

1 Systematic comparison of family history and polygenic risk 2 across 24 common diseases

3

4

5 Nina Mars,^{1,2} Joni V. Lindbohm,^{3,4} Pietro della Briotta Parolo,¹ Elisabeth Widén,¹ Jaakko Kaprio,^{1,3}
6 Aarno Palotie,^{1,2,5} *FinnGen*,* Samuli Ripatti^{1,3,6} #

7

8 1) Institute for Molecular Medicine Finland, FIMM, HiLIFE, University of Helsinki, Helsinki, Finland, 2) Stanley Center
9 for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA, 3) Clincum, Department of Public
10 Health, University of Helsinki, Helsinki, Finland, 4) Department of Epidemiology and Public Health, University College
11 London, London, UK, 5) Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General
12 Hospital, Boston, MA, USA, 6) Broad Institute of MIT and Harvard, Cambridge, MA, USA

13

14

15

16

17

18

19

20 Words: 3,092

21

22 **# Corresponding author**

23 Tel: +358 40 567 0826

24 Email: samuli.ripatti@helsinki.fi

25 Address: Institute for Molecular Medicine (FIMM), Biomedicum 2U, Tukholmankatu 8, 00290, Helsinki,
26 Finland

27

28 * A list of FinnGen contributors is listed in the supplement

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

1 **ABSTRACT**

2 Family history is the standard indirect measure of inherited susceptibility in clinical care, while
3 polygenic risk scores (PRS) have more recently demonstrated potential for more directly capturing
4 genetic risk in many diseases. No studies have systematically compared how these overlap and
5 complement each other across common diseases. Within FinnGen (N=306,418), we leverage family
6 relationships, up to 50 years of nationwide registries, and genome-wide genotyping to examine the
7 interplay of family history and genome-wide PRSs. We explore the dynamic for three types of family
8 history across 24 common diseases: first- and second-degree family history, and parental causes of
9 death. Covering a large proportion of the burden of non-communicable diseases in adults, we show
10 that family history and PRS are independent and not interchangeable measures, but instead provide
11 complementary information of inherited disease susceptibility. The PRSs explained on average 10%
12 of the effect of first-degree family history, and first-degree family history 3% of PRSs, and PRS
13 effects were independent of both early- and late-onset family history. The PRS stratified the risk
14 similarly in individuals with and without family history. In most diseases, including coronary artery
15 disease, glaucoma, and type 2 diabetes, a positive family history with a high PRS was associated with
16 a considerably elevated risk, whereas a low PRS compensated completely for the risk implied by
17 positive family history. This study provides a catalogue of risk estimates for both family history of
18 disease and PRSs, and highlights opportunities for a more comprehensive way of assessing inherited
19 disease risk across common diseases.

20

21

22 **KEYWORDS**

23 Polygenic risk score; family history; risk assessment

24

1 INTRODUCTION

2 Family history (FH) is a risk factor in most common, non-communicable diseases.¹ With multiple
3 advantages including low cost and non-invasiveness, it captures both genetic and non-genetic familial
4 risk and is therefore widely applied for risk stratification and health promotion. Common clinical
5 applications include assessment of FH of breast cancer for targeted screening, earlier initiation of
6 cardiovascular disease prevention, and evaluating the likelihood of rheumatic disease in patients with
7 inflammatory arthritis.²⁻⁴ Despite the advantages, assessment of FH also has important limitations in
8 capturing inherited disease risk. Many individuals with common diseases have no FH, or may not
9 know the diseases their relatives have, and the same level of familial risk is assigned to all relatives
10 of similar degree. The accuracy of FH is fairly low owing to factors such as recall bias, and sensitivity
11 to wording in queries may lead to misinterpretation of risk.^{5,6} With average family sizes declining in
12 many developed countries,⁷ FH will also provide increasingly less information for a comprehensive
13 assessment of familial risk.

14 The algorithmic developments and rapid growth in genome-wide genetic testing provides a
15 more personalized approach for measuring genetic susceptibility through polygenic risk scores
16 (PRS).^{8,9} PRSs employ information from large-scale genetic screens comparing allele frequencies in
17 thousands of disease cases and healthy controls and have identified numerous genetic loci for virtually
18 all common diseases.¹⁰ To estimate polygenic risks, the common genetic variation and the effects on
19 the disease risks are integrated into a single metric, the PRS. The effectiveness of PRSs in risk
20 stratification has been demonstrated for many diseases, with predictive value demonstrated alongside
21 established clinical risk assessment tools.¹¹ Similarly, PRSs modify risk among carriers of high-risk
22 variants and identify high-risk individuals for whom existing prediction tools are suboptimal.¹¹⁻¹⁶

23 Given the initial expense of implementing PRS estimation in a clinical setting relative to the
24 seemingly simple questions pertaining to family history, systematic evaluation of the independent
25 added benefit of PRS across common diseases is essential. Studies on individual diseases have
26 observed fairly independent effects of PRS and first-degree FH,^{11,15,17-27} but no studies have
27 systematically compared the relative contributions and overlap of PRS and FH across different types
28 of familial risk, across varying genetic architectures, and across a wide range of diseases. Moreover,
29 only few studies have used genome-wide PRSs, although these contemporary PRSs containing a large

1 number of variants have demonstrated improved performance beyond PRSs with less variants due to
2 high polygenicity in common diseases.^{13,28-30} Here we study the interplay of first and second-degree
3 FH, parental causes of death, and genome-wide PRSs for 24 diseases using FinnGen (N=306,418),
4 showing that FH and PRSs are largely independent and provide complementary information in risk
5 assessment.

7 **MATERIAL AND METHODS**

8 *Participants and diseases*

9 This observational study uses FinnGen study Data Freeze 7, a collection of 306,418 adults (age ≥ 18)
10 from epidemiological cohorts, disease-based cohorts, and hospital biobanks (**Supplementary Table**
11 **1**). We used three binary definitions for FH: 1) any type of first-degree FH (FH_{1st} morbidity or
12 mortality), 2) any type of second-degree FH (FH_{2nd}), and 3) parental cause of death (FH_P). Both for
13 the index patients and their relatives (i.e. how FH was obtained), cases were identified through
14 nationwide healthcare registries. The 24 diseases were chosen based on availability of large published
15 genome-wide association studies (GWAS) with full summary statistics available for genome-wide
16 PRSs (**Supplementary Table 2**). Case definitions are in **Supplementary Table 3**. Registry follow-
17 up ended at Dec 31, 2019, with parental causes of death available until Dec 31, 2018. For FH_P, we
18 studied 15 out of the 24 diseases, identifying causes of death (immediate, contributing, and underlying
19 causes of death). Details on genotypes, PRS generation and inference of relatedness is in
20 Supplementary Methods.

22 *Polygenic risk scores*

23 For each of the 24 diseases, we constructed disease-specific PRSs in a systematic manner
24 (**Supplementary Table 2**). The number of cases in the GWASs ranged from 3,769 (epilepsy) to
25 567,460 (eGFR used for chronic kidney disease). PRS-CS was used for inferring posterior effect sizes
26 for calculation of genome-wide PRSs³¹ with 1000 Genomes Project European sample (N = 503) as
27 the external linkage disequilibrium (LD) reference panel.³² The PRS was analyzed primarily as a
28 continuous variable, with selected analyses applying either a 1) binary definition of FH, with high
29 PRS defined as a PRS in the top decile of the distribution, with the rest as the reference group, or 2)

1 PRS categories 0-10%, 10-20%, 20-40%, 40-60%, 60-80%, 80-90%, and 90-100%, with the
2 reference group being 40-60%. To assess the impact of high vs low PRS, the reference category was
3 33rd to 90th percentiles, and low PRS was defined as the lowest tertile of the distribution, to allow for
4 a sufficient number of cases with low PRS.

5

6 ***Statistical analysis***

7 Associations between FH, PRS, and risk of disease was assessed with logistic regression, with models
8 adjusted for sex, birth year, genotyping array, cohort, and the first ten genetic principal components
9 of ancestry. The same adjustments were applied when calculating area under the receiver operator
10 characteristic curve (AUC) estimates. AUC estimates with 95% CIs were calculated with R package
11 pROC. Interactions between FH and the continuous PRS (scaled to zero mean and unit variance) was
12 assessed by introducing their interaction term to the regression model, assessing statistical
13 significance set at a p-value threshold of 0.0013 (Bonferroni-correction for 24+15 tests). Cumulative
14 incidences by age 80 were estimated with Kaplan-Meier survival curves (R package survminer).
15 Statistical analyses were performed using R, version 4.1.0.

16

17

18 **RESULTS**

19 FinnGen comprises 306,418 individuals (56.3% women; mean age 59.8 at the end of follow-up in
20 2019, s.d. 17.3). For the 24 diseases, FH was defined as 1) first-degree family history, FH_{1st}
21 (morbidity or mortality), 2) second-degree family history, FH_{2nd}, and 3) parental cause of death, FH_p.
22 Each identifies the relatives' diagnoses systematically through nationwide registries, including the
23 hospital discharge registry (available from 1968-), causes of death registry (1964-), the Finnish
24 Cancer Registry (1953-). FH_{1st} and FH_{2nd} leverage the genetic relatedness within FinnGen: out of
25 306,418 individuals, we identified 39,444 with first-degree relative pairs based on the KING kinship
26 coefficient³³ (see methods for details; 60.3% women; mean age 53.0, s.d. 16.5; parent-offspring
27 relationship in 19,261 individuals, full-sibling in 20,183). For breast cancer, we studied only women
28 (15,281 individuals, mother-daughter relationship in 7,770; full sisters in 7,511) and for prostate
29 cancer, only men (9,473 individuals; father-son relationship in 3,932; full brothers in 5,541).

1 Similarly, we identified 47,154 individuals with a second-degree relative in the dataset (63.2%
2 women, mean age 47.5, s.d. 15.0; N=18,973 for breast cancer; N=12,355 for prostate cancer).
3 Parental causes of death (FH_P), were linked through causes of death registry available from 1964 to
4 2019 and we excluded 78,436 whose both parents have died before 1964 or who had missing data on
5 both parents (e.g. due to emigration), resulting in 227,982 individuals (mean age 53.6, s.d. 15.1;
6 N=133,653 for breast cancer; N=94,329 for prostate cancer; 70,225 (30.1%) with one and 73,299
7 (32.2%) with two dead parents). See **Supplementary Figure 1** for study flow diagram.

8

9 ***Family history and risk of disease***

10 First, we systematically evaluated the effects of FH on risk of disease. **Figure 1** shows the prevalence
11 of the diseases, and the prevalence and effect sizes for positive FH. The most common diseases were
12 cardiometabolic diseases, followed by knee osteoarthritis and hypothyroidism. Positive FH_{1st} was
13 significantly associated with higher risk of disease in all 24 diseases except stroke. The effect sizes
14 ranged from odds ratio (OR) 3.25 (95% confidence interval, CI, 2.41–4.37) in chronic kidney disease
15 to OR 1.17 (0.98–1.39) in stroke (**Supplementary Table 4**). For FH_{2nd}, 18/24 diseases showed
16 evidence of an association, with their effect sizes ranging from OR 1.85 (1.19–2.89) in colorectal
17 cancer to OR 1.17 (1.09–1.25) in hypertension (**Supplementary Table 5**). Compared to FH_{1st}, the
18 effect sizes for FH_{2nd} were on average 69.1% lower (s.d. 25.0%; calculated from log odds), i.e. a third
19 of the effect of FH_{1st}. For FH_P, out of the 24 diseases, we studied 15 diseases that are well captured
20 by causes of death and used information from all recorded causes of death (immediate, contributing,
21 and underlying causes of death on the death certificate). For all 15 diseases, we observed an
22 association between FH_P and risk of disease, with effect sizes ranging from OR 2.82 (2.25–3.53) in
23 seropositive rheumatoid arthritis to OR 1.12 (1.04–1.20) in stroke (**Supplementary Table 6**).
24 Compared to FH_{1st}, the effect sizes for FH_P were on average 30.1% lower (s.d. 22.4%), i.e. two thirds
25 of the effect of FH_{1st}.

26

27 ***Overlap of family history and polygenic risk***

28 Next, we compared the overlap between FH and PRSs. We constructed 24 genome-wide PRSs with
29 uniform methodology using PRS-CS³¹, one for each disease (**Supplementary Table 2**). We first

1 compared the effects sizes per standard deviation (SD) increase for PRS, and FH_{1st} (**Figure 2,**
2 **Supplementary Table 4**). The PRS was associated with elevated risk in all 24 diseases. The higher
3 the PRS, the higher was the proportion of positive FH (**Supplementary Figure 2**). Effect sizes for
4 the PRS ranged from OR 2.33 (95%CI 2.10–2.558) in prostate cancer to OR 1.12 (1.05-1.20) in
5 epilepsy. Adjusting the PRS effect size with FH_{1st}, the change in effect size was small (mean decrease
6 as log odds -3.0%, s.d. 1.3%). Adjusting the effect of FH_{1st} with PRS led to a mean decrease of -
7 10.3% (s.d. 6.0%), i.e. PRS explained of tenth of first-degree family history. No effect size decrease
8 was observed for PRS adjusting with FH_{2nd} (**Supplementary Table 5**). We observed similar results
9 for FH_P (**Supplementary Table 6**; effect size decrease adjusting PRS effects with FH_P -0.7%, s.d.
10 0.6%, vice versa -14.5%, s.d. 9.2%). Proportional decreases in log odds by disease for all definitions
11 of FH are in **Figure 3**. FH generally explained a much smaller fraction of the effect of PRS than vice
12 versa. A similar pattern was observed categorizing the PRS and comparing high PRS (>90th
13 percentile) to the rest of the distribution (**Supplementary Table 7, Supplementary Figure 3**). A
14 high PRS conferred on average similar effect sizes as FH_{1st}. The effect sizes particularly in common
15 cancers and cardiometabolic diseases were higher for the PRS, whereas the effect sizes for psychiatric
16 diseases were higher for FH_{1st}.

17 As early-onset FH is considered a particularly important familial risk factor, we also assessed
18 the impact of FH_P divided into tertiles of age at death. The largest effect size was observed for FH_P
19 with the lowest age tertile, in line with early-onset FH being a stronger risk factor than late-onset FH.
20 Adjusting the PRS with this FH_P divided into age tertiles had no impact on the effect sizes of the
21 PRSs. Adjusting this FH_P by PRS resulted in the largest effect size decreases for the youngest age
22 tertile, but the decreases were overall small. These show that the PRS was independent of both early-
23 and late-onset FH_P (**Supplementary Table 8, Figure 4**).

24 With formal interaction testing, we did not identify any systematic interactions between FH
25 and PRS (**Supplementary Figure 4**), which was further supported by observing similar PRS effect
26 sizes in individuals with positive and negative FH_{1st} (**Figure 5**).

27
28

1 ***Polygenic risk in individuals with a positive family history***

2 Next, having assessed the overlap between FH and the PRSs, we estimated how high and low PRS
3 impact disease risk in individuals with positive FH_{1st}. Looking at cumulative incidence of risk of
4 disease with the PRS divided into three groups (high PRS >90%, average PRS 33-90%, and low PRS
5 <33%), we observed that a low PRS systematically compensated for the impact of positive FH_{1st}, and
6 individuals with a combination of high PRS and positive FH_{1st} had a particularly high risk
7 (**Figure 6**). Survival curves for a broader set of diseases, and survival curves stratifying individuals
8 with no FH_{1st} into similar PRS groups are in **Supplementary Figures 5–6**.

9
10 ***Concordance of high polygenic risk in relatives***

11 Lastly, we assessed concordance – detection of a high PRS among first- and second-degree relatives,
12 relevant for cascade screening in relatives of individuals with high PRS. We evaluated two questions:
13 1) “What is the probability of having high PRS, if a relative has high PRS?” and 2) “How does this
14 probability differ with relative’s disease status?”. For 1), on average 33.7% of the first-degree and
15 19.8% of second-degree relatives had a similarly high PRS (**Supplementary Figures 7–8**). For 2),
16 the concordance was somewhat higher with positive FH_{1st} than with negative FH_{1st}, with an average
17 difference of 2.5% (range 0.0–7.9%). For FH_{2nd}, no difference with disease status was observed
18 (average 0.6%).

19
20
21 **DISCUSSION**

22 Covering a large proportion of the burden of non-communicable diseases in adults, we systematically
23 compared the overlap of polygenic risk and different types of family history, showing that they
24 provide independent and complementary information of inherited disease susceptibility in all 24
25 studied diseases. PRS explained on average 10% of the effect of FH_{1st}, but FH_{1st} only 3% of the PRSs,
26 and the PRSs were independent of both early- and late-onset family history. The PRS estimates
27 stratified risk similarly in individuals with and without positive FH: a high PRS conferred a
28 considerably elevated risk, whereas a low PRS compensated for the effect of FH.

1 To our knowledge, no previous studies have systematically compared how FH and PRS
2 overlap and complement each other across common diseases. Our results are in line with previous
3 disease-specific reports observing at most a modest attenuation in the effect of FH adjusting for PRS
4 in cardiometabolic diseases, cancers, and depression.^{11,15,17-27} We extend these by a systematic
5 comparison across 24 common diseases, using genome-wide PRSs generated with uniform
6 methodology, by measuring FH uniformly through nationwide healthcare registries, and by
7 leveraging genetic relatedness. Our results show that effects of FH and polygenic risk scores are
8 independent, indicating that these measures complement each other for assessment of inherited
9 disease risk. Compared to prevention guidelines that do not recommend use of PRS when FH is
10 available,³ these results provide important novel data supporting the use of PRS for improving risk
11 assessment of several diseases with major public health importance.

12 The largely independent effects have several potential explanations. In addition to capturing
13 shared DNA, FH measures non-genetic exposures and behaviors shared by families. In contrast, PRSs
14 capture each person's unique combinations of common, disease-associated genetic variants,
15 including genetic risk variation not shared by the relatives. PRSs can be measured in any phase of
16 life, whereas FH relies on disease events having actualized in relatives with utility in late-onset
17 diseases. FH also assigns a similar risk for all relatives of the same degree, despite everyone carrying
18 a unique set of genetic variants measurable through PRSs. Our observation of independent effects is
19 also in line with earlier reports showing the importance of FH of breast and ovarian cancers in *BRCA1*
20 and *BRCA2* mutation carriers.³⁴

21 Genetic information is typically considered in clinical care only when evidence-based
22 prevention strategies to attenuate risk are available.³⁵ For instance, risk assessment of cancers has
23 long tradition of comprehensive ascertainment of FH to identify familial clustering³⁶ when targeted
24 interventions and screening tools are available.^{2,37} Our results indicate that PRSs could be used to
25 refine risk assessment of breast, prostate, and colorectal cancer, even when information about FH is
26 available. In glaucoma, a high PRS and FH had equal and largely independent effects, but only FH is
27 currently used for assessing risk of glaucoma in patients with ocular hypertension.³⁸ The risk of
28 coronary artery disease and type 2 diabetes can be decreased by lifestyle interventions and
29 medications, and FH is commonly used for assessing their risk.^{3,39} For both diseases, we observed

1 larger effects for high PRS than for FH. Moreover, a high PRS may identify individuals more likely
2 to benefit from preventive treatments: for coronary artery disease, a high PRS can result in higher
3 relative efficacy of statins and disclosing PRS risk together with traditional risk factors can motivate
4 lifestyle changes.⁴⁰⁻⁴² In contrast, stroke PRSs and FH show lower effect sizes than other
5 cardiovascular diseases, likely owing to the heterogeneity of the disease and differing etiological
6 patterns of stroke subtypes.^{43,44}

7 This study has multiple strengths. FH was assessed systematically and comprehensively by
8 using linkages to high-quality nationwide registries, including hospital discharges, causes-of-death,
9 and medication reimbursement registries, overcoming several limitations of self-reported FH such as
10 recall bias, sensitivity to wording, and inter-individual differences in knowledge about FH.^{5,6,45} We
11 report effects of FH for disorders challenging to capture precisely from self-reported data, such as
12 alcohol use disorder and atrial fibrillation, and show effects for diseases less studied in the field of
13 PRSs including glaucoma and hypothyroidism. Moreover, we looked at the effect of a dichotomized
14 PRS for comparison purposes and applying PRSs on the continuous scale would provide further risk
15 stratification as unlike FH, extremes of PRSs can also be used to identify individuals at particularly
16 high or low risk. FinnGen's wide age range is a key strength of the study, allowing systematic
17 comparison of polygenic risk and FH across 24 diseases. Our results are also supported by
18 quantitative genetic theory.^{46,47} Average concordances of a high PRS among first- and second-degree
19 relatives was 33.7% and 19.8%, in line with estimates on cardiometabolic diseases in UK Biobank,⁴⁸
20 and in agreement theoretically derived concordance estimates of 32.4% and 19.3%.⁴⁷ Moreover, the
21 study provides catalogue of risk estimates for both FH of disease and PRSs in a large-scale biobank
22 study.

23 The study was limited to individuals of European ancestry among whom current PRSs have
24 the highest utility.⁴⁹ Although our recording of FH_{1st} and FH_{2nd} was primarily based on only one
25 relative, FH_p contained both parents, and FH estimates are well in line with earlier reports from
26 epidemiological cohorts and large registry studies (**Supplementary Table 9**). Although the various
27 registries are efficient in capturing cases, milder disease forms such as mild osteoarthritis or atopic
28 dermatitis may remain uncaptured. Similarly, common conditions such as depression or alcohol use
29 disorder are often underreported unless severe or contributing to somatic pathologies. With over half

1 of the study participants in the dataset ascertained from hospital biobanks or disease cohorts, the data
2 is somewhat enriched in cases, resulting in cumulative incidences may not be fully generalizable to
3 the population.

4 In conclusion, we studied the interplay of family history and genome-wide PRSs,
5 systematically comparing effects across 24 common diseases. The effects of family history and PRS
6 were largely independent, and the pattern was observed across the diseases. We demonstrate that
7 polygenic risk and family history are not interchangeable measures of genetic susceptibility. Instead,
8 they provide complementary information, bringing opportunities for a more comprehensive way of
9 assessing inherited risk. A PRS can be calculated early of life to serve as risk indicator in individuals
10 without family history of disease, while providing effective risk stratification also among individuals
11 with positive family history.

12
13
14 **Acknowledgements** We thank the researchers who have openly shared the GWAS summary statistics that were
15 used in this study to generate PRSs. We would like to thank Mervi Aavikko and Risto Kajanne for management
16 assistance. The FinnGen project is funded by two grants from Business Finland (HUS 4685/31/2016 and UH
17 4386/31/2016) and the following industry partners: AbbVie Inc., AstraZeneca UK Ltd, Biogen MA Inc., Bristol
18 Myers Squibb, Genentech Inc., Merck Sharp & Dohme Corp, Pfizer Inc., GlaxoSmithKline Intellectual Property
19 Development Ltd., Sanofi US Services Inc., Maze Therapeutics Inc., Janssen Biotech Inc, and Novartis Pharma
20 AG. Following biobanks are acknowledged for delivering biobank samples to FinnGen: Auria Biobank
21 (www.auria.fi/biopankki), THL Biobank (www.thl.fi/biobank), Helsinki Biobank (www.helsinginbiopankki.fi),
22 Biobank Borealis of Northern Finland ([https://www.ppshep.fi/Tutkimus-ja-opetus/Biopankki/Pages/Biobank-](https://www.ppshep.fi/Tutkimus-ja-opetus/Biopankki/Pages/Biobank-Borealis-briefly-in-English.aspx)
23 [Borealis-briefly-in-English.aspx](https://www.ppshep.fi/Tutkimus-ja-opetus/Biobank-Borealis-briefly-in-English.aspx)), Finnish Clinical Biobank Tampere ([www.tays.fi/en-](http://www.tays.fi/en-US/Research_and_development/Finnish_Clinical_Biobank_Tampere)
24 [US/Research_and_development/Finnish_Clinical_Biobank_Tampere](http://www.tays.fi/en-US/Research_and_development/Finnish_Clinical_Biobank_Tampere)), Biobank of Eastern Finland ([www.ita-](http://www.ita-suomenbiopankki.fi/en)
25 [suomenbiopankki.fi/en](http://www.ita-suomenbiopankki.fi/en)), Central Finland Biobank (www.ksshp.fi/fi-FI/Potilaalle/Biopankki), Finnish Red Cross
26 Blood Service Biobank (www.veripalvelu.fi/verenluovutus/biopankkitoiminta) and Terveystalo Biobank
27 (www.terveystalo.com/fi/Yritystietoa/Terveystalo-Biopankki/Biopankki/). All Finnish Biobanks are members of
28 BBMRI.fi infrastructure (www.bbMRI.fi) and FINBB biobank cooperative (<https://finbb.fi/>) is the coordinator of the
29 BBMRI-ERIC operations in Finland covering all Finnish biobanks.

30 This work has been supported by Academy of Finland (grant number 331671 to N.M., 285380 to S.R.,
31 128650 to A.P., 308248 to J.K.); Academy of Finland Center of Excellence in Complex Disease Genetics (grant

1 number 312062 to S.R., 312074 to A.P., 312073 to J.K.); European Union's Horizon 2020 research and innovation
2 programme under grant agreement No 101016775; the Sigrid Jusélius Foundation (to S.R. and A.P.); University of
3 Helsinki HiLIFE Fellow grants 2017-2020 (to S.R.); The Finnish Innovation Fund Tekes (grant number
4 2273/31/2017 to E.W.); Foundation and the Horizon 2020 Research and Innovation Programme (grant number
5 667301 (COSYN) to A.P.); JVL was supported by Academy of Finland (grant number 311492) and Helsinki
6 Institute of Life Science (H970).

7
8 **Declaration of Interests** A.P. is a member of the Pfizer Genetics Scientific Advisory Panel. The remaining
9 authors declare no competing interests.

10

11 **Data availability** The FinnGen data may be accessed through Finnish Biobanks' FinBB portal (web link:
12 www.finbb.fi, email: info.fingenious@finbb.fi). Download links for the GWAS summary statistics used for
13 constructing are provided in Supplementary Table 1. The weights for our polygenic risk scores will be made
14 available at PGS Catalog (pgs-info@ebi.ac.uk), with the submitted data embargoed until publication and
15 PGS Catalog identifiers reported in the final version of the manuscript.

16

17 **Code availability** The full genotyping and imputation protocol for FinnGen is described
18 at <https://doi.org/10.17504/protocols.io.xbgfijw>. The PRS-CS pipeline in FinnGen is described at
19 <https://github.com/FINNGEN/CS-PRS-pipeline>.

20

21 **Author contributions**

22 N.M., J.V.L., and S.R. conceived and designed the study. N.M. and P.D.B.P carried out the statistical and
23 computational analyses with advice from S.R. and J.V.L. Quality control of the data was carried out by N.M.
24 and P.D.B.P. All authors provided critical inputs to interpretation of the data. The manuscript was written and
25 revised by N.M. and S.R., with comments from all of the co-authors. All co-authors have approved the final
26 version of the manuscript.

27

1 **References**

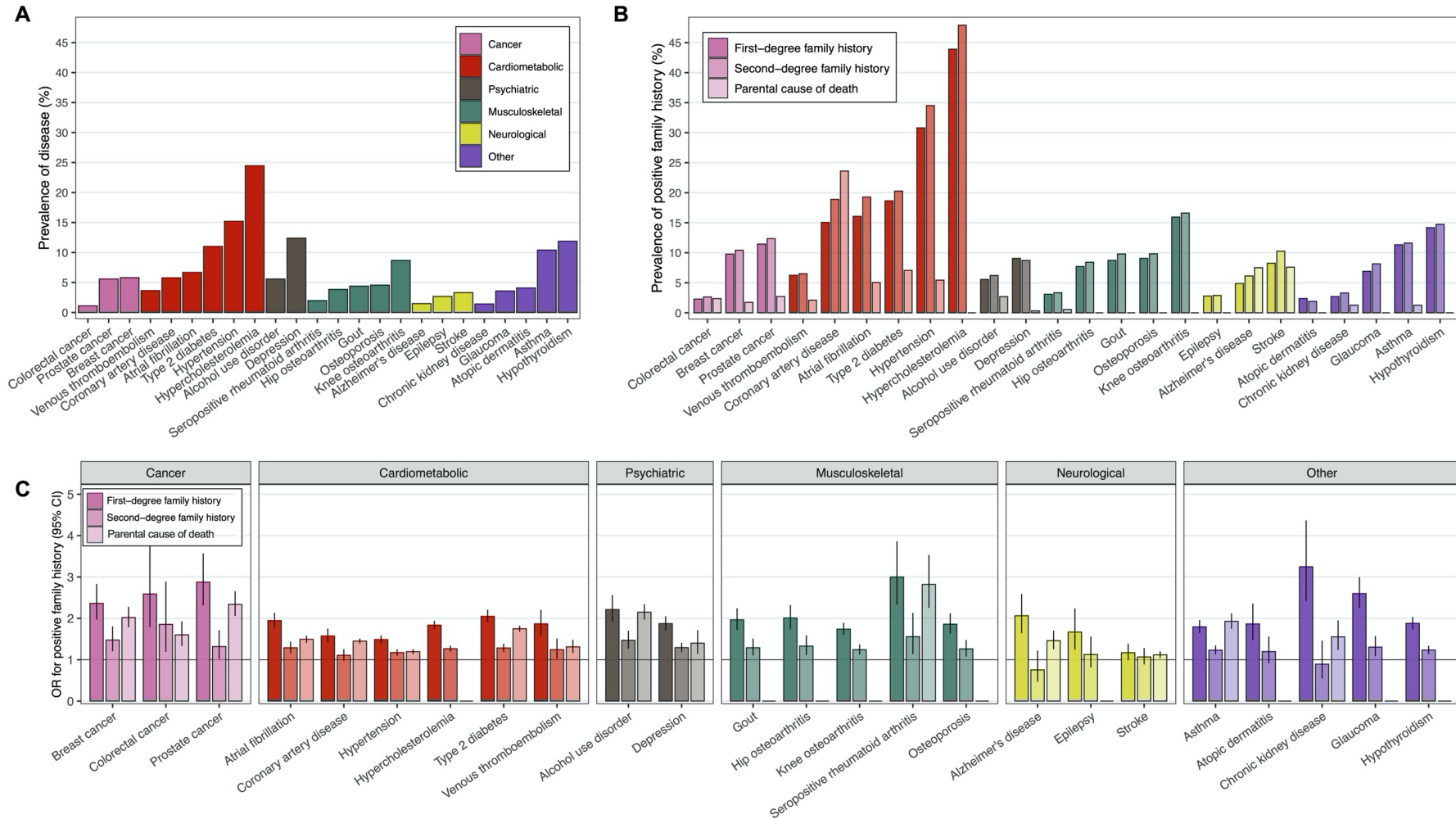
- 2
- 3 1. Valdez, R., Yoon, P.W., Qureshi, N., Green, R.F., and Khoury, M.J. (2010). Family history
4 in public health practice: a genomic tool for disease prevention and health promotion. *Annu.*
5 *Rev. Public Health* *31*, 69-87 61 p following 87.
6 10.1146/annurev.publhealth.012809.103621.
 - 7 2. Daly, M.B., Pal, T., Berry, M.P., Buys, S.S., Dickson, P., Domchek, S.M., Elkhanany, A.,
8 and Friedman, S. NCCN Clinical Practice Guidelines in Oncology: Genetics/Familial High-
9 Risk Assessment: Breast, Ovarian, and Pancreatic Version 2.2021. NCCN.org.
 - 10 3. Piepoli, M.F., Hoes, A.W., Agewall, S., Albus, C., Brotons, C., Catapano, A.L., Cooney,
11 M.T., Corra, U., Cosyns, B., Deaton, C., et al. (2016). 2016 European Guidelines on
12 cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the
13 European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention
14 in Clinical Practice (constituted by representatives of 10 societies and by invited
15 experts)Developed with the special contribution of the European Association for
16 Cardiovascular Prevention & Rehabilitation (EACPR). *Eur. Heart J.* *37*, 2315-2381.
17 10.1093/eurheartj/ehw106.
 - 18 4. Frisell, T., Saevarsdottir, S., and Askling, J. (2016). Family history of rheumatoid arthritis:
19 an old concept with new developments. *Nat Rev Rheumatol* *12*, 335-343.
20 10.1038/nrrheum.2016.52.
 - 21 5. Conway-Pearson, L.S., Christensen, K.D., Savage, S.K., Huntington, N.L., Weitzman, E.R.,
22 Ziniel, S.I., Bacon, P., Cacioppo, C.N., Green, R.C., and Holm, I.A. (2016). Family health
23 history reporting is sensitive to small changes in wording. *Genet. Med.* *18*, 1308-1311.
24 10.1038/gim.2016.45.
 - 25 6. Wilson, B.J., Qureshi, N., Santaguida, P., Little, J., Carroll, J.C., Allanson, J., and Raina, P.
26 (2009). Systematic review: family history in risk assessment for common diseases. *Ann.*
27 *Intern. Med.* *151*, 878-885. 10.7326/0003-4819-151-12-200912150-00177.
 - 28 7. United Nations, Department of Economic and Social Affairs, Population Division. Patterns
29 and trends in household size and composition: Evidence from a United Nations dataset.
30 (ST/ESA/SER.A/433). (2019).
 - 31 8. Lewis, C.M., and Vassos, E. (2020). Polygenic risk scores: from research tools to clinical
32 instruments. *Genome Med* *12*, 44. 10.1186/s13073-020-00742-5.
 - 33 9. Chatterjee, N., Shi, J., and Garcia-Closas, M. (2016). Developing and evaluating polygenic
34 risk prediction models for stratified disease prevention. *Nat. Rev. Genet.* *17*, 392-406.
35 10.1038/nrg.2016.27.
 - 36 10. Loos, R.J.F. (2020). 15 years of genome-wide association studies and no signs of slowing
37 down. *Nature Communications* *11*, 5900. 10.1038/s41467-020-19653-5.
 - 38 11. Mars, N., Koskela, J.T., Ripatti, P., Kiiskinen, T.T.J., Havulinna, A.S., Lindbohm, J.V.,
39 Ahola-Olli, A., Kurki, M., Karjalainen, J., Palta, P., et al. (2020). Polygenic and clinical risk
40 scores and their impact on age at onset and prediction of cardiometabolic diseases and
41 common cancers. *Nat. Med.* 10.1038/s41591-020-0800-0.
 - 42 12. Fahed, A.C., Wang, M., Homburger, J.R., Patel, A.P., Bick, A.G., Neben, C.L., Lai, C.,
43 Brockman, D., Philippakis, A., Ellinor, P.T., et al. (2020). Polygenic background modifies
44 penetrance of monogenic variants for tier 1 genomic conditions. *Nat Commun* *11*, 3635.
45 10.1038/s41467-020-17374-3.
 - 46 13. Mars, N., Widén, E., Kerminen, S., Meretoja, T., Pirinen, M., della Briotta Parolo, P., Palta,
47 P., Havulinna, A., Elliott, A., Shcherban, A., et al. (2020). The role of polygenic risk and

- 1 susceptibility genes in breast cancer over the course of life. *Nature Communications* 11,
2 6383. [10.1038/s41467-020-19966-5](https://doi.org/10.1038/s41467-020-19966-5).
- 3 14. Isgut, M., Sun, J., Quyyumi, A.A., and Gibson, G. (2021). Highly elevated polygenic risk
4 scores are better predictors of myocardial infarction risk early in life than later. *Genome*
5 *Med* 13, 13. [10.1186/s13073-021-00828-8](https://doi.org/10.1186/s13073-021-00828-8).
- 6 15. Archambault, A.N., Su, Y.R., Jeon, J., Thomas, M., Lin, Y., Conti, D.V., Win, A.K.,
7 Sakoda, L.C., Lansdorp-Vogelaar, I., Peterse, E.F.P., et al. (2020). Cumulative Burden of
8 Colorectal Cancer-Associated Genetic Variants Is More Strongly Associated With Early-
9 Onset vs Late-Onset Cancer. *Gastroenterology* 158, 1274-1286 e1212.
10 [10.1053/j.gastro.2019.12.012](https://doi.org/10.1053/j.gastro.2019.12.012).
- 11 16. Conti, D.V., Darst, B.F., Moss, L.C., Saunders, E.J., Sheng, X., Chou, A., Schumacher,
12 F.R., Olama, A.A.A., Benlloch, S., Dadaev, T., et al. (2021). Trans-ancestry genome-wide
13 association meta-analysis of prostate cancer identifies new susceptibility loci and informs
14 genetic risk prediction. *Nat. Genet.* 53, 65-75. [10.1038/s41588-020-00748-0](https://doi.org/10.1038/s41588-020-00748-0).
- 15 17. Inouye, M., Abraham, G., Nelson, C.P., Wood, A.M., Sweeting, M.J., Dudbridge, F., Lai,
16 F.Y., Kaptoge, S., Brozynska, M., Wang, T., et al. (2018). Genomic Risk Prediction of
17 Coronary Artery Disease in 480,000 Adults: Implications for Primary Prevention. *J. Am.*
18 *Coll. Cardiol.* 72, 1883-1893. [10.1016/j.jacc.2018.07.079](https://doi.org/10.1016/j.jacc.2018.07.079).
- 19 18. Tikkanen, E., Havulinna, A.S., Palotie, A., Salomaa, V., and Ripatti, S. (2013). Genetic risk
20 prediction and a 2-stage risk screening strategy for coronary heart disease. *Arterioscler.*
21 *Thromb. Vasc. Biol.* 33, 2261-2266. [10.1161/ATVBAHA.112.301120](https://doi.org/10.1161/ATVBAHA.112.301120).
- 22 19. Hughes, E., Tshiaba, P., Gallagher, S., Wagner, S., Judkins, T., Roa, B., Rosenthal, E.,
23 Domchek, S., Garber, J., Lancaster, J., et al. (2020). Development and Validation of a
24 Clinical Polygenic Risk Score to Predict Breast Cancer Risk. *JCO Precis Oncol* 4.
25 [10.1200/PO.19.00360](https://doi.org/10.1200/PO.19.00360).
- 26 20. Agerbo, E., Trabjerg, B.B., Borglum, A.D., Schork, A.J., Vilhjalmsen, B.J., Pedersen,
27 C.B., Hakulinen, C., Albinana, C., Hougaard, D.M., Grove, J., et al. (2021). Risk of Early-
28 Onset Depression Associated With Polygenic Liability, Parental Psychiatric History, and
29 Socioeconomic Status. *JAMA Psychiatry* 78, 387-397. [10.1001/jamapsychiatry.2020.4172](https://doi.org/10.1001/jamapsychiatry.2020.4172).
- 30 21. InterAct Consortium, Scott, R.A., Langenberg, C., Sharp, S.J., Franks, P.W., Rolandsson,
31 O., Drogan, D., van der Schouw, Y.T., Ekelund, U., Kerrison, N.D., et al. (2013). The link
32 between family history and risk of type 2 diabetes is not explained by anthropometric,
33 lifestyle or genetic risk factors: the EPIC-InterAct study. *Diabetologia* 56, 60-69.
34 [10.1007/s00125-012-2715-x](https://doi.org/10.1007/s00125-012-2715-x).
- 35 22. Tada, H., Melander, O., Louie, J.Z., Catanese, J.J., Rowland, C.M., Devlin, J.J., Kathiresan,
36 S., and Shiffman, D. (2016). Risk prediction by genetic risk scores for coronary heart
37 disease is independent of self-reported family history. *Eur. Heart J.* 37, 561-567.
38 [10.1093/eurheartj/ehv462](https://doi.org/10.1093/eurheartj/ehv462).
- 39 23. Muranen, T.A., Mavaddat, N., Khan, S., Fagerholm, R., Pelttari, L., Lee, A., Aittomaki, K.,
40 Blomqvist, C., Easton, D.F., and Nevanlinna, H. (2016). Polygenic risk score is associated
41 with increased disease risk in 52 Finnish breast cancer families. *Breast Cancer Res. Treat.*
42 158, 463-469. [10.1007/s10549-016-3897-6](https://doi.org/10.1007/s10549-016-3897-6).
- 43 24. Agerbo, E., Sullivan, P.F., Vilhjalmsen, B.J., Pedersen, C.B., Mors, O., Borglum, A.D.,
44 Hougaard, D.M., Hollegaard, M.V., Meier, S., Mattheisen, M., et al. (2015). Polygenic Risk
45 Score, Parental Socioeconomic Status, Family History of Psychiatric Disorders, and the
46 Risk for Schizophrenia: A Danish Population-Based Study and Meta-analysis. *JAMA*
47 *Psychiatry* 72, 635-641. [10.1001/jamapsychiatry.2015.0346](https://doi.org/10.1001/jamapsychiatry.2015.0346).
- 48 25. Mavaddat, N., Michailidou, K., Dennis, J., Lush, M., Fachal, L., Lee, A., Tyrer, J.P., Chen,
49 T.H., Wang, Q., Bolla, M.K., et al. (2019). Polygenic Risk Scores for Prediction of Breast

- 1 Cancer and Breast Cancer Subtypes. *Am. J. Hum. Genet.* *104*, 21-34.
2 10.1016/j.ajhg.2018.11.002.
- 3 26. Li, H., Feng, B., Miron, A., Chen, X., Beesley, J., Bimeh, E., Barrowdale, D., John, E.M.,
4 Daly, M.B., Andrulis, I.L., et al. (2017). Breast cancer risk prediction using a polygenic risk
5 score in the familial setting: a prospective study from the Breast Cancer Family Registry and
6 kConFab. *Genet. Med.* *19*, 30-35. 10.1038/gim.2016.43.
- 7 27. Moll, M., Lutz, S.M., Ghosh, A.J., Sakornsakolpat, P., Hersh, C.P., Beaty, T.H., Dudbridge,
8 F., Tobin, M.D., Mittleman, M.A., Silverman, E.K., et al. (2020). Relative contributions of
9 family history and a polygenic risk score on COPD and related outcomes: COPDGene and
10 ECLIPSE studies. *BMJ Open Respir Res* *7*. 10.1136/bmjresp-2020-000755.
- 11 28. Khera, A.V., Chaffin, M., Aragam, K.G., Haas, M.E., Roselli, C., Choi, S.H., Natarajan, P.,
12 Lander, E.S., Lubitz, S.A., Ellinor, P.T., and Kathiresan, S. (2018). Genome-wide polygenic
13 scores for common diseases identify individuals with risk equivalent to monogenic
14 mutations. *Nat. Genet.* *50*, 1219-1224. 10.1038/s41588-018-0183-z.
- 15 29. Dikilitas, O., Schaid, D.J., Kosel, M.L., Carroll, R.J., Chute, C.G., Denny, J.A., Fedotov, A.,
16 Feng, Q., Hakonarson, H., Jarvik, G.P., et al. (2020). Predictive Utility of Polygenic Risk
17 Scores for Coronary Heart Disease in Three Major Racial and Ethnic Groups. *Am. J. Hum.*
18 *Genet.* *106*, 707-716. 10.1016/j.ajhg.2020.04.002.
- 19 30. Mars, N., Kerminen, S., Feng, Y.-C.A., Kanai, M., Läll, K., Thomas, L.F., Skogholt, A.H.,
20 della Briotta Parolo, P., Neale, B.M., Smoller, J.W., et al. (2022). Genome-wide risk
21 prediction of common diseases across ancestries in one million people. *Cell Genomics* *2*.
22 10.1016/j.xgen.2022.100118.
- 23 31. Ge, T., Chen, C.Y., Ni, Y., Feng, Y.A., and Smoller, J.W. (2019). Polygenic prediction via
24 Bayesian regression and continuous shrinkage priors. *Nat Commun* *10*, 1776.
25 10.1038/s41467-019-09718-5.
- 26 32. Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang,
27 H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R.
28 (2015). A global reference for human genetic variation. *Nature* *526*, 68-74.
29 10.1038/nature15393.
- 30 33. Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.M. (2010).
31 Robust relationship inference in genome-wide association studies. *Bioinformatics* *26*, 2867-
32 2873. 10.1093/bioinformatics/btq559.
- 33 34. Kuchenbaecker, K.B., Hopper, J.L., Barnes, D.R., Phillips, K.A., Mooij, T.M., Roos-Blom,
34 M.J., Jervis, S., van Leeuwen, F.E., Milne, R.L., Andrieu, N., et al. (2017). Risks of Breast,
35 Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers.
36 *JAMA* *317*, 2402-2416. 10.1001/jama.2017.7112.
- 37 35. ACMG Board of Directors (2015). Clinical utility of genetic and genomic services: a
38 position statement of the American College of Medical Genetics and Genomics. *Genet.*
39 *Med.* *17*, 505-507. 10.1038/gim.2015.41.
- 40 36. Bennett, R.L. (2010). *The practical guide to genetic family history*, 2nd edition. (Wiley-
41 Blackwell).
- 42 37. Provenzale, D., Gupta, S., Ahnen, D.J., Chen, L.-M., Chung, D.C., Cooper, G., Earlu, D.S.,
43 Giardiello, F.M., and Grady, W. (2019). NCCN Clinical Practice Guidelines in Oncology:
44 Genetic/Familial High-Risk Assessment: Colorectal Version 3.2019. *J. Natl. Compr. Canc.*
45 *Netw.* *14*, 1010-1030. 10.6004/jnccn.2016.0108.
- 46 38. National Institute for Health and Care Excellence (2017). Glaucoma: diagnosis and
47 management [NICE guideline No. 81] <https://www.nice.org.uk/guidance/ng81>.
- 48 39. American Diabetes Association (2021). 3. Prevention or Delay of Type 2 Diabetes:
49 Standards of Medical Care in Diabetes-2021. *Diabetes Care* *44*, S34-S39. 10.2337/dc21-
50 S003.

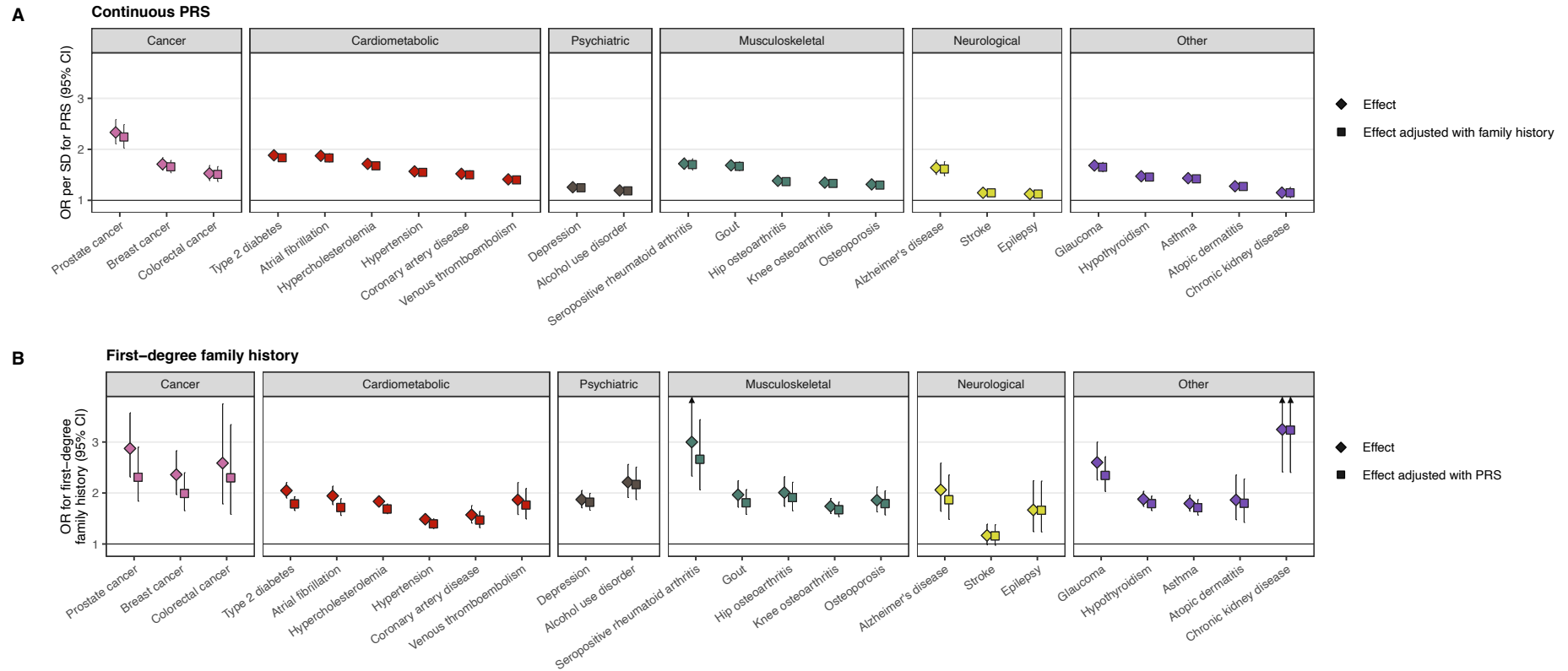
- 1 40. Natarajan, P., Young, R., Stitzel, N.O., Padmanabhan, S., Baber, U., Mehran, R., Sartori,
2 S., Fuster, V., Reilly, D.F., Butterworth, A., et al. (2017). Polygenic Risk Score Identifies
3 Subgroup With Higher Burden of Atherosclerosis and Greater Relative Benefit From Statin
4 Therapy in the Primary Prevention Setting. *Circulation* *135*, 2091-2101.
5 10.1161/CIRCULATIONAHA.116.024436.
- 6 41. Mega, J.L., Stitzel, N.O., Smith, J.G., Chasman, D.I., Caulfield, M., Devlin, J.J., Nordio, F.,
7 Hyde, C., Cannon, C.P., Sacks, F., et al. (2015). Genetic risk, coronary heart disease events,
8 and the clinical benefit of statin therapy: an analysis of primary and secondary prevention
9 trials. *Lancet* *385*, 2264-2271. 10.1016/S0140-6736(14)61730-X.
- 10 42. Widén, E., Junna, N., Ruotsalainen, S., Surakka, I., Mars, N., Ripatti, P., Partanen, J.J., Aro,
11 J., Mustonen, P., Tuomi, T., et al. (2020). Communicating polygenic and non-genetic risk
12 for atherosclerotic cardiovascular disease - An observational follow-up study. medRxiv,
13 2020.2009.2018.20197137. 10.1101/2020.09.18.20197137.
- 14 43. Abraham, G., Malik, R., Yonova-Doing, E., Salim, A., Wang, T., Danesh, J., Butterworth,
15 A.S., Howson, J.M.M., Inouye, M., and Dichgans, M. (2019). Genomic risk score offers
16 predictive performance comparable to clinical risk factors for ischaemic stroke. *Nat*
17 *Commun* *10*, 5819. 10.1038/s41467-019-13848-1.
- 18 44. Malik, R., Chauhan, G., Traylor, M., Sargurupremraj, M., Okada, Y., Mishra, A., Rutten-
19 Jacobs, L., Giese, A.K., van der Laan, S.W., Gretarsdottir, S., et al. (2018). Multiancestry
20 genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke
21 and stroke subtypes. *Nat. Genet.* *50*, 524-537. 10.1038/s41588-018-0058-3.
- 22 45. Drews, C.D., and Greeland, S. (1990). The impact of differential recall on the results of
23 case-control studies. *Int. J. Epidemiol.* *19*, 1107-1112. 10.1093/ije/19.4.1107.
- 24 46. Do, C.B., Hinds, D.A., Francke, U., and Eriksson, N. (2012). Comparison of family history
25 and SNPs for predicting risk of complex disease. *PLoS Genet.* *8*, e1002973.
26 10.1371/journal.pgen.1002973.
- 27 47. Carmi, S. (2021). Cascade screening following a polygenic risk score test: what is the
28 disease risk of a sibling conditional on a high score in a proband? bioRxiv,
29 2021.2004.2011.439329. 10.1101/2021.04.11.439329.
- 30 48. Reid, N.J., Brockman, D.G., Elisabeth Leonard, C., Pelletier, R., and Khera, A.V. (2021).
31 Concordance of a High Polygenic Score Among Relatives: Implications for Genetic
32 Counseling and Cascade Screening. *Circ Genom Precis Med* *14*, e003262.
33 10.1161/CIRCGEN.120.003262.
- 34 49. Martin, A.R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B.M., and Daly, M.J. (2019).
35 Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.*
36 *51*, 584-591. 10.1038/s41588-019-0379-x.
- 37

1 **Figure 1. Disease prevalence and prevalence and effect sizes of positive family history. Panel A:** disease prevalence in individuals for whom we studied risk of
 2 first-degree family history **Panel B:** Prevalence of first-degree family history (left column), second-degree family history (middle column), and parental cause of
 3 death (right column). **Panel C:** Effect size of first-degree family history (left column), second-degree family history (middle column), and parental cause of death
 4 (right column) with respective diseases. For parental causes of death, we studied 15 out of the 24 diseases.



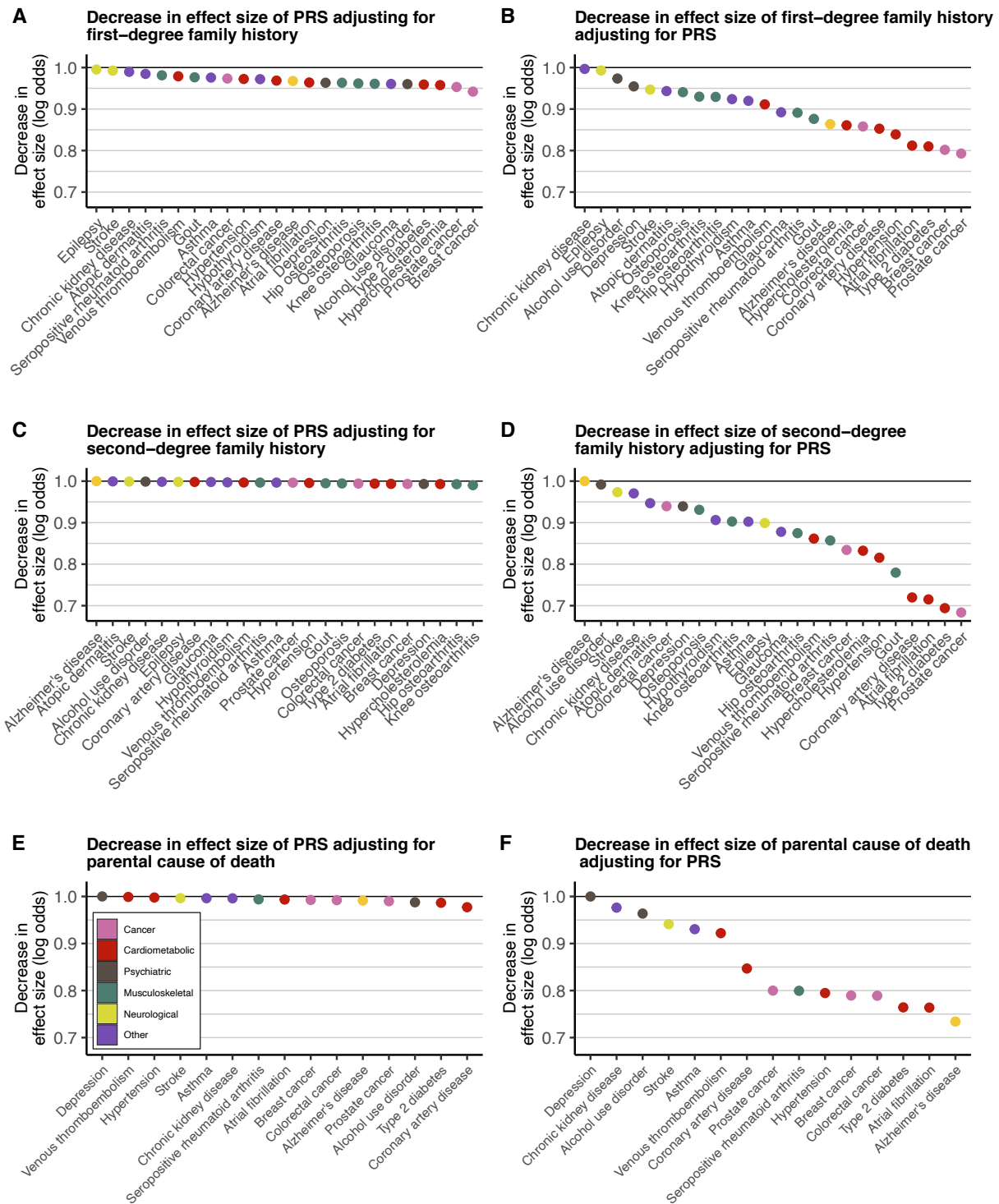
5 Sample size in panel A: total N = 39,444, N = 15,281 for breast cancer, N = 9,473 for prostate cancer. Sample sizes in panels B-C: first-degree family history as in panel A, second-degree family history total N = 47,154,
 6 N = 18,973 for breast cancer, N = 12,355 for prostate cancer, and parental causes total N = 227,982, N = 133,653 for breast cancer; N = 94,329 for prostate cancer. Odds ratios (OR) were obtained from logistic regression
 7 models adjusted for sex (except for breast and prostate cancer), birth year, genotyping array, cohort, and the first ten genetic principal components of ancestry.
 8

1 **Figure 2.** Cross-adjustment effects for first-degree family history (FH_{1st}), and respective polygenic risk scores (PRS). The impact of adjusting the PRS effect with first-
 2 degree FH_{1st} (**panel A**) and vice versa (**panel B**). The diamonds represent the unadjusted effects and the squares the adjusted effects. The PRS explained on average
 3 10% of the effect of FH_{1st} , but FH_{1st} only 3% of the PRSs. The PRS effect is shown per one SD increase.
 4



5
 6
 7 Total N = 39,444, N = 15,281 for breast cancer, N = 9,473 for prostate cancer. Odds ratios (OR) were obtained from logistic regression models adjusted for sex (except for breast and prostate cancer), birth year,
 8 genotyping array, cohort, and the first ten genetic principal components of ancestry.

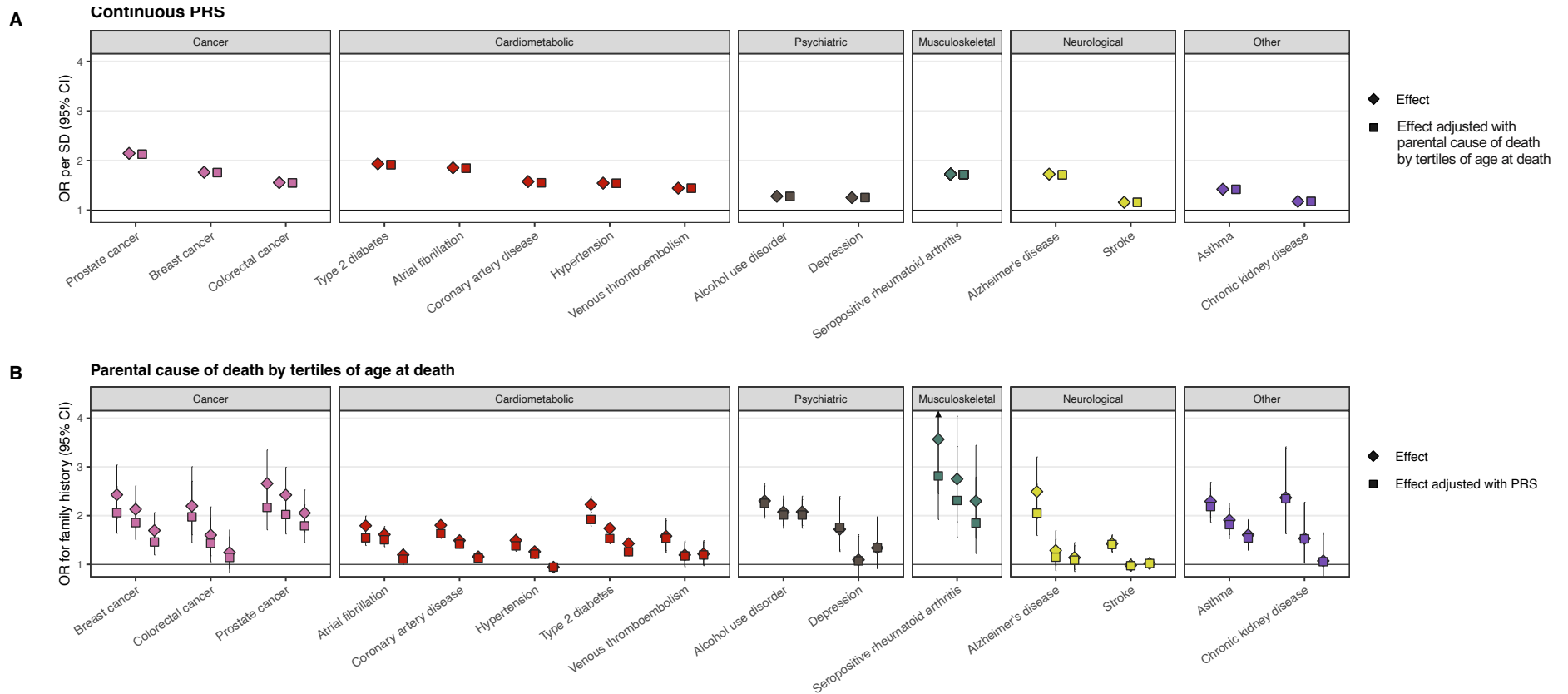
1 **Figure 3.** Proportional decreases in log odds by disease for first-degree family history FH_{1st} (panels **A-B**), for
 2 second-degree family history FH_{2nd} (panels **C-D**), and parental causes of death FH_P (panels **E-F**). The left column
 3 (panels **A, C, E**) represents decreases in effect size of high polygenic risk score (PRS; defined as top 10% of the
 4 distribution) adjusting for family history. The right column (panels **B, D, F**) represents decreases in effect size of
 5 family history adjusting for high PRS. The y axis represents the decrease in the effect size, calculated by dividing
 6 the log odds from the adjusted logistic regression model with the log odds from the non-adjusted model. For
 7 instance, in panel **A**, the y axis represents the following quantity: (log odds of PRS adjusting for FH_{1st}) / (log odds
 8 of PRS without adjusting for FH_{1st}).



9
10
11

The reference category for PRS was the rest of the distribution (individuals below the 90th percentile). In panel **D**, the proportion of Alzheimer's disease was set at 1.00 as we did not observe any association for second-degree family history of Alzheimer's disease.

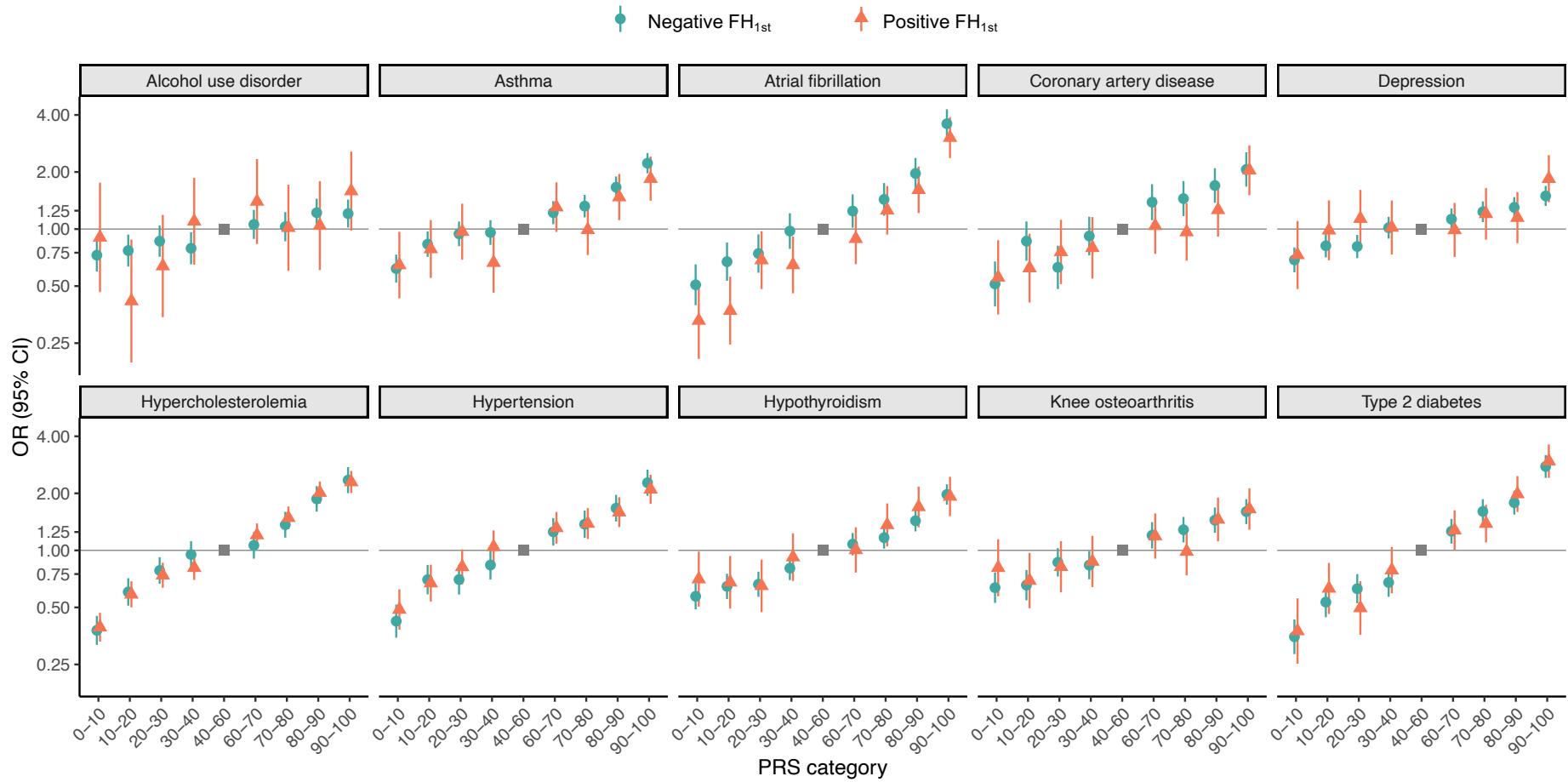
1 **Figure 4.** Polygenic risk is independent of both early- and late-onset family history. As early-onset family history is considered a particularly important familial risk
 2 factor, we also assessed the impact of FH_p divided into tertiles of age at death. **Panel A:** Adjusting the effect of polygenic risk score (PRS; per SD) by parental causes
 3 of death (FH_p) divided into tertiles of age at death had no impact on the effect sizes of the PRSs. **Panel B:** Adjusting the effects FH_p by tertiles of age at death by PRS
 4 resulted in the largest effect size decreases for the youngest age tertile, however, for most diseases the difference by age tertile was small. The diamonds represent the
 5 unadjusted effects and the squares the adjusted effects. In panel B, the effect sizes from with lowest to highest age at death are displayed from left to right, and the
 6 reference group for each disease is individuals with negative FH_p.



8
 9
 10 Sample size: total N = 227,982; N = 133,653 for breast cancer; N = 94,329 for prostate cancer. Odds ratios (OR) were obtained from logistic regression models adjusted for sex (except for breast
 11 and prostate cancer), birth year, genotyping array, cohort, and the first ten genetic principal components of ancestry. Age limits for tertiles of FH_p, and the number of individuals with parental
 12 cause of death in each tertile are reported in Supplementary Table 7.

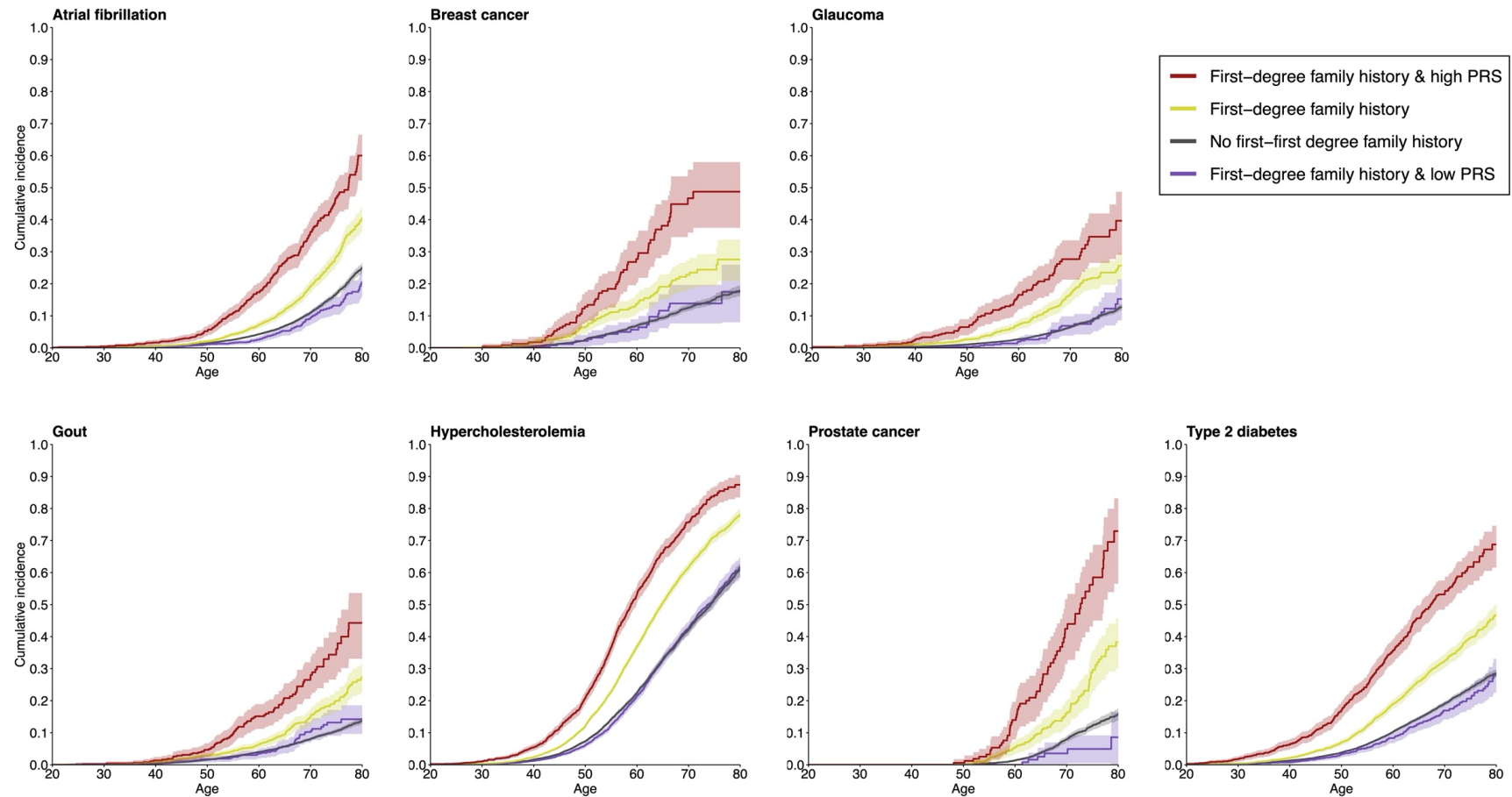
1
2
3

Figure 5. Effect sizes of polygenic risk scores stratified by first-degree family history (FH_{1st}), assessed for the 10 most prevalent diseases.



4

1 **Figure 6.** The impact of polygenic risk on disease risk in individuals with positive family history. The survival curves show cumulative incidences for individuals with
 2 positive first-degree family (FH_{1st}), stratified by level of polygenic risk score (PRS). High PRS was defined as top decile of the PRS distribution and low PRS as the
 3 bottom tertile of the PRS distribution. The figure shows results for the five diseases with the largest effect sizes for PRS, and for breast and prostate cancer. Survival
 4 curves for a broader set of diseases, and survival curves stratifying individuals with no FH_{1st} into similar PRS groups are in **Supplementary Figures 5–6.**
 5



6
 7
 8 Total N = 39,444, N = 15,281 for breast cancer, N = 9,473 for prostate cancer. Analyses were performed for diseases with an OR >2 for high PRS in Figure 2 (15/24 diseases), and over 10 cases
 9 in each subgroup, excluding Alzheimer's disease due to its average onset late in life.