F. Easton <sup>a</sup> <https://orcid.org/0000-0003-2444-3247>

Marc Tischkowitz <sup>f</sup> <https://orcid.org/0000-0002-7880-0628>

Zsofia Kote-Jarai <sup>c†</sup> <https://orcid.org/0000-0002-6329-5155>

Rosalind Eeles <sup>c,e†</sup> <https://orcid.org/0000-0002-3698-6241>

Antonis C. Antoniou <sup>a†</sup> <https://orcid.org/0000-0001-9223-3116>

<sup>a</sup> Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.

<sup>b</sup> MRC Biostatistics Unit, University of Cambridge, Cambridge, UK.

<sup>c</sup> Oncogenetics Team, Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK.

<sup>d</sup> Division of Population Health, Health Services Research and Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK.

<sup>e</sup> Cancer Genetics Unit, The Royal Marsden NHS Foundation Trust, London, UK.

<sup>f</sup> Department of Medical Genetics, University of Cambridge, Cambridge, UK.

\* Corresponding author.

Address: Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Cambridge, CB1 8RN, Cambridge, UK.

Email: [ten25@cam.ac.uk](mailto:ten25@cam.ac.uk)

Twitter: @DrTommyNyberg

† Joint senior authors.

**Keywords:** external validation; genetic risk; kin-cohort; prospective cohort; prostate cancer; risk model; UK Biobank; UK Genetic Prostate Cancer Study

**Running head:** The CanRisk-Prostate validated genetic risk model.

**Word count (abstract):** 253 words.

**Word count (main text):** 3023 words.

## Abstract

**Purpose:** Prostate cancer (PCa) is highly heritable. No validated PCa risk model currently exists. We therefore sought to develop a genetic risk model that can provide personalised predicted PCa risks based on known moderate-to-high-risk pathogenic variants, low-risk common genetic variants and explicit cancer family history, and to externally validate the model in an independent prospective cohort.

**Patients and methods:** We developed a risk model using a kin-cohort comprising individuals from 16,633 PCa families ascertained in the UK in 1993-2017 from the UK Genetic Prostate Cancer Study, and complex segregation analysis adjusting for ascertainment. The model was externally validated in 170,850 unaffected men (7,624 incident PCas) recruited in 2006-2010 to the independent UK Biobank prospective cohort study.

**Results:** The most parsimonious model included the effects of pathogenic variants in *BRCA2*, *HOXB13* and *BRCA1* and a polygenic score based on 268 common low-risk variants. Residual familial risk was modelled by a hypothetical recessively inherited variant and a polygenic component whose standard deviation decreased log-linearly with age. The model predicted familial risks that were consistent with those reported in previous observational studies. In the validation cohort, the model discriminated well between unaffected men and men with incident PCas within 5yr (C-index=0.790, 95% CI 0.783-0.797) and 10yr (C-index=0.772, 95% CI 0.768-0.777). The 50% of men with highest predicted risks captured 86.3% of PCa cases within 10yr.

**Conclusion:** This is the first validated risk model offering personalised PCa risks. The model will assist in counselling men concerned about their risk and can facilitate future risk-stratified population screening approaches.

## Context summary

**Key objective:** Can a genetic risk model that uses information on all known high-, moderate- and low-risk prostate cancer genetic susceptibility variants, together with residual cancer family history information, accurately predict men's risk of developing prostate cancer in the future?

**Knowledge generated:** We developed a genetic risk model using data from 16,633 prostate cancer families. The model uses data on rare pathogenic variants in the moderate-to-high-risk genes *BRCA2*, *HOXB13* and *BRCA1*, a polygenic score based on 268 common low-risk variants, and detailed cancer family history, to predict the future risks. The risk model predicted incident prostate cancers in an independent cohort of 170,850 prospectively followed men with high discrimination and good calibration. The majority, 86%, of incident prostate cancers occurred among the half of men with the highest predicted risks.

## Introduction

Prostate cancer (PCa) exhibits marked familial aggregation and has one of the highest heritabilities of any common cancer.<sup>1–4</sup> This is explained in part by rare pathogenic variants (PVs) in *BRCA2*, *HOXB13* and possibly *BRCA1*, that are associated with moderate-to-high PCa risks,<sup>5–14</sup> together with several hundred commoner variants conferring lower risks, identified through genome-wide association studies.<sup>15–18</sup>

Men currently seen in family or genetics clinics are counselled based on descriptive family history (FH) and ethnicity-specific risk estimates<sup>19,20</sup> and/or average PV risk estimates.<sup>20–22</sup> However, risks for *BRCA1/2* and *HOXB13* PV carriers have been found to vary by PCa FH.<sup>9,10</sup> In addition, polygenic scores (PGS) based on common variants can provide considerable risk stratification,<sup>18,23–25</sup> in the general population and in men with FH,<sup>23</sup> *BRCA1/2*<sup>26–28</sup> or *HOXB13* PVs.<sup>13,27</sup> A comprehensive risk model, incorporating the joint effects of known and unknown genetic factors, should therefore provide better risk stratification and hence a more rational basis for counselling. Such models are now in widespread use in the management of breast and ovarian cancer risk.<sup>29–33</sup> A PCa model would address similar clinical needs. Some genetic PCa risk models exist,<sup>34–41</sup> but none combine data on detailed FH, PVs and the latest PGS. None have been externally validated.

To support improved and consistent counselling of at-risk men based on personalised future PCa risks, and to enable risk-stratified interventions, we developed a risk model based on data from a large kin-cohort study; and validated the model in an independent prospective cohort.

## Material and Methods

### Study participants: UKGPCS

The UK Genetic Prostate Cancer Study (UKGPCS; <http://www.icr.ac.uk/ukgpcs/>) recruited individuals with histologically confirmed PCa in three arms: a population-based arm that recruited men independent of age or FH, and arms enriched for young-age-at-onset PCa or PCa FH. Self-reported cancer FH data were collected through a questionnaire. We used data on the families of 16,633 European ancestry probands recruited in 1993-2017. Subsets had data available on *HOXB13* G84E (n=11,500),<sup>10,13,42</sup> *BRCA1* (n=2,148), *BRCA2* PVs (n=3,077),<sup>43,44</sup> and a 268-SNP PGS (n=11,149; Supplementary material).<sup>18</sup>

### Population controls

To estimate the PGS population-distribution, we included 4,319 controls genotyped using the same SNP array as the cases, from: (1) men without PCa personal or FH recruited through UKGPCS participating clinics, and (2) ProtecT trial participants with PSA<0.5 ng/ml.<sup>44,45</sup>

### Study participants: UK Biobank

The model was externally validated in UK Biobank (<https://www.ukbiobank.ac.uk/>), a prospective cohort study of volunteers recruited in 2006-2010. Data were available on 170,850 white British male participants without any cancer at recruitment (except non-melanoma skin cancers).

Participants provided baseline cancer FH information, and were followed-up prospectively through linkage with national registries. Data were available on a modified 268-SNP PGS and on the *HOXB13* G84E variant for all participants,<sup>14,46</sup> and on *BRCA1/2* protein-truncating variants for 40% of the participants (Supplementary material).<sup>47,48</sup>

## Descriptive familial relative risks (FRR)

To explore familial aggregation patterns in UKGPCS families, we estimated FRRs to relatives of the probands (Supplementary material).

## Risk model development

We used complex segregation analysis to fit genetic models for the observed cancer inheritance patterns in UKGPCS families.<sup>49</sup> PCa incidence was assumed to depend on *BRCA2*, *HOXB13* and *BRCA1* PVs, together with a polygenic component (PGC) to model residual familial risk. The PGC was assumed normally distributed, reflecting the combined effects of a large number of low-risk alleles. Additional models were considered which allowed for a fourth hypothetical major gene; following recessive, dominant or multiplicative models of Mendelian inheritance. The average age-specific incidences across all genotypes and polygenotypes were constrained to agree with calendar-period- and birth-cohort-specific population incidences.<sup>29,30,50,51</sup> Female relatives were assumed to be at risk of breast and ovarian cancer, following a similar model but without PGC. The models were parametrised by logit-transformed allele frequencies and log-relative risks (RRs) for genetic components; the log-SD of PGC which was assumed constant or age-dependent; and the logit-transformed proportion of the PGC that was explained by the PGS. Parameters were estimated by maximising the joint likelihoods of the family members' phenotypes under the assumed genetic model, using MENDEL software (version 3.3).<sup>52</sup> We adjusted for the non-random ascertainment of families by conditioning on data that may have influenced the ascertainment.<sup>53</sup> The fit of different models were compared using the Akaike information criterion and likelihood ratio tests (Supplementary material).

### Known genetic components

For *BRCA1/2*, given the small number of carriers in UKGPCS, we assumed external estimates of age-specific RRs of PCa,<sup>5</sup> breast and ovarian cancer and allele frequencies.<sup>29–31,33</sup> *HOXB13* G84E frequencies and RRs were estimated here: guided by a previous study, we assumed a multiplicative per-allele effect, with birth-cohort-specific RRs (born<1930/≥1930).<sup>10</sup>

We used the best-fitting model to include a PCa PGS based on 268 SNPs

(<https://www.pgscatalog.org/score/PGS000662/>).<sup>18</sup> We decomposed the PGC into one part explained by the PGS and an independent residual part due to unidentified genetic effects,<sup>31,33</sup> and estimated the fraction of the PGC explained by the PGS as a model parameter.

Guided by observations that FH is associated with higher PCa risk also for PV carriers,<sup>9,10</sup> and that PGSs modify the risk for PV carriers,<sup>13,26–28</sup> we assumed that the joint effects of PGC, PGS and PVs on PCa risk are multiplicative.

### Sensitivity analyses

We assessed the effect of the ascertainment adjustment based on method of PCa diagnosis (symptomatic, PSA testing, or unknown), and refitted the model in subgroups (supplementary material).

### Model-predicted risks

We compared age-specific model-predicted FRRs with FRRs reported in observational studies.<sup>1</sup> The model was used to estimate absolute PCa risks in example scenarios (supplementary material).

## External validation

We predicted 5 and 10yr prospective risks of developing PCa for the UK Biobank participants, using the data on age and FH available at baseline, PVs and PGS. Only *BRCA1/2* protein-truncating variants were available, and hence *BRCA1/2* PVs did not include pathogenic missense variants or large rearrangements; we therefore assumed testing sensitivities of 83% for *BRCA2* and 65% for *BRCA1*. We compared the predicted and observed risks of PCa diagnosis, and assessed the model discriminatory ability and calibration (supplementary material). We also assessed the model sensitivity and specificity at different quantiles of the risk distribution.

## Ethics

All participants provided written informed consent. UKGPCS was approved by the London Central Research Ethics Committee. UK Biobank was approved by the North West Multi-Centre Research Ethics Committee.

## Results

**Figure S1** details the inclusion and **Table S1** the characteristics of the UKGPCS probands and their relatives. Thirty percent reported at least one PCa diagnosis in first- or second-degree relatives (FDRs/SDRs). Fifty percent were diagnosed by clinical symptoms, 24% by PSA screening and for 26% the method of detection was unknown.

The descriptive PCa FRR was 3.18 (95% CI 2.92-3.45) for male FDRs in the population-based families. The FRRs were higher for brothers than fathers, and for FDRs of men diagnosed through PSA testing than for FDRs of men diagnosed through clinical symptoms (Table S2).

## Model development

A detailed description of the model-fitting process is available in the supplementary material and Table S3. The most parsimonious model is summarised in Table 1, and included the effects of *BRCA2*, *HOXB13* and *BRCA1*, together with a hypothetical recessively inherited allele and a PGC with age-dependent SD. The SD was 2.13 (95% CI 2.00-2.27) at age 70yr and decreased at a relative rate of 0.989 (95% CI 0.985-0.994) per year of age. The PGS explained 52.3% (95% CI 50.3%-54.4%) of the polygenic SD. The predicted age-specific FRRs were consistent with previously published FRR estimates (Figure S2A-B).<sup>1</sup>

## Sensitivity analyses

Ignoring the method of PCa detection in the ascertainment adjustment had a marked effect on the model parameters (Supplementary material, Table S4), but resulted in model-predicted FRRs that were considerably higher and inconsistent to those reported in large epidemiological studies (Figure S2C-D).<sup>1</sup> This was driven by the subgroup of families ascertained through PSA-screened probands (Table S4). We therefore did not pursue these models further.

## Model-predicted absolute risks

The average population risk is 16% by age 85yr. The corresponding model-predicted risk is 54% for *BRCA2* carriers, 39% for *HOXB13* G84E carriers, 17% for *BRCA1* carriers and 16% for non-carriers (Figure 1). Based on FH alone, the predicted risk for men with a relative diagnosed at age 50yr is 42%

when the father is affected and 43% when the brother is affected. These risks reduce to 27% and 26%, respectively, when the age-at-diagnosis is 80yr (Figure 1). Based on the PGS alone, the predicted risk varies between 4%-36% between the 5<sup>th</sup>-95<sup>th</sup> percentiles of the PGS distribution (Figure 1). The absolute risk differences by PGS are larger in those with FH (Figure 2) and those carrying PVs (Figure 3).

### External validation

Figure S3 summarises the inclusion and Table S5 the characteristics of the UK Biobank participants. The supplementary material and Table S6 detail the modified 268-SNP PGS used. There were 3,456 incident PCa cases within 5yr and 7,624 within 10yr.

### Discrimination

The predicted risk based on age had a C-index of 0.716 (95% CI 0.709-0.723) for prospective PCa diagnosis within 5yr and 0.693 (95% CI 0.688-0.698) within 10yr. Adding FH, PV or PGS information increased the C-indices. Including all available information, the corresponding C-indices were 0.790 (95% CI 0.783-0.797) and 0.772 (95% CI 0.768-0.777) respectively (Table 2).

In subgroups defined by age, FH, PV status or PGS quartile, the corresponding C indices ranged between 0.702-0.806 for 5yr and 0.692-0.789 for 10yr risks (Table S7).

### Calibration

The predicted risks based on all available information appeared to systematically underestimate the observed risks (Figure 4A-B). The underestimation was however apparent also when based only on

the year-and-age-specific population incidence (Figure S4A-B), and in age-, FH- or PV-status-based subgroups (Figures S5-S11A-B), indicating a higher PCa incidence in UK Biobank participants compared to the UK population incidence. After re-calibrating the predicted risks to account for the excess overall risk in UK Biobank (Supplementary material; Figure S12),<sup>54</sup> the model-predicted and observed risks were generally similar, both in the full dataset (Figure 4C-D) and in subgroups (Figures S5-S11C-D). The results indicated that the re-calibrated risks might be somewhat overestimated in the highest risk-decile (Figure 4C-D), but the difference was small (ratio of observed/predicted 10yr risks=0.90, 95% CI 0.87-0.93), and in participants with FH (Figure S8C-D).

### Risk classification

The participants with the top 1% of the predicted risks included 7.2% and 5.8% of the observed PCa cases within 5yr and 10yr respectively. Expanding to the top 10% of the predicted risks identified 38.5% and 34.8% of the cases respectively. 89.1% and 86.3% of the cases had above-median predicted risks (Table S8; Figure S13).

## Discussion

We have developed a comprehensive genetic PCa risk model for European ancestry men, using UKGPCS, the largest family-based PCa study of its kind. The model allows for personalising PCa risks based on a consultant's age, detailed cancer FH, moderate-to-high risk *BRCA2*, *HOXB13* and *BRCA1* PVs, and a 268-SNP PGS. In the large independent prospective UK Biobank cohort, the model discriminated well between individuals unaffected or affected with PCa within 5 or 10yr, and the predicted risks were in line with the observed risks after re-calibration to accommodate a higher-than-population-risk in the cohort.

In the model, familial PCa aggregation is explained by the known PVs, a PGC with a SD that decreases with age, together with an additional high-risk recessive allele. The 268-SNP PGS explains 52.3% (95% CI 50.3%-54.4%) of the PGC's SD. The putative recessive high-risk allele is consistent with the higher FRRs observed between brothers than in father-son pairs in this study and in previous observational studies.<sup>1,2</sup> The result is also consistent with previous segregation analysis studies.<sup>55-57</sup> However, to date, no PCa recessive susceptibility loci have been identified, and it is more likely that such a recessive component reflects several alleles that collectively behave in a recessive manner, or potentially other factors that explain the FRR patterns. In particular, the patterns might be driven by more frequent PSA testing in brothers than sons of affected men, as men with PCa FH are more likely to be PSA tested than other men<sup>58</sup> and PCa FRRs are higher during the first year after a FDR's PCa diagnosis,<sup>59,60</sup> particularly after a brother's diagnosis.<sup>59</sup> The estimated RR for homozygote carriers was higher when the method of diagnosis was ignored in the ascertainment adjustment and in the subgroup of families of probands diagnosed by PSA test, indicating that the result may partially be driven by PSA screening effects. However, early reports also suggested higher risks for brothers of affected men than for sons, even before widespread PSA test availability.<sup>61</sup> In addition, twin studies found that little PCa risk variation is attributable to shared familial non-genetic factors.<sup>3,4</sup> Taken together, these suggest that variants which act in a recessive manner may explain some of the higher FRR to brothers of cases, but direct identification of such variants in association studies will be required to confirm this. Notwithstanding, the model provides a good fit to the data and hence a rational basis for risk prediction.

In family-based studies, relatives are ascertained through an affected family member and are generally at a higher-than-average risk of disease. Therefore, it is critical to adjust for the ascertainment to avoid biased parameter estimates.<sup>62-64</sup> The participants diagnosed by PSA testing had FRRs that were higher than FRR estimates reported in population-based studies.<sup>1,2</sup> This may

reflect a greater PSA screening rate by FH.<sup>58</sup> To address this, we adjusted for potential ascertainment due to family phenotypes in all families of probands who were not diagnosed through symptomatic PCas. This provided FRR estimates that are consistent with those reported in large population-based studies.<sup>1,2</sup>

The PCa risks observed for UK Biobank participants were higher than corresponding year-and-age-specific population incidences. The UK Biobank participants have been reported to have higher socioeconomic status than the general UK population.<sup>65</sup> PSA testing rates vary by socioeconomic status,<sup>66</sup> and might explain this excess PCa risk. Consistently, the model-predicted risks underestimated those in UK Biobank, but after adjusting for the overall excess PCa risk in the cohort, the predicted risks were consistent with the observed risks in most risk categories.

The model can be expanded with the inclusion of new PVs, as evidence and reliable risk estimates become available for additional genes associated with PCa risk.<sup>44,67–72</sup> Similarly, although the model incorporates the latest 268-SNP PGS,<sup>18</sup> the model is flexible and can incorporate alternative PGSs, provided that an estimate of the proportion of the PGC that is explained by the PGS is available.<sup>73</sup> As further risk variants are identified, the model discrimination is expected to improve.

The validation results demonstrate that the model provides high levels of PCa risk-stratification in the population, and hence might facilitate the identification of men who could benefit from screening and other early detection interventions. For example, the half with above-median predicted risks included 89.1% of all prospective PCa cases observed within 5yr. Previous research has suggested that targeted PSA-based screening of *BRCA2* PV carriers<sup>8,74</sup> or based on PGS-

stratification could reduce overdiagnosis rates<sup>75</sup> and be cost-effective.<sup>76</sup> Future studies should evaluate the impact of risk-stratified screening based on a more comprehensive risk prediction model such the model presented here.

The study has limitations. The ascertainment adjustment is limited by a lack of data on PSA testing history in the UKGPCS families and data on whether FH influenced screening decisions of PSA-test-diagnosed probands; it may be an overadjustment that has resulted in reduced precision in the parameter estimates compared to the estimates that could have been achievable if exact information were available. A growing body of evidence suggests that the risk to *BRCA2* carriers varies by the location of the PV within the gene.<sup>77–79</sup> The model does not incorporate this variation. This requires more precise estimates of the risks associated with PVs in each region than are currently available. The use of self-reported cancer FH data may be limited by underreporting and inaccuracies.<sup>80</sup> However, model-predicted FRRs were consistent with FRRs reported in observational studies. Furthermore, the participants were unaware of their genotypic information at study entry, and so differential reporting of FH by PV status or PGS is unlikely. In the validation cohort, the FH data did not include information on relatives' age-at-diagnosis or information on unaffected relatives. We inferred plausible ages-at-diagnosis based on assumed familial age structures, but did not make assumptions about the unaffected relatives. This may explain the somewhat higher than expected risks in the FH-positive subgroup, as inclusion of unaffected relatives would have attenuated the risks. Despite these limitations, there was a clear gradient toward higher observed risks with higher predicted risks, and the predicted risks discriminated well between cases and non-cases also in the subgroup with FH. *BRCA2* PVs are associated with high-grade PCa,<sup>5,8,9</sup> but previous evidence suggests that overall risks based on *HOXB13*<sup>11–13</sup> or *BRCA1* PVs<sup>5,8,9</sup> or the 268-SNP PGS<sup>25</sup> are similarly predictive of high-/low-grade PCa. Both UKGPCS and UK Biobank lacked grade data on the self-reported PCas in relatives, so we could not estimate grade-specific FRRs, despite some previous

observational evidence suggesting that brothers tend to develop similar-grade PCas.<sup>81</sup> Grade data on UKBiobank participants' incident PCAs are not currently available; therefore, validation of grade-specific risks was also not possible. However, the majority of the UKGPCS probands had symptomatic PCAs, which tend to be more aggressive than preclinical PCAs.<sup>82</sup> Taken together with the *BRCA2* risks<sup>5,8,9</sup> and evidence suggesting grade-specific FRRs,<sup>81</sup> it is likely that the model predictions reflect more clinically significant disease risks. This may also partly explain the under-predicted risks in UK Biobank, prior to re-calibration. However, further research is needed on genetic predictors for aggressive PCa and on validating the prediction of specifically aggressive PCa risks. The model does not incorporate non-familial/genetic factors, such as PSA or other clinical measurements. Importantly, the model was developed and validated in men of European ancestry. PCa risks are higher in men of African ancestry and lower in men of Asian ancestry,<sup>83</sup> and further adaptation will be required to provide calibrated risks across all ancestries.

## Conclusion

This multifactorial risk prediction model is the first to incorporate the effects of the currently known moderate-to-high-risk and common low-risk PCa risk variants together with detailed FH information. The model predicts consistent familial risks, and shows good discrimination and calibration in an independent prospective validation cohort. The model will be beneficial for counselling of men in cancer family clinics, and can form the basis for future research evaluating risk-stratified population screening approaches.

## Manuscript data

**Acknowledgements:** We thank the participants in the UKGPCS and UK Biobank studies. We acknowledge the NCRN nurses and consultants for their work in the UKGPCS study, and the PRACTICAL consortium (<http://practical.icr.ac.uk/>) for organising genotyping which provided data for UKGPCS.

This research has been conducted using the UK Biobank Resource under Application Number 28126.

**Funding:** The risk modelling work was supported by Cancer Research UK (grants C12292/A20861, C12292/A22820 and PPRPGM-Nov20\100002).

The UKGPCS was supported by Cancer Research UK (grant numbers C5047/A7357, C5047/A3354, C5047/A10692, C5047/A15007 and C5047/A17528).

Further support was provided by: Prostate Research Campaign UK (now Prostate Cancer UK), The Institute of Cancer Research and The Everyman Campaign, The National Cancer Research Network UK, The National Cancer Research Institute (NCRI) UK.

We are grateful for support of National Institute for Health Research (NIHR) funding to: the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014); and the NIHR Biomedical Research Centre at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

The sponsors played no direct role in the study.

**Author contributions:**

Conception and design: Nyberg, Lee, Easton, Tischkowitz, Kote-Jarai, Eeles, Antoniou.

Provision of study materials or patients: Govindasami, Lophatananon, Muir, UK Genetic Prostate Cancer Study Collaborators, Bancroft, Kote-Jarai, Eeles.

Collection and assembly of data: Brook, Dennis, Yang, Wilcox, Dadaev, Govindasami, Leslie, Lophatananon, Muir, UK Genetic Prostate Cancer Study Collaborators, Bancroft, Kote-Jarai, Eeles.

Data analysis and interpretation: Nyberg, Brook, Ficorella, Lee, Lush, Easton, Antoniou.

Supervision: Tischkowitz, Antoniou.

Funding acquisition: Easton, Kote-Jarai, Eeles, Antoniou.

Manuscript writing: Nyberg, Antoniou.

Manuscript revision: All authors.

Final approval of manuscript: All authors.

Accountable for all aspects of the work: All authors.

**Conflicts of interest:** None.

**Prior presentation:** The development of the risk model was presented in part at the 70th Annual Meeting of the American Society of Human Genetics, virtual, October 27-30, 2020; the external validation has not previously been presented.

**Data sharing statement:** Individual pedigree-level data from UKGPCS are not publicly available as individuals could potentially be identifiable from the family structure. However, we confirm that summary-level data are available on request. The data that were used for validation are available by application to UK Biobank (<https://www.ukbiobank.ac.uk/enable-your-research>). Sufficient information on the risk prediction algorithm and on the genetic and familial predictive components

to allow replication is provided in the manuscript and supplementary material. The algorithms are also available on request for research purposes from the authors.

## References

1. Brandt A, Bermejo JL, Sundquist J, et al: Age-specific risk of incident prostate cancer and risk of death from prostate cancer defined by the number of affected family members. *Eur Urol* 58:275–80, 2010
2. Kiciński M, Vangronsveld J, Nawrot TS: An epidemiological reappraisal of the familial aggregation of prostate cancer: a meta-analysis. *PLoS One* 6:e27130, 2011
3. Page WF, Braun MM, Partin AW, et al: Heredity and prostate cancer: A study of World War II veteran twins. *Prostate* 33:240–245, 1997
4. Mucci LA, Hjelmborg JB, Harris JR, et al: Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. *JAMA* 315:68, 2016
5. Nyberg T, Tischkowitz M, Antoniou AC: BRCA1 and BRCA2 pathogenic variants and prostate cancer risk: systematic review and meta-analysis. *Br J Cancer* 126:1067–1081, 2022
6. Breast Cancer Linkage Consortium: Cancer risks in BRCA2 mutation carriers. *J Natl Cancer Inst* 91:1310–6, 1999
7. Thompson D, Easton DF, Breast Cancer Linkage Consortium: Cancer incidence in BRCA1 mutation carriers. *J Natl Cancer Inst* 94:1358–1365, 2002
8. Page EC, Bancroft EK, Brook MN, et al: Interim Results from the IMPACT Study: Evidence for Prostate-specific Antigen Screening in BRCA2 Mutation Carriers. *Eur Urol* 76:831–842, 2019
9. Nyberg T, Frost D, Barrowdale D, et al: Prostate Cancer Risks for Male BRCA1 and BRCA2 Mutation Carriers: A Prospective Cohort Study. *Eur Urol* 77:24–35, 2020
10. Nyberg T, Govindasami K, Leslie G, et al: Homeobox B13 G84E Mutation and Prostate Cancer Risk. *Eur Urol* 75:834–845, 2019

11. Laitinen VH, Wahlfors T, Saaristo L, et al: HOXB13 G84E mutation in Finland: population-based analysis of prostate, breast, and colorectal cancer risk. *Cancer Epidemiol Biomarkers Prev* 22:452–60, 2013
12. Karlsson R, Aly M, Clements M, et al: A population-based assessment of germline HOXB13 G84E mutation and prostate cancer risk. *Eur Urol* 65:169–76, 2014
13. Kote-Jarai Z, Mikropoulos C, Leongamornlert DA, et al: Prevalence of the HOXB13 G84E germline mutation in British men and correlation with prostate cancer risk, tumour characteristics and clinical outcomes. *Ann Oncol* 26:756–61, 2015
14. Wei J, Shi Z, Na R, et al: Germline HOXB13 G84E mutation carriers and risk to twenty common types of cancer: results from the UK Biobank. *Br J Cancer* 123:1356–1359, 2020
15. Schumacher FR, Al Olama AA, Berndt SI, et al: Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat Genet* 50:928–936, 2018
16. Dadaev T, Saunders EJ, Newcombe PJ, et al: Fine-mapping of prostate cancer susceptibility loci in a large meta-analysis identifies candidate causal variants. *Nat Commun* 9:2256, 2018
17. Matejic M, Saunders EJ, Dadaev T, et al: Germline variation at 8q24 and prostate cancer risk in men of European ancestry. *Nat Commun* 9:4616, 2018
18. Conti D V, Darst BF, Moss LC, et al: Trans-ancestry genome-wide association meta-analysis of prostate cancer identifies new susceptibility loci and informs genetic risk prediction. *Nat Genet* 53:65–75, 2021
19. Nieder A, Taneja S, Zeegers M, et al: Genetic counseling for prostate cancer risk. *Clin Genet* 63:169–76, 2003
20. National Comprehensive Cancer Network: NCCN Clinical Practice Guidelines in Oncology: Prostate Cancer Early Detection. Version 1.2022 [Internet], 2022[cited 2022 May 29] Available from:

<https://www.nccn.org>

- 21.** Liede A, Metcalfe K, Hanna D, et al: Evaluation of the Needs of Male Carriers of Mutations in BRCA1 or BRCA2 Who Have Undergone Genetic Counseling. *Am J Hum Genet* 67:1494–1504, 2000
- 22.** Giri VN, Knudsen KE, Kelly WK, et al: Role of genetic testing for inherited prostate cancer risk: Philadelphia prostate cancer consensus conference 2017. *J Clin Oncol* 36:414–424, 2018
- 23.** Al Olama AA, Benlloch S, Antoniou AC, et al: Risk analysis of prostate cancer in PRACTICAL, a multinational consortium, using 25 known prostate cancer susceptibility loci. *Cancer Epidemiol Prev biomarkers* 24:1121–1129, 2015
- 24.** Szulkin R, Whittington T, Eklund M, et al: Prediction of individual genetic risk to prostate cancer using a polygenic score. *Prostate* 75:1467–1474, 2015
- 25.** Plym A, Penney KL, Kalia S, et al: Evaluation of a Multiethnic Polygenic Risk Score Model for Prostate Cancer. *JNCI J Natl Cancer Inst* 114:771–774, 2022
- 26.** Lecarpentier J, Silvestri V, Kuchenbaecker KB, et al: Prediction of breast and prostate cancer risks in male BRCA1 and BRCA2 mutation carriers using polygenic risk scores. *J Clin Oncol* 35:2240–2250, 2017
- 27.** Darst BF, Sheng X, Eeles RA, et al: Combined Effect of a Polygenic Risk Score and Rare Genetic Variants on Prostate Cancer Risk. *Eur Urol* 80:134–138, 2021
- 28.** Barnes DR, Silvestri V, Leslie G, et al: Breast and Prostate Cancer Risks for Male BRCA1 and BRCA2 Pathogenic Variant Carriers Using Polygenic Risk Scores. *JNCI J Natl Cancer Inst* 114:109–122, 2022
- 29.** Antoniou AC, Pharoah PPD, Smith P, et al: The BOADICEA model of genetic susceptibility to breast and ovarian cancer. *Br J Cancer* 91:1580–1590, 2004
- 30.** Antoniou A, Cunningham A, Peto J, et al: The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br J Cancer* 98:1457–1466, 2008

31. Lee A, Mavaddat N, Wilcox AN, et al: BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genet Med* 21:1708–1718, 2019
32. Carver T, Hartley S, Lee A, et al: CanRisk Tool—A Web Interface for the Prediction of Breast and Ovarian Cancer Risk and the Likelihood of Carrying Genetic Pathogenic Variants. *Cancer Epidemiol Biomarkers Prev* 30:469–473, 2021
33. Lee A, Mavaddat N, Cunningham AP, et al: Enhancing the BOADICEA cancer risk prediction model to incorporate new data on RAD51C, RAD51D, BARD1, updates to tumour pathology and cancer incidence. *J Med Genet* (in press), 2022
34. Xu J, Sun J, Kader AK, et al: Estimation of absolute risk for prostate cancer using genetic markers and family history. *Prostate* 69:1565–1572, 2009
35. Sun J, Kader AK, Hsu F-C, et al: Inherited genetic markers discovered to date are able to identify a significant number of men at considerably elevated risk for prostate cancer. *Prostate* 71:421–430, 2011
36. Macinnis RJ, Antoniou AC, Eeles RA, et al: A risk prediction algorithm based on family history and common genetic variants: application to prostate cancer with potential clinical impact. *Genet Epidemiol* 35:549–56, 2011
37. So H-C, Kwan JSH, Cherny SS, et al: Risk Prediction of Complex Diseases from Family History and Known Susceptibility Loci, with Applications for Cancer Screening. *Am J Hum Genet* 88:548–565, 2011
38. Lindström S, Schumacher FR, Cox D, et al: Common genetic variants in prostate cancer risk prediction—results from the NCI Breast and Prostate Cancer Cohort Consortium (BPC3). *Cancer Epidemiol Biomarkers Prev* 21:437–444, 2012
39. Chen H, Liu X, Brendler CB, et al: Adding genetic risk score to family history identifies twice as many high-risk men for prostate cancer: Results from the prostate cancer prevention trial. *Prostate*

76:1120–1129, 2016

40. Mars N, Koskela JT, Ripatti P, et al: Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. *Nat Med* 26:549–557, 2020
41. Shi Z, Platz EA, Wei J, et al: Performance of Three Inherited Risk Measures for Predicting Prostate Cancer Incidence and Mortality: A Population-based Prospective Analysis. *Eur Urol* 79:419–426, 2021
42. Amos CI, Dennis J, Wang Z, et al: The OncoArray Consortium: A Network for Understanding the Genetic Architecture of Common Cancers. *Cancer Epidemiol Biomarkers Prev* 26:126–135, 2017
43. Mijuskovic M, Saunders EJ, Leongamornlert DA, et al: Rare germline variants in DNA repair genes and the angiogenesis pathway predispose prostate cancer patients to develop metastatic disease. *Br J Cancer* 119:96–104, 2018
44. Leongamornlert DA, Saunders EJ, Wakerell S, et al: Germline DNA Repair Gene Mutations in Young-onset Prostate Cancer Cases in the UK: Evidence for a More Extensive Genetic Panel. *Eur Urol* 76:329–337, 2019
45. Lane JA, Donovan JL, Davis M, et al: Active monitoring, radical prostatectomy, or radiotherapy for localised prostate cancer: study design and diagnostic and baseline results of the ProtecT randomised phase 3 trial. *Lancet Oncol* 15:1109–1118, 2014
46. Bycroft C, Freeman C, Petkova D, et al: The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562:203–209, 2018
47. Szustakowski JD, Balasubramanian S, Kvikstad E, et al: Advancing human genetics research and drug discovery through exome sequencing of the UK Biobank. *Nat Genet* 53:942–948, 2021
48. Wilcox N, Dumont M, González-Neira A, et al: Exome sequencing identifies novel susceptibility genes and defines the contribution of coding variants to breast cancer risk. *medRxiv* , 2022
49. Thomas DC: *Statistical methods in genetic epidemiology*. New York, Oxford University Press,

2004

**50.** International Agency for Research on Cancer: Cancer incidence in five continents (CI5), Volumes I to X [Internet] Available from: <https://ci5.iarc.fr/CI5I-X/>

**51.** Office for National Statistics: Cancer registration statistics, England, 2010-2017

[Internet] Available from:

<https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatistics/cancerregistrationstatisticsengland>

**52.** Lange K, Papp JC, Sinsheimer JS, et al: Mendel: the Swiss army knife of genetic analysis programs. *Bioinformatics* 29:1568–70, 2013

**53.** Shute NC, Ewens WJ: A resolution of the ascertainment sampling problem. III. Pedigrees. *Am J Hum Genet* 43:387–95, 1988

**54.** Pennells L, Kaptoge S, Wood A, et al: Equalization of four cardiovascular risk algorithms after systematic recalibration: individual-participant meta-analysis of 86 prospective studies. *Eur Heart J* 40:621–631, 2019

**55.** Cui J, Staples MP, Hopper JL, et al: Segregation Analyses of 1,476 Population-Based Australian Families Affected by Prostate Cancer. *Am J Hum Genet* 68:1207–1218, 2001

**56.** Pakkanen S, Baffoe-Bonnie AB, Matikainen MP, et al: Segregation analysis of 1,546 prostate cancer families in Finland shows recessive inheritance. *Hum Genet* 121:257–267, 2007

**57.** MacInnis RJ, Antoniou AC, Eeles RA, et al: Prostate cancer segregation analyses using 4390 families from UK and Australian population-based studies. *Genet Epidemiol* 34:42–50, 2010

**58.** Shah M, Zhu K, Palmer RC, et al: Family history of cancer and utilization of prostate, colorectal and skin cancer screening tests in U.S. men. *Prev Med (Baltim)* 44:459–464, 2007

**59.** Bermejo JL, Hemminki K: Familial Risk of Cancer Shortly After Diagnosis of the First Familial

Tumor. *JNCI J Natl Cancer Inst* 97:1575–1579, 2005

**60.** Bratt O, Garmo H, Adolfsson J, et al: Effects of Prostate-Specific Antigen Testing on Familial Prostate Cancer Risk Estimates. *JNCI J Natl Cancer Inst* 102:1336–1343, 2010

**61.** Woolf CM: An investigation of the familial aspects of carcinoma of the prostate. *Cancer* 13:739–744, 1960

**62.** Wacholder S, Hartge P, Struwing JP, et al: The kin-cohort study for estimating penetrance. *Am J Epidemiol* 148:623–30, 1998

**63.** Hopper JL, Bishop DT, Easton DF: Population-based family studies in genetic epidemiology. *Lancet (London, England)* 366:1397–406, 2005

**64.** Easton DF, Pharoah PDP, Antoniou AC, et al: Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med* 372:2243–57, 2015

**65.** Fry A, Littlejohns TJ, Sudlow C, et al: Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am J Epidemiol* 186:1026–1034, 2017

**66.** Morgan RM, Steele RJC, Nabi G, et al: Socioeconomic Variation and Prostate Specific Antigen Testing in the Community: A United Kingdom Based Population Study. *J Urol* 190:1207–1212, 2013

**67.** Raymond VM, Mukherjee B, Wang F, et al: Elevated Risk of Prostate Cancer Among Men With Lynch Syndrome. *J Clin Oncol* 31:1713–1718, 2013

**68.** Dominguez-Valentin M, Sampson JR, Seppälä TT, et al: Cancer risks by gene, age, and gender in 6350 carriers of pathogenic mismatch repair variants: findings from the Prospective Lynch Syndrome Database. *Genet Med* 22:15–25, 2020

**69.** Bancroft EK, Page EC, Brook MN, et al: A prospective prostate cancer screening programme for men with pathogenic variants in mismatch repair genes (IMPACT): initial results from an

international prospective study. *Lancet Oncol* 22:1618–1631, 2021

**70.** Wu Y, Yu H, Zheng SL, et al: A comprehensive evaluation of CHEK2 germline mutations in men with prostate cancer. *Prostate* 78:607–615, 2018

**71.** Pritchard CC, Mateo J, Walsh MF, et al: Inherited DNA-repair gene mutations in men with metastatic prostate cancer. *N Engl J Med* 375:443–453, 2016

**72.** Karlsson Q, Brook MN, Dadaev T, et al: Rare Germline Variants in ATM Predispose to Prostate Cancer: A PRACTICAL Consortium Study. *Eur Urol Oncol* 4:570–579, 2021

**73.** Mavaddat N, Ficoella L, Carver T, et al: Incorporating alternative Polygenic Risk Scores into the BOADICEA risk prediction model. *medRxiv* , 2022

**74.** Gulati R, Cheng HH, Lange PH, et al: Screening men at increased risk for prostate cancer diagnosis: Model estimates of benefits and harms. *Cancer Epidemiol Biomarkers Prev* 26:222–227, 2017

**75.** Pashayan N, Duffy SW, Neal DE, et al: Implications of polygenic risk-stratified screening for prostate cancer on overdiagnosis. *Genet Med* 17:789–795, 2015

**76.** Callender T, Emberton M, Morris S, et al: Polygenic risk-tailored screening for prostate cancer: A benefit–harm and cost-effectiveness modelling study. *PLOS Med* 16:e1002998, 2019

**77.** Thompson D, Easton D, Breast Cancer Linkage Consortium: Variation in cancer risks, by mutation position, in BRCA2 mutation carriers. *Am J Hum Genet* 68:410–419, 2001

**78.** Patel VL, Busch EL, Friebel TM, et al: Association of Genomic Domains in BRCA1 and BRCA2 with Prostate Cancer Risk and Aggressiveness. *Cancer Res* 80:624–638, 2020

**79.** Nyberg T, Frost D, Barrowdale D, et al: Prostate Cancer Risk by BRCA2 Genomic Regions. *Eur Urol* 78:494–497, 2020

**80.** Murff HJ, Spigel DR, Syngal S: Does this patient have a family history of cancer? An evidence-

based analysis of the accuracy of family cancer history. *JAMA* 292:1480–9, 2004

**81.** Jansson KF, Akre O, Garmo H, et al: Concordance of Tumor Differentiation Among Brothers with Prostate Cancer. *Eur Urol* 62:656–661, 2012

**82.** Pashayan N, Pharoah P, Neal DE, et al: Stage Shift in PSA-detected Prostate Cancers – Effect Modification by Gleason Score. *J Med Screen* 16:98–101, 2009

**83.** Siegel DA, O’Neil ME, Richards TB, et al: Prostate Cancer Incidence and Survival, by Stage and Race/Ethnicity — United States, 2001–2017. *MMWR Morb Mortal Wkly Rep* 69:1473–1480, 2020

## Figure legends

Figure 1: Predicted cumulative prostate cancer risks for a 45 year old consultand by

- (A) father's age at PCa diagnosis,
- (B) brother's age at PCa diagnosis,
- (C) pathogenic variants, or
- (D) polygenic score percentile.

For comparison, all graphs show the population average risk (grey curve). Consultands and brothers were assumed to be born after 1960 and fathers were assumed to be born in the 1930-39 birth cohort.

Figure 2: Predicted cumulative prostate cancer risks for a 45 year old consultand by combinations of family history and polygenic score percentile:

- (A) father diagnosed at age 50, by polygenic score;
- (B) father diagnosed at age 60, by polygenic score;
- (C) father diagnosed at age 70, by polygenic score; or
- (D) father diagnosed at age 80, by polygenic score.

For comparison, all graphs show the population average risk (grey curve). Consultands were assumed to be born after 1960 and fathers were assumed to be born in the 1930-39 birth cohort.

Figure 3: Predicted cumulative prostate cancer risks for a 45 year old consultand, by combinations of family history, pathogenic variant and polygenic score percentile:

- (A) positive for *BRCA2*, by family history and polygenic score percentile;
- (B) positive for *HOXB13* G84E, by family history and polygenic score percentile;
- (C) positive for *BRCA1* PV, by family history and polygenic score percentile; or

(D) negative for pathogenic variants in *BRCA2*, *HOXB13* and *BRCA1*, by family history and polygenic score percentile.

For comparison, all graphs show the population average risk (grey curve). Consultands were assumed to be born after 1960 and fathers were assumed to be born in the 1930-39 birth cohort.

Figure 4: Calibration plots of the model-predicted and observed prostate cancer risks in the external UK Biobank validation cohort. The graphs show the mean predicted risk under the risk model based on all age, family history, pathogenic variant and polygenic score information available at baseline, within groups defined by the deciles of the model-predicted risks, against the corresponding observed prospective risks based on the Kaplan-Meier estimator: prostate cancer risk within

(A) 5 yr;

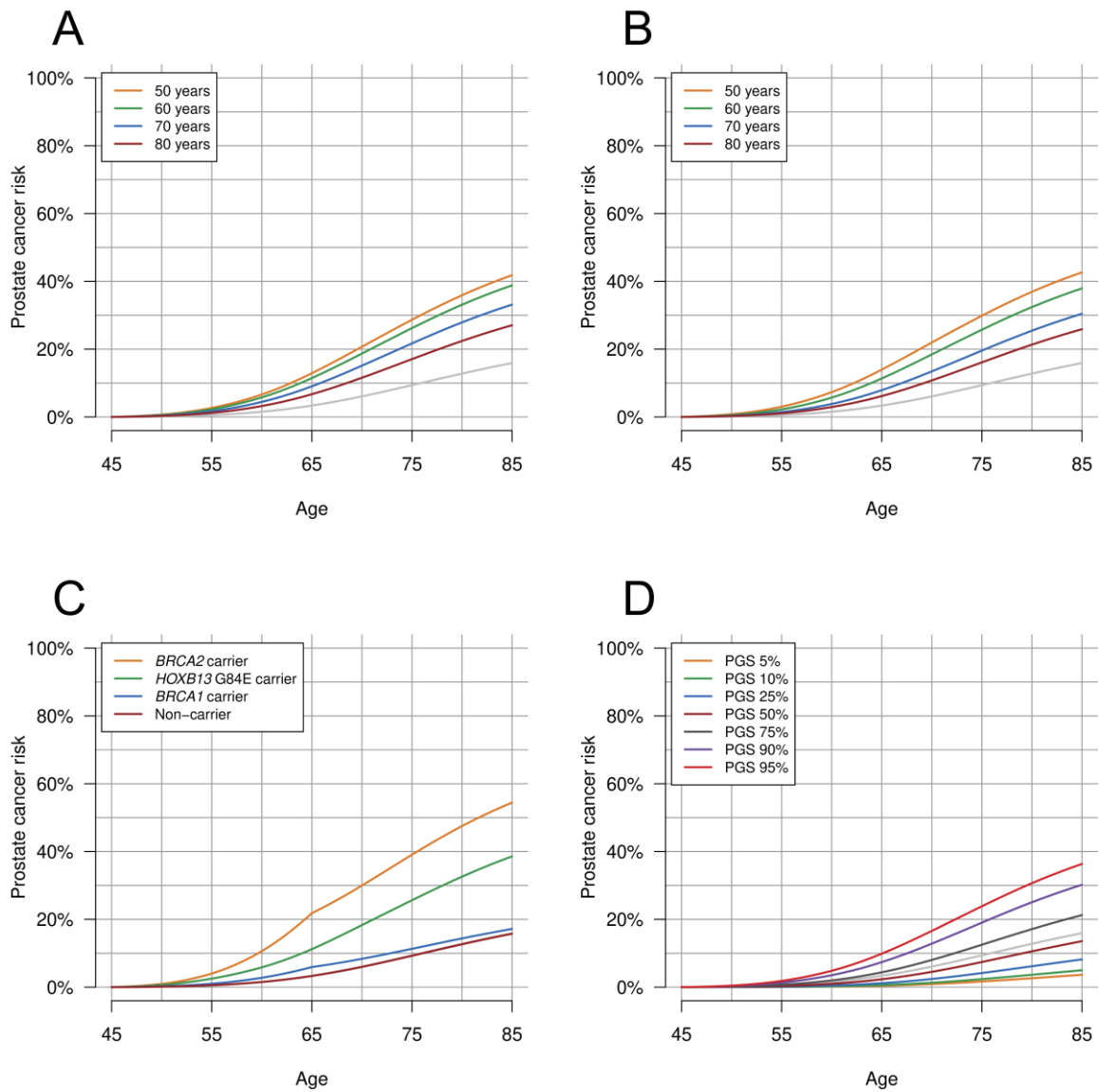
(B) 10 yr;

(C) 5 yr, after re-calibrating the risks to account for the excess prostate cancer risk observed in the UK Biobank participants;

(D) 10 yr, after re-calibrating the risks to account for the excess prostate cancer risk observed in the UK Biobank participants (see supplementary material and Figure S12).

**Figure 1:** Predicted cumulative prostate cancer risks for a 45 year old consultant by  
 (A) father's age at PCa diagnosis,  
 (B) brother's age at PCa diagnosis,  
 (C) pathogenic variants, or  
 (D) polygenic score percentile.

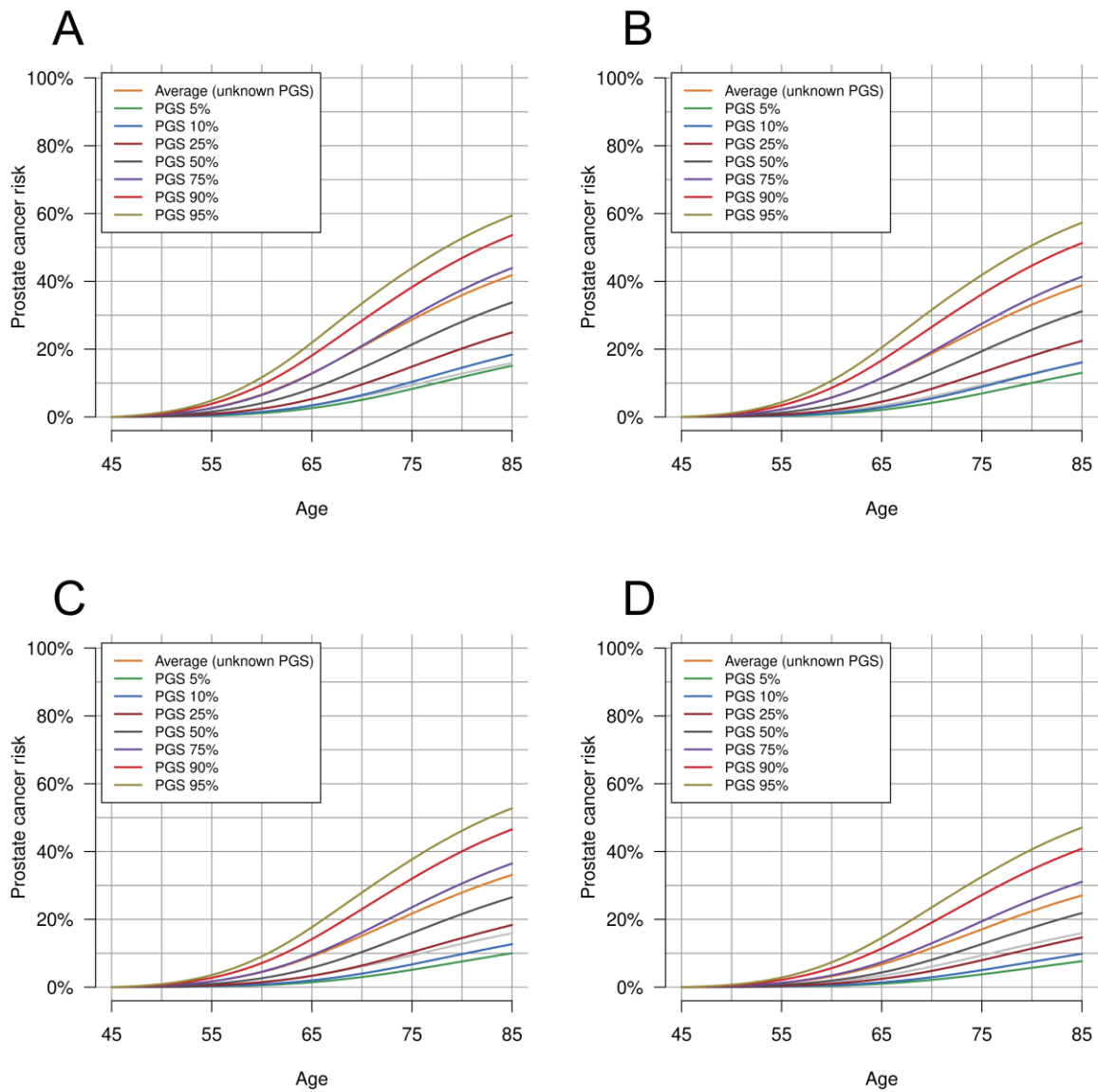
For comparison, all graphs show the population average risk (grey curve). Consultants and brothers were assumed to be born after 1960 and fathers were assumed to be born in the 1930-39 birth cohort.



**Figure 2:** Predicted cumulative prostate cancer risks for a 45 year old consultant by combinations of family history and polygenic score percentile:

- (A) father diagnosed at age 50, by polygenic score;  
 (B) father diagnosed at age 60, by polygenic score;  
 (C) father diagnosed at age 70, by polygenic score; or  
 (D) father diagnosed at age 80, by polygenic score.

For comparison, all graphs show the population average risk (grey curve). Consultants were assumed to be born after 1960 and fathers were assumed to be born in the 1930-39 birth cohort.



**Figure 3:** Predicted cumulative prostate cancer risks for a 45 year old consultand, by combinations of family history, pathogenic variant and polygenic score percentile:

(A) positive for *BRCA2*, by family history and polygenic score percentile;

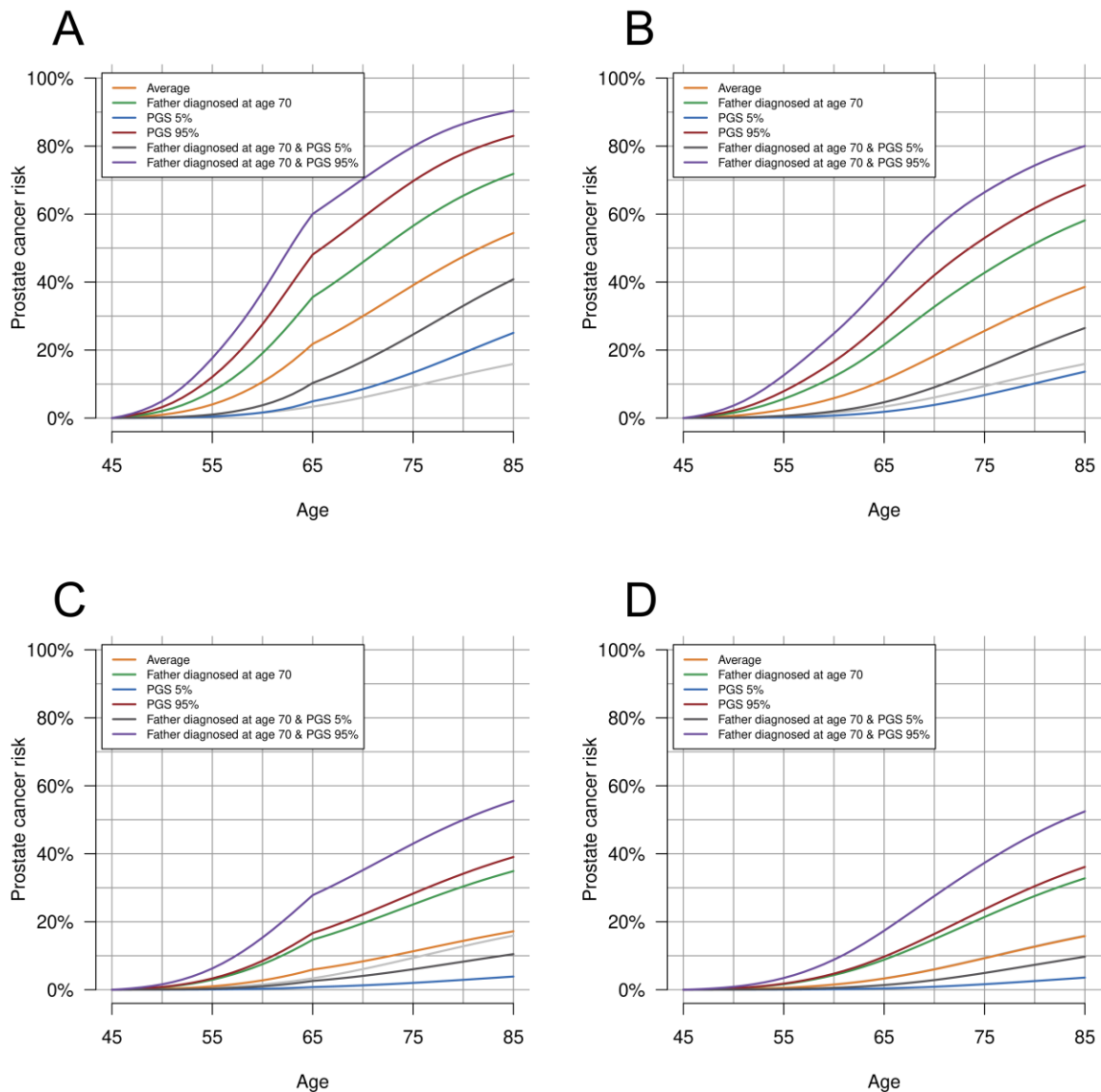
(B) positive for *HOXB13* G84E, by family history and polygenic score percentile;

(C) positive for *BRCA1* PV, by family history and polygenic score percentile; or

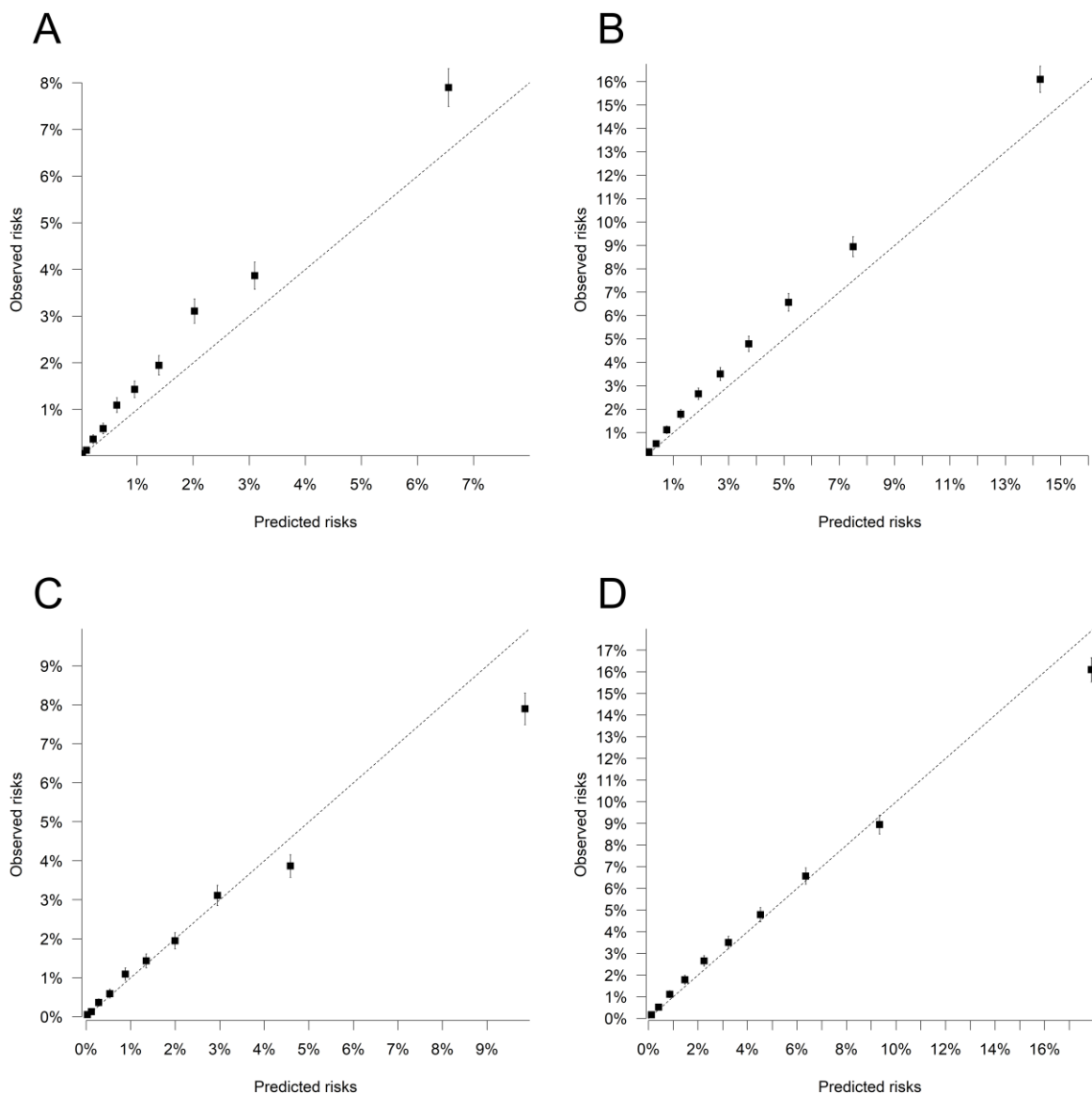
(D) negative for pathogenic variants in *BRCA2*, *HOXB13* and *BRCA1*, by family history and polygenic score percentile.

For comparison, all graphs show the population average risk (grey curve). Consultands were

assumed to be born after 1960 and fathers were assumed to be born in the 1930-39 birth cohort.



**Figure 4:** Calibration plots of the model-predicted and observed prostate cancer risks in the external UK Biobank validation cohort. The graphs show the mean predicted risk under the risk model based on all age, family history, pathogenic variant and polygenic score information available at baseline, within groups defined by the deciles of the model-predicted risks, against the corresponding observed prospective risks based on the Kaplan-Meier estimator: prostate cancer risk within (A) 5 yr; (B) 10 yr; (C) 5 yr, after re-calibrating the risks to account for the excess prostate cancer risk observed in the UK Biobank participants; (D) 10 yr, after re-calibrating the risks to account for the excess prostate cancer risk observed in the UK Biobank participants (see supplementary material and Figure S12).



## Tables

**Table 1: Risk model parameters.**

Parameter estimates used by the final prostate cancer risk model.

Abbreviations: LL, log-likelihood. k, no. of parameters. AIC, Akaike information criterion. LRT, likelihood-ratio test. RR, relative risk. PCa, prostate cancer. CI, confidence interval. RAF, risk allele frequency. SD, standard deviation. PV, pathogenic variant. PGS, polygenic score.

Genetic component	Parameter	Subgroup, if applicable	Estimate (95% CI)
<i>BRCA2</i>	Risk allele frequency		0.10% [a]
	Relative risk of prostate cancer [b]	Age <65	7.14 [c]
		Age ≥65	3.84 [c]
<i>HOXB13</i> G84E	Risk allele frequency		0.21% (0.14%-0.32%)
	Relative risk of prostate cancer (per-copy)	Birth cohort <1930	3.17 (1.78-5.65)
		Birth cohort ≥1930	5.93 (3.40-10.4)
<i>BRCA1</i>	Risk allele frequency		0.06% [a]
	Relative risk of prostate cancer [b]	Age <65	1.78 [c]
		Age ≥65	0.91 [c]
Hypothetical recessive locus	Risk allele frequency		6.33% (2.81%-13.7%)
	Relative risk of prostate cancer (homozygous allele carriers vs non-carriers/heterozygous carriers)		48.1 (11.2-206)
Polygenic component	Polygenic SD (log-linear model)[d]	Age 70 (intercept)	2.13 (2.00-2.27)
		Per year of age	0.989 (0.985-0.994)
Polygenic score (PGS)	Proportion of the polygenic SD that is explained by the PGS		52.3% (50.3%-54.4%)

[a] External estimates from the BOADICEA model.<sup>29–31,33</sup>

[b] In addition to the parameters shown, the models incorporate external piecewise linear age-specific RR estimates for female breast and ovarian cancer from the BOADICEA model.<sup>29–31,33</sup>

[c] External estimates from a previous meta-analysis.<sup>5</sup>

[d] The log-linear age-specific polygenic SD at age  $t$  was specified as

$$\ln SD(t) = a_0 + a_1 \times (t-70) \text{ at age } t,$$

where  $a_0$  corresponds to the estimated SD at age 70 and  $a_1$  the change in SD per year of age.

**Table 2: Risk model discrimination performance in the external UK Biobank prospective cohort.**

Time frame	n prostate cancer/total (%)	Predictors used	Model-predicted risk, median (interquartile range)		Concordance index (95% CI)
			Participants without prostate cancer	Participants diagnosed with prostate cancer	
5 years	3456/170850 (2.02%)	Age only	1.40% (0.44%-2.51%)	2.51% (1.63%-2.91%)	0.716 (0.709-0.723)
		Age + Family history	1.36% (0.37%-2.38%)	2.42% (1.65%-2.90%)	0.720 (0.714-0.727)
		Age + PVs	1.38% (0.43%-2.47%)	2.47% (1.61%-2.89%)	0.718 (0.711-0.725)
		Age + PGS	0.77% (0.21%-1.94%)	2.86% (1.49%-5.18%)	0.787 (0.780-0.794)
		Age + Family history + PVs	1.34% (0.37%-2.36%)	2.41% (1.66%-2.88%)	0.723 (0.716-0.729)
		Age + Family history + PGS	0.77% (0.21%-1.95%)	2.97% (1.50%-5.44%)	0.788 (0.782-0.795)
		Age + PVs + PGS	0.76% (0.21%-1.93%)	2.89% (1.48%-5.24%)	0.789 (0.782-0.796)
		Age + Family history + PVs + PGS	0.76% (0.20%-1.94%)	2.97% (1.51%-5.50%)	0.790 (0.783-0.797)
10 years	7624/170850 (4.46%)	Age only	3.83% (1.29%-5.46%)	5.46% (4.21%-6.41%)	0.693 (0.688-0.698)
		Age + Family history	3.69% (1.29%-5.51%)	5.41% (4.06%-6.32%)	0.696 (0.691-0.701)
		Age + PVs	3.80% (1.27%-5.50%)	5.42% (4.17%-6.37%)	0.695 (0.690-0.700)
		Age + PGS	2.18% (0.74%-4.82%)	6.52% (3.53%-10.95%)	0.770 (0.765-0.775)
		Age + Family history + PVs	3.64% (1.28%-5.48%)	5.37% (4.02%-6.28%)	0.698 (0.693-0.703)
		Age + Family history + PGS	2.16% (0.72%-4.87%)	6.62% (3.55%-11.35%)	0.771 (0.766-0.775)
		Age + PVs + PGS	2.16% (0.73%-4.79%)	6.55% (3.53%-11.04%)	0.772 (0.767-0.777)
		Age + Family history + PVs + PGS	2.14% (0.71%-4.84%)	6.65% (3.56%-11.43%)	0.772 (0.768-0.777)