Polygenic scores in disease prediction: evaluation using the relevant performance metrics

Hingorani AD.[1,2,3,4], Gratton J.[1, 2], Finan C.[1,2,3,4], Schmidt AF.[1,2,3,4,5], Patel R[1,2,3,4]., Sofat R.[3,4,6], Kuan V.[1,2,3], Langenberg C[7]., Hemingway H.[2,3,4,8], Morris JK.[9], Wald NJ.[8,9]

[1] Institute of Cardiovascular Science, University College London, UK.

[2] UCL British Heart Foundation Research Accelerator.

[3] UCL National Institute of Health Research Biomedical Research Centre.

[4] Health Data Research, UK.

[5] University Medical Centre, Utrecht, Netherlands.

[6] Department of Pharmacology and Therapeutics, Institute of Systems, Molecular and Integrative Biology, University of Liverpool.

[7] Berlin Institute of Health @Charite, Berlin, Germany.

[8] Institute of Health Informatics, University College London, UK.

[9] Population Health Research Institute, St. George's, University of London.

Abstract

We examine the performance of polygenic scores in screening and disease prediction using metrics which are firmly established for non-genetic tests but rarely reported in evaluations of polygenic scores. Using performance metrics in the Polygenic Score Catalog (odds and hazard ratios, area under the receiver operating characteristic curve or C-index), we calculate the sensitivity or detection rate (DR_5) for polygenic score cut-offs that define a 5% false positive rate (FPR). To examine clinical and public health relevance, we use information on disease incidence to calculate the odds of being affected given a positive result ($OAPR$; the ratio of true to false positives) both for an individual and a population. Polygenic scores typically detect $7.6 - 16.4\%$ (and therefore miss $83.6 - 92.4\%$) of affected individuals with a 5% FPR. For a polygenic score for coronary artery disease (CAD) with a DR_5 of 12% (88% of cases missed), the $OAPR$ is $1:3.75$ if used in a population with an average 10-year CAD risk of 10%, and $1:41$ in a population with an average 10-year CAD risk of 1%. For a polygenic score for breast cancer also with a DR_5 of 12% and a population average 10-year risk of 1 in 65 (odds =$1:64$), odds are reduced to $1:91$ for a woman with a polygenic score at the 25th centile and increased to $1:53$ for a woman with a result at the 75th centile. Analysis using the relevant metrics reveals the weak predictive performance of polygenic risk scores, which limits their effectiveness in screening and disease prevention.

Introduction

There is widespread interest in the use of polygenic scores in the estimation of disease risk[1]. The public can already access this type of information by purchasing DNA testing kits directly from commercial providers[2]. Recently, Lambert and colleagues produced the PGS Catalog, a publicly accessible directory of studies on polygenic scores[3]. There have also been position papers on reporting standards for such studies from the Clinical Genome Resource (ClinGen) Complex Disease Working Group together with the PGS Catalog[4], and on the responsible clinical use of polygenic scores from the Polygenic Risk Score Task Force of the International Common Disease Alliance[5]. However, the performance of polygenic scores in disease prediction has been debated and their eventual place in medicine and public health is uncertain[6–9]. Here we use data curated in the PGS Catalog to examine the performance of polygenic scores in screening and disease prediction.

To do so, we convert the originally reported metrics for all scores into the common established measures of screening performance: the detection rate (DR; or sensitivity) for a pre-specified false positive rate (FPR). This enables fair comparisons between different

polygenic scores, as well as between polygenic scores and non-genetic tests such as conventional cardiovascular disease (CVD) risk factors or mammography. Using CVD and breast cancer as examples, we illustrate how these measures, derived from the originally reported metrics, can be combined with estimates of disease incidence to estimate screening performance in a population and assess the implication of a particular polygenic score value for an individual [10].

Methods

Each score in the PGS Catalog is associated with one or more 'performance metrics'. For disease prediction, recorded metrics include hazard (HR) or odds ratios (OR), both per 1-standard deviation (SD) increment in score, or the area under the receiver operating characteristic curve (AUC), sometimes expressed as the C-index. However, the well-established performance metrics for screening tests or disease prediction are missing, possibly because they were rarely or inconsistently reported by the original publications. The missing metrics are the detection rate (DR; or sensitivity) and false positive rate (FPR) which are, respectively, the proportion of people with positive tests among those affected and unaffected by disease. Also missing is the positive predictive value (PPV), which is the proportion people affected among those with positive tests, more usefully expressed as the odds of being affected given a positive result ($OAPR$): the ratio of true to false positives[11].

Since polygenic risk scores have a Gaussian distribution, we use the reported metrics to derive the difference in mean values between affected and unaffected individuals (standardised using the Z-score) and to assess the overlap in the distribution of score values. We use this overlap to calculate the DR for any given FPR or *vice versa.* For simplicity and consistency, we set polygenic score cut-offs that define a 5% FPR and calculate the corresponding detection rate (DR_5) [10,12]. The assumptions, methods and calculations are described in more detail in the **Extended methods** and **Supplementary Tables**.

Results

Overview of the performance of polygenic scores

At the time of publication in March 2021, the PGS Catalog had curated 1835 performance metrics for 749 polygenic scores, involving 197 diseases or traits, reported in 154 publications (**Figure 1**). Of the 1835 metrics, 1610 (87.7%) concerned disease end points,

reported as OR per-SD in 617, HR per-SD in 221, AUC in 625 and C-index in 147 instances (**Supplementary Tables**). The median DR_5 values [interquartile range: IQR] for polygenic scores whose performance was reported using HR or OR per-SD were 7.6% $[6.6 - 9.3]$ and 9.7% $[7.6 - 12.1]$ respectively, excluding 51 instances where the HR or OR per-SD was recorded as < 1. For polygenic risk score performance reported using AUC or C-index, the corresponding DR_5 values were 10.5% $[8.3 - 14.2]$ and 16.4% $[10.8 - 23.1]$ respectively (**Figure 2**).

This means that at a 5% false positive rate, polygenic risk scores typically detect $7.6 - 16.4\%$ (and therefore miss $83.6 - 92.4\%$) of affected individuals. For a condition, for example CVD, with a 10% incidence in a middle-aged population in a given period (odds $= 1:9$) and a polygenic score with $DR_5 = 12\%$ (88% of cases missed), the $OAPR = (0.12 \times 1):(0.05 \times 9) = 1:3.75$, i.e., false positives outnumber true positives by nearly four to one. Changing the cut-off to reduce the FPR (e.g., to 1%) also reduces the DR (to 3% in this example, with 97% of cases missed), and the $OAPR$ is increased to $1:3$. Retaining a 5% FPR but applying the test in a population with 1% CVD incidence over the same period (odds $= 1:99$), e.g., in individuals at a younger age, yields an $OAPR$ of $1:41$.

Achieving more effective discrimination requires much greater separation of polygenic score distributions of affected and unaffected individuals than is observed in practice. For instance, achieving an 80% detection rate for a 5% false positive rate ($DR_5 = 80\%$) requires an OR per-SD of 12 (compared to the median observed value of 1.41) or an AUC of 0.96 (compared to the median observed value of 0.61) (**Figure 3**). Only 8.2% of AUC values in the PGS Catalog exceeded 0.8 which equates to a DR_5 of 32%, with most of these reflecting large effect variants at the HLA locus in a few autoimmune diseases.

Comparison of polygenic scores with non-genetic screening tests

This performance falls well short of that of established screening tests. For example, the UK National Screening Committee[13] has approved population screening for breast, cervical and colon cancer but not prostate cancer or type 2 diabetes. In each case, a relevant polygenic score from the PGS Catalog would detect fewer cases at comparable false positive rates than the established non-genetic tests. Polygenic risk scores for coronary artery disease (CAD) perform about as well as conventional risk factors (e.g. blood pressure and cholesterol) which are individually poor predictors[14] even when combined with other variables (**Table**).

Adding polygenic scores to prediction based on non-genetic risk factors

Nevertheless, it has been proposed that the addition of polygenic scores could improve cardiovascular (CVD) prediction based on conventional risk factors. For example, Sun *et al*.[15] estimated the effect of adding polygenic scores for CAD (PGS Catalog reference 000018) and stroke (PGS 000039) to conventional risk factors in the prediction of incident CVD events. Using data from primary care records, they constructed a hypothetical cohort of 100,000 people aged 40-75 years with representative risk factor profiles for people in England. They assumed a 10-year CVD risk cut-off of 10% for providing statin treatment[16] and that statins reduce the risk of a CVD event by 20%. It can be shown from data presented in their paper that the average 10-year CVD risk in this population is 8% (odds= $1{:}\,12.5$); that using conventional risk factors (including age) detects 60% of cases at a 24% FPR ($DR_{24} = 60\%$); and that adding the polygenic risk scores detected 61% of cases for a 23% FPR ($DR_{23} = 61\%$) (**Figure 4**). Given the study population was aged 40 and over, with a baseline risk of 8%, an alternative strategy, in line with current US guidance, would be to offer statins to all $100\,000$ regardless of polygenic score or conventional risk factors. Using the same assumptions, this would prevent 1600 cardiovascular events instead of the 957 they estimated would be prevented using conventional risk factor prediction alone, 974 using conventional risk factors and polygenic scores together, or 1029 using a hybrid model, where conventional risk factor assessment is followed by polygenic scores only for those at intermediate (5-10%) risk (**Figure 4**). Since age is a major determinant of CVD risk, age alone performs about as well as risk models that include age[17]. Given the rarity of cardiovascular events below 50, using an age cut-off of 50 instead of 40 is likely to prevent almost as many events with fewer false positives[18].

Interpretation of a polygenic score in an individual

Information on disease incidence can be used to estimate the odds of disease for an individual based on their position in a polygenic score distribution[10], and to indicate the absolute numbers of true and false positives in a given population (see **Extended methods**).

For example, the population average 10-year breast cancer risk in a woman aged 40 is estimated as 1 in 65 (odds = $1{:}\,64$). A breast cancer polygenic risk score (PGS000004) has been reported with an OR per-SD of 1.61, corresponding to a DR_5 of 12% (**Table**). For women at age 40 with polygenic risk scores between the 20th and 80th centiles, the addition of genetic information produces only a modest adjustment in risk (**Figure 5**). For example,

for a 40-year-old woman with a polygenic score at the 25th centile, the 10-year odds of breast cancer are reduced from $1:64$ to $1:91$, while for a woman with a result at the 75th centile the odds are increased to $1:53$. For women with a polygenic score in the tails of the distribution, the change in odds is greater: odds are reduced from $1:64$ to $1:160$ at the 2.5^{th} centile and increased to $1:28$ at the 97.5^{th} centile, which is higher than the average 10-year odds of breast cancer ($1:42$) at age 50, the UK age threshold for screening by mammography. Offering mammography in their 40th year to all $397\ 469$ women (UK Office of National Statistics 2018 estimate) based on a breast cancer polygenic score cut-off at or above the 97.5^{th} centile among the unaffected would lead to $10\ 206$ women being offered mammography based on a DR of 7% for a 2.5% FPR. A two-stage screen (where screen positives based on the polygenic score are then offered a mammogram) is estimated to detect 316 cancers, and miss 5798, with 783 false positives, based on a reported DR_8 of 75% for mammography[19] (**Figure 5**).

Discussion

The PGS Catalog provides a comprehensive listing of studies and polygenic scores, but the reported performance metrics limit comparisons among polygenic scores, and between polygenic scores and non-genetic tests, making it difficult to judge their place in disease prediction. We exploited the Gaussian distribution of polygenic scores and the mathematical relationships to generate a common, relevant predictive performance metric for each score – the DR_5 value. We were able to infer a substantial overlap of polygenic risk score distributions among affected and unaffected individuals for almost all conditions, yielding DR_5 values that are indicative of weak predictive performance. This was the case even for conditions such as CVD where some have argued polygenic scores are ready to incorporate in practice[20].

The findings may seem surprising given the level of enthusiasm in polygenic scores, but this may reflect the presentation of results in research papers and company materials. What is relevant in screening is the risk of an event in a group compared to that of the whole population, but publications often depict comparisons between mutually exclusive groups, e.g., those in opposite tails of the polygenic score distribution[21]. As shown in **Figure 3**, seemingly impressive odds ratios of 34, 15, 10 or 6 in comparisons of the top vs. the bottom 1%, 5%, 10% or 20% of the polygenic score distribution, respectively, correspond to substantial overlap of score distributions for affected and unaffected individuals, and equate, in the example shown to a DR_5 of only 16%.

Because polygenic risk scores, like many non-genetic risk factors[22], display a Gaussian distribution in the population[23] and a relatively weak log-linear association with disease risk, more cases occur among those at average than those at extreme risk[24]. This means that where there are safe, inexpensive preventative interventions (e.g., statins for CVD) there is greater public health benefit in extending rather than limiting their access[25] and this has been achieved by the progressive lowering of the risk threshold for statin treatment over time. Many of those at high polygenic risk are therefore already likely to be eligible for statin treatment. Although CVD risk in individuals with high polygenic scores could be as great as that of individuals with a monogenic condition (e.g., autosomal dominant familial hypercholesterolaemia; FH), identifying those with a monogenic condition enables cascade testing of family members, which is not relevant in families of those at high polygenic risk.

Ascertaining a minority of individuals at very high genetic risk may be justified if the preventative intervention is costly, resource limited, or potentially harmful. However, screening to identify those at very high risk entails genetic testing in all and, aside from missing the many cases among those at average risk, generates many false positives. This could have substantial downstream resource implications for healthcare systems if genetic testing is to be followed by a more definitive test., e.g., early mammography in the case of breast cancer.

Using polygenic scores for disease prediction[26] has the appeal of the low cost of genotyping, and the invariant nature of an individual's genotype, which means testing need be done only once to compute the risk of a wide range of diseases. However, the appeal is diminished by the weak predictive performance. This may be because causal risk factors (though important in unveiling targets for pharmaceutical intervention) are less able to differentiate affected individuals than those markers that change as a consequence of the disease itself [8,27]. Thus, analytes measured using different -omics technologies, such as proteomics and metabolomics, which can be perturbed by the presence of early disease, may emerge as better predictors.

The current analysis underlines the need for more pertinent presentation of the performance of polygenic risk scores for disease prediction. We suggest the following: (1) the PGS Catalog and other resources that curate such information derive the DR for a specified FPR as we show, and include it as a standard curated performance metric; (2) primary studies always report the mean and SD of polygenic score values among affected and unaffected individuals, from which the overlap in distributions and the relevant performance metrics can then be derived[12]; (3) authors report the performance of polygenic scores with and without the inclusion of other variables (particularly age and sex) which can markedly influence

predictive performance so that users can judge the increment provided by the polygenic score itself; (4) commercial providers communicate individual test results to customers with greater clarity and relevance to screening performance, e.g., by presenting the $OAPR$, which requires additional information on population average risk at a particular age over a specified time; and (5) as others have already suggested[28], policy makers should consider tighter regulation of commercial polygenic scores, based on clinical not just analytical performance, to protect already stretched public health systems from being burdened by management of false positive results.

The limitations of polygenic scores in disease prediction does not preclude their use in other applications. For example, they may inform the variable penetrance of rare mutations in monogenic diseases, e.g., hypertrophic cardiomyopathy. There are also other predictive applications of genotyping, e.g., in pharmacogenetic testing, to optimise efficacy and safety of medicines. Genotyping may also be of value in blood and tissue matching. However, perhaps counterintuitively, the main healthcare benefit of common disease genomics may come from understanding the causes of disease and drug target discovery[29] rather than disease prediction.

Acknowledgements

Supported by the UCL British Heart Foundation Accelerator (AA/18/6/34223), the UCL NIHR Biomedical Research Centre, and the UKRI/NIHR funded Multimorbidity Mechanism and Therapeutics Research Collaborative (MR/V033867/1). ADH is an NIHR Senior Investigator.

Ethics declaration

ADH is a member of the NIHR Biomedical Research Centres Polygenic Risk Score Collaborative and is a Co-Investigator on a grant from Pfizer to identify potential therapeutic targets for heart failure using human genomics.

References

1. Genome UK: the future of healthcare - GOV.UK. https://www.gov.uk/government/publications/genome-uk-the-future-of-healthcare.
2. Horton, R. *et al.* Direct-to-consumer genetic testing. *BMJ* (2019) doi:10.1136/bmj.l5688.
3. Lambert, S. A. *et al.* The Polygenic Score Catalog as an open database for reproducibility and systematic evaluation. *Nature Genetics* (2021) doi:10.1038/s41588-021-00783-5.
4. Wand, H. *et al.* Improving reporting standards for polygenic scores in risk prediction studies. *Nature* **591**, 211–219 (2021).
5. Responsible use of polygenic risk scores in the clinic: potential benefits, risks and gaps. *Nat. Med.* **27**, 1876–1884 (2021).
6. Torkamani, A., Wineinger, N. E. & Topol, E. J. The personal and clinical utility of polygenic risk scores. *Nat. Rev. Genet.* (2018) doi:10.1038/s41576-018-0018-x.
7. Lambert, S. A., Abraham, G. & Inouye, M. Towards clinical utility of polygenic risk scores. *Human Molecular Genetics* (2019) doi:10.1093/hmg/ddz187.
8. Wald, N. J. & Old, R. The illusion of polygenic disease risk prediction. *Genetics in Medicine* (2019) doi:10.1038/s41436-018-0418-5.
9. Sud, A., Turnbull, C. & Houlston, R. Will polygenic risk scores for cancer ever be clinically useful? *npj Precision Oncology* (2021) doi:10.1038/s41698-021-00176-1.
10. Wald, N. & Leck, I. *Antenatal and neonatal screening*. *Antenatal and Neonatal Screening* (2009). doi:10.1093/acprof:oso/9780192628268.001.0001.
11. Wald, N. J. Guidance on terminology. *Journal of Medical Screening* (2008) doi:10.1258/jms.2008.008got.
12. Wald, N. J. & Morris, J. K. Assessing risk factors as potential screening tests: A simple assessment tool. *Arch. Intern. Med.* (2011) doi:10.1001/archinternmed.2010.378.
13. UK National Screening Committee - GOV.UK. https://www.gov.uk/government/organisations/uk-national-screening-committee.
14. Wald, N. *et al.* Apolipoproteins and ischaemic heart disease: implications for screening. *Lancet* (1994) doi:10.1016/S0140-6736(94)90814-1.
15. Sun, L. *et al.* Polygenic risk scores in cardiovascular risk prediction: A cohort study and modelling analyses. *PLoS Med.* (2021) doi:10.1371/JOURNAL.PMED.1003498.
16. Overview | Cardiovascular disease: risk assessment and reduction, including lipid modification | Guidance | NICE. https://www.nice.org.uk/guidance/cg181.
17. Wald, N. J., Simmonds, M. & Morris, J. K. Screening for future cardiovascular disease using age alone compared with multiple risk factors and age. *PLoS One* (2011)

doi:10.1371/journal.pone.0018742.

18. Wald, N. J. & Law, M. R. A strategy to reduce cardiovascular disease by more than 80%. *Br. Med. J.* (2003) doi:10.1136/bmj.326.7404.1419.
19. Elmore, J. G., Armstrong, K., Lehman, C. D. & Fletcher, S. W. Screening for breast cancer. *Journal of the American Medical Association* (2005) doi:10.1001/jama.293.10.1245.
20. Knowles, J. W. & Ashley, E. A. Cardiovascular disease: The rise of the genetic risk score. *PLoS Med.* (2018) doi:10.1371/journal.pmed.1002546.
21. Polygenic Risk Scores - Genomics plc. https://www.genomicsplc.com/resources/polygenic-risk-scores/.
22. Law, M. R. & Wald, N. J. Risk factor thresholds: Their existence under scrutiny. *British Medical Journal* (2002) doi:10.1136/bmj.324.7353.1570.
23. Pain, O., Gillett, A. C., Austin, J. C., Folkersen, L. & Lewis, C. M. A Tool for Translating Polygenic Scores onto the Absolute Scale Using Summary Statistics. *medRxiv* 2021.04.16.21255481 (2021) doi:10.1101/2021.04.16.21255481.
24. Rose, G. Sick individuals and sick populations. *Int. J. Epidemiol.* (1985) doi:10.1093/ije/14.1.32.
25. Hingorani, A. D. & Hemingway, H. How should we balance individual and population benefits of statins for preventing cardiovascular disease? *BMJ* (2011) doi:10.1136/bmj.c6244.
26. Holmes, M. V., Harrison, S., Talmud, P. J., Hingorani, A. D. & Humphries, S. E. Utility of genetic determinants of lipids and cardiovascular events in assessing risk. *Nat. Rev. Cardiol.* **8**, 207–221 (2011).
27. Wald, N. J., Hackshaw, A. K. & Frost, C. D. When can a risk factor be used as a worthwhile screening test? *BMJ* (1999) doi:10.1136/bmj.319.7224.1562.
28. Direct-to-consumer genomic testing: Government Response to the Committee's First Report - Science and Technology Committee - House of Commons. https://publications.parliament.uk/pa/cm5802/cmselect/cmsctech/805/80502.htm.
29. Schmidt, A. F., Hingorani, A. D. & Finan, C. Human Genomics and Drug Development. *Cold Spring Harb. Perspect. Med.* (2021) doi:10.1101/cshperspect.a039230.

			Non-genetic screening					Corresponding polygenic score		
Disease	**Test**	**Status of screening in UK**	**Positive result or cut-off**	**Reference**	**Detection rate (%)**	**False positive rate (%)**	**PGS Catalogue identifiers**	**Reported performance**	**Estimated detection rate**	**False positive rate§**
Breast cancer	Mammography	Offered to women from age 50 to their 71st birthday	Malignancy score	JAMA. 2005; 293:1245. Ann Intern Med. 2003; 138: 168.	75	8	PPM000004 PGS000004 PSS000004 PGP000002	OR per SD=1.61 AUC=0.63	12 12	5 5
Colon cancer	Faecal immunochemical test (FIT)	Offered to men and women aged 60-74 every 2 years	10µg/g	Gut 2015; 64:1327.	79	6	PPM00161 PGS000376 PSS000538 PGP000118	OR per SD=1.177 AUC=0.547	7 7	5 5
Cervical cancer	Cervical cytology (up to 2017), testing for high-risk human papilloma virus (HPV) thereafter	Offered to women aged 25-49 very 3 years and women aged 50-64 every 5 years.	PCR positive for high-risk HPV subtypes	JAMA 2002; 288: 1749	88	21	PPM000193 PGS000073 PSS000112 PGP000050	OR per SD =1.22	7	5
Type 2 diabetes	e.g., Leicester Risk Assessment Score	Screening for type 2 diabetes is not currently recommended by UK National Screening Committee†	≥25	Diabetic Medicine 2010; 27, 887.	29	11	PPM001605 PGS000713 PSS000754 PGP000128	AUC=0.69	17	5
Prostate cancer	Prostate specific antigen	Screening for prostate cancer using PSA is not currently recommended by UK National Screening Committee	4ng/mL	CA Cancer J Clin. 2010;60(2):70.	21	9	PPM001644 PGS000719 PSS000860 PGP000135	AUC=0.66	15	5
Coronary artery disease	Total cholesterol Diastolic blood pressure QRISK2	Not used in isolation Not used in isolation Recommended as part of NICE Guideline (CG 181); also used as part of the Vascular Health Check scheme (which is under review)	8mmol/L 105mmHg 20% estimated 10-year risk	BMJ 2003; 326:1419. BMJ 2008;336: a332. BMJ 2010:340:c2442. and rapid response	15 13 40 (26*)	5 5 13 (6*)	PPM000027 PGS000018 PSS000018 PGP000007	HR per SD=1.706 AUC=0.79 C-index=0.623	13 31 12	5 5 5

*Values are for men and in brackets for women. §Corresponds to a cut point value for a polygenic score at the 95th centile of the distribution among those unaffected. †The National Institute for Health and Care Excellence recommends validated self-assessment questionnaires or validated web-based tools for type 2 diabetes prevention in people at high risk.
PPM=PGS performance metric; PGS=Evaluated score; PSS=PGS sample set; PGP=publication ID

Table. Comparison of the performance of selected non-genetic tests and a corresponding polygenic risk score

Figure legends

Figure 1. Summary of data included in the PGS Catalog as of March 2021. For ease of illustration, two outlier publications were removed from the first histogram (one with 302 polygenic scores and the other with 221 polygenic scores).

Figure 2. Distribution of detection rates for a 5% false positive rate ($DR5$) derived from (a) HR per-SD; (b) OR per-SD; (c) AUC and (d) C-index values listed in the PGS Catalog as of March 2021.

Figure 3. Relationship between AUC or C-index, $DR5$ and OR for different proportions of the top and bottom tails of the polygenic risk score distribution.

Figure 4. Modelling of the effect of CVD prediction using (a) conventional risk factors alone and conventional risk factors with polygenic risk assessment in individuals at $5-10\%$ 10-year CVD risk; (b) conventional risk factors with polygenic risk in all; and (c) population-wide prevention without CVD prediction. The schema utilises estimates from Sun et al., PLoS Medicine 2021, who used age- and sex-specific incidence data from 2.1 million NHS primary care records translated to a hypothetical group of 100,000 adults aged 40-75 years. In scenarios (a) and (b) statins are assumed to be prescribed for those with a 10-year CVD risk ≥10% and in all scenarios assumed to produce a 20% relative risk reduction. Polygenic risk estimation utilised PGS 000018 for CAD (HR per-SD 1.31) and PGS 000039 for stroke (HR 1.19 per-SD).

Figure 5. Modelling the impact of a polygenic risk score for invasive breast cancer implemented in women on their 40th birthday, based on an estimated 10-year breast cancer risk of 1 in 65 (odds $1:64$). The analysis is based on the PGS Catalog breast cancer score PGS000004 with a reported OR per-SD of 1.61 (PPM000004) derived in dataset PSS000004. Panel (a) illustrates the change in odds for women with a score result corresponding to the 2.5th, 25th, 75th and 97.5th centiles of the polygenic score distribution (with reference to the distribution among unaffected). Panel (b) illustrates the number of anticipated invasive breast cancer cases detected and missed, the number of false positives, and the number of additional mammograms for a two-stage screening test using this polygenic score (stage 1) with a 97.5th centile cut-off, yielding a 7% detection rate for a 2.5% false positive rate and mammographic screening with a detection rate of 75% for an 8% false positive rate. Estimates are based on 397469 women aged 40 (Office of National

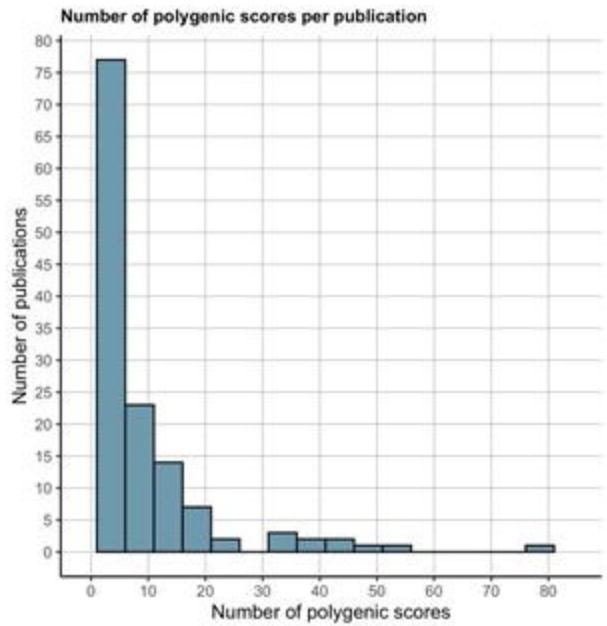
Statistics 2018). Performance of mammography in the detection of breast cancer uses estimates from JAMA. 2005; 293:1245

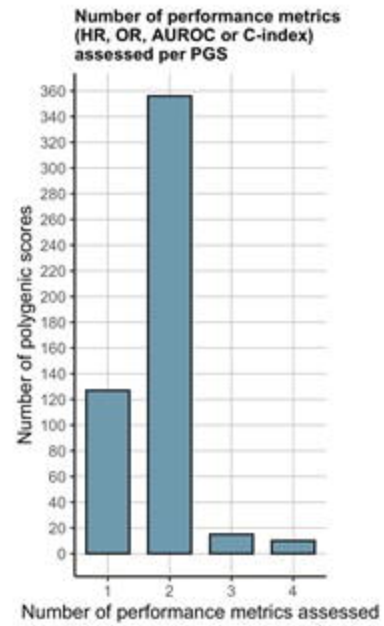
154 publications

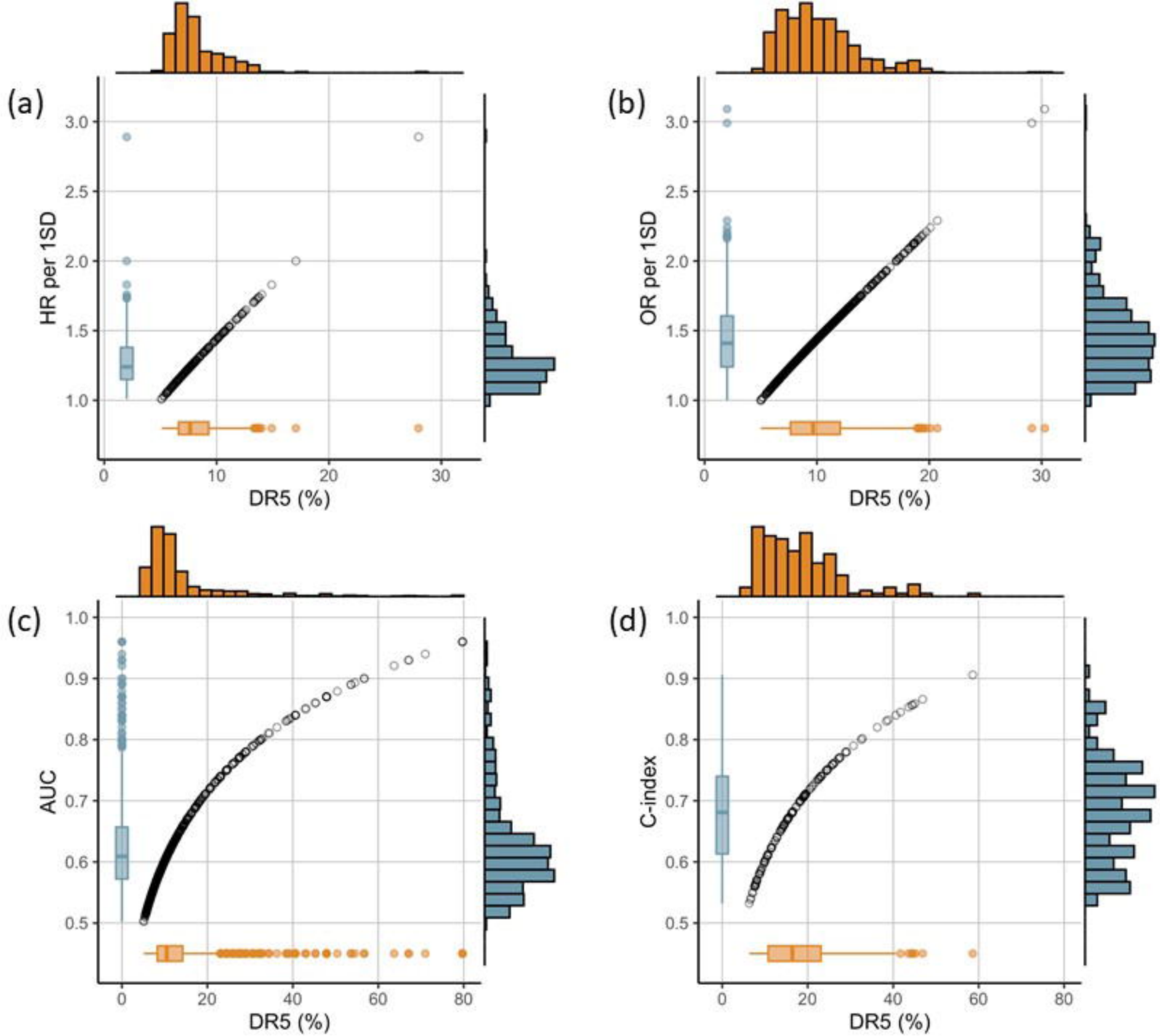
197 diseases or traits
(EFO terms)

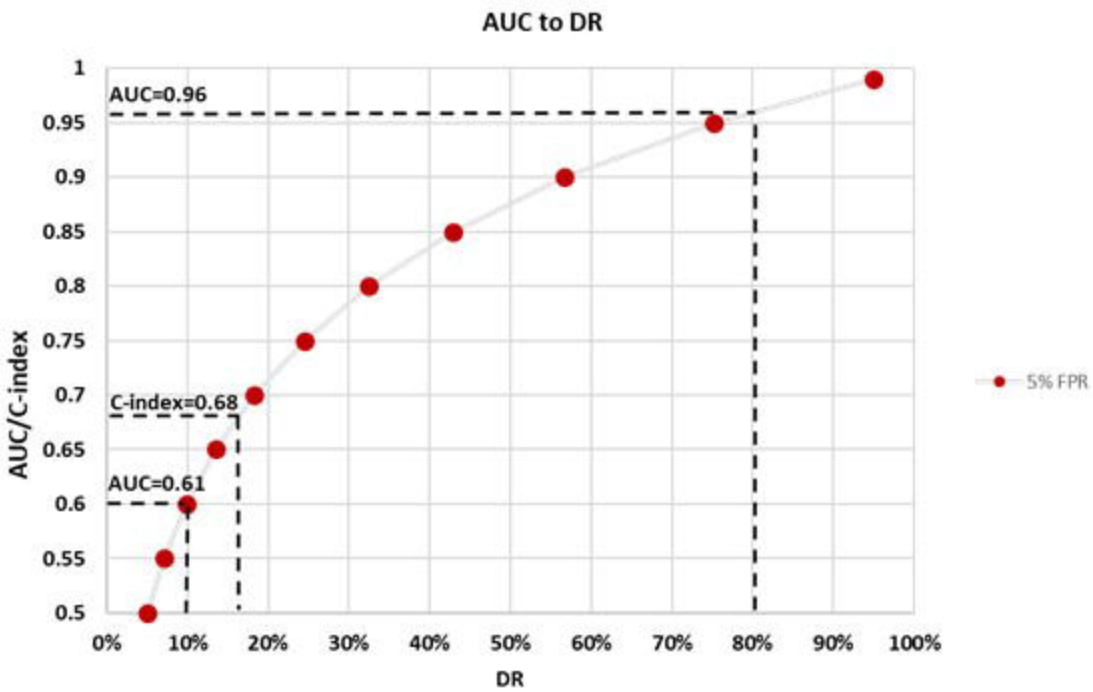
749 polygenic scores

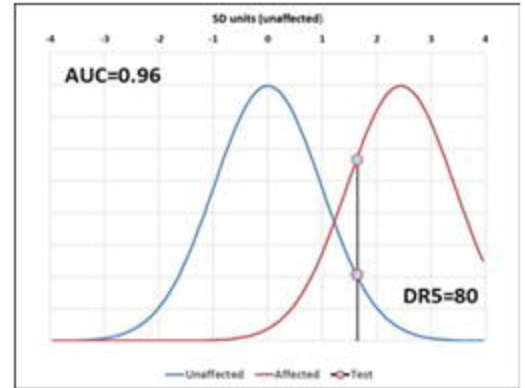
1835 performance metrics:
221 HR per SD, 617 OR per SD, 625 AUC,
147 C-index, 41 other



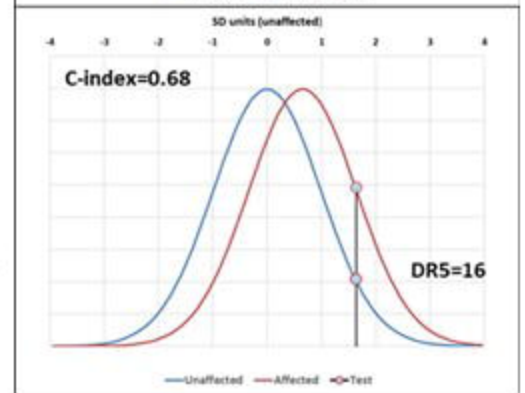



(a)
HR per 1SD
DR5 (%)
(b)
OR per 1SD
DR5 (%)
(c)
AUC
DR5 (%)
(d)
C-index
DR5 (%)

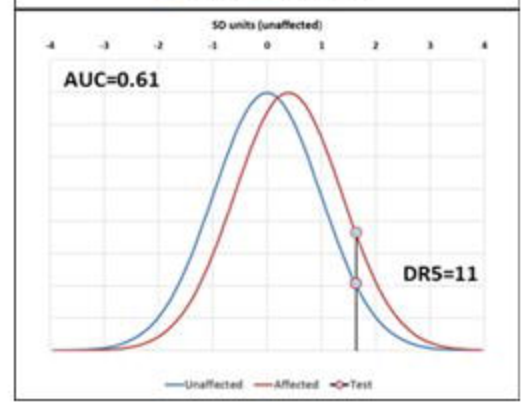




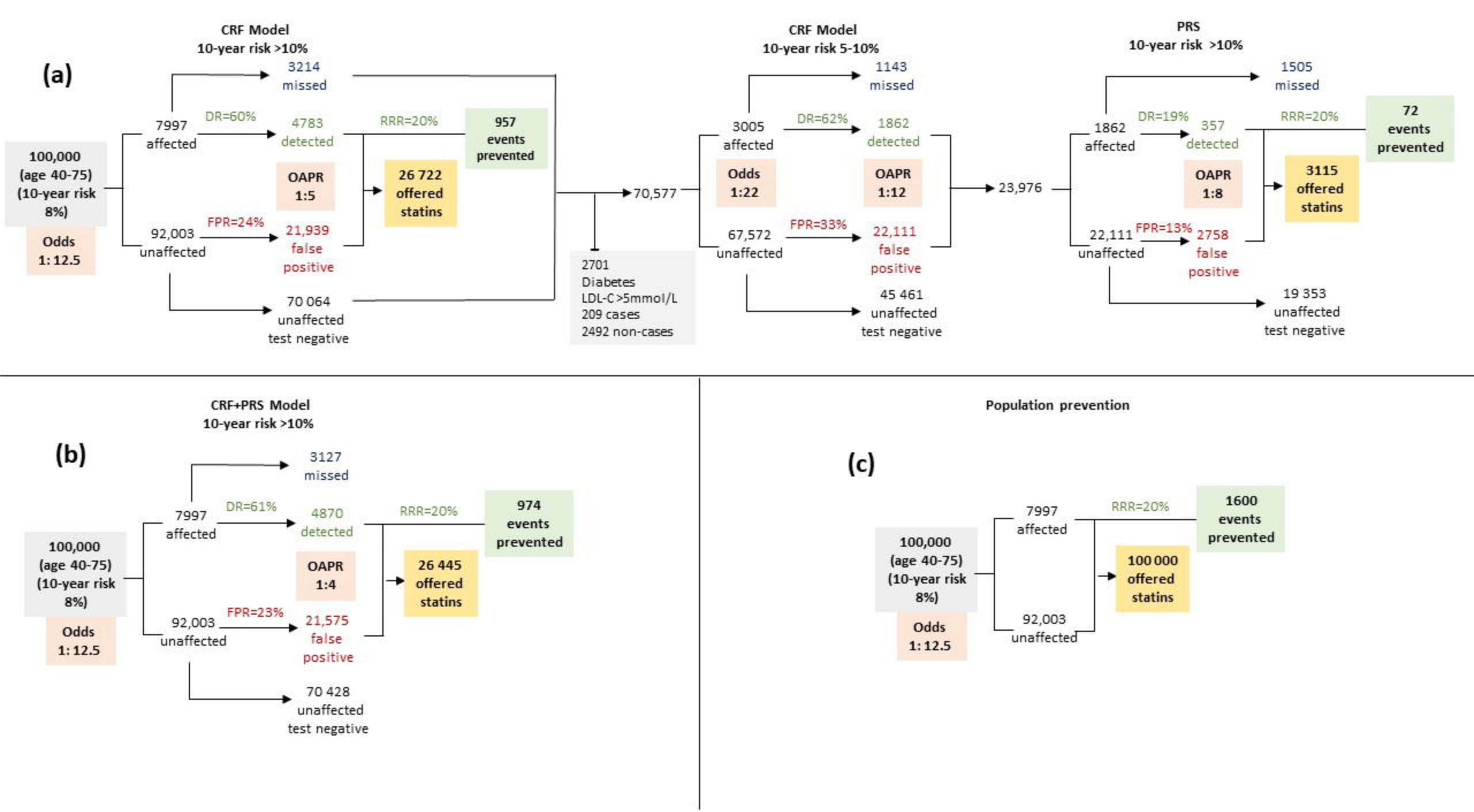
OR Top vs Bottom 1%	712916
OR Top vs Bottom 5%	42200
OR Top vs Bottom 10%	10295
OR Top vs Bottom 20%	2089
OR Top vs Bottom 25%	1182
OR Top vs Bottom 33%	536
Detection rate	0.80
False positive rate	0.05



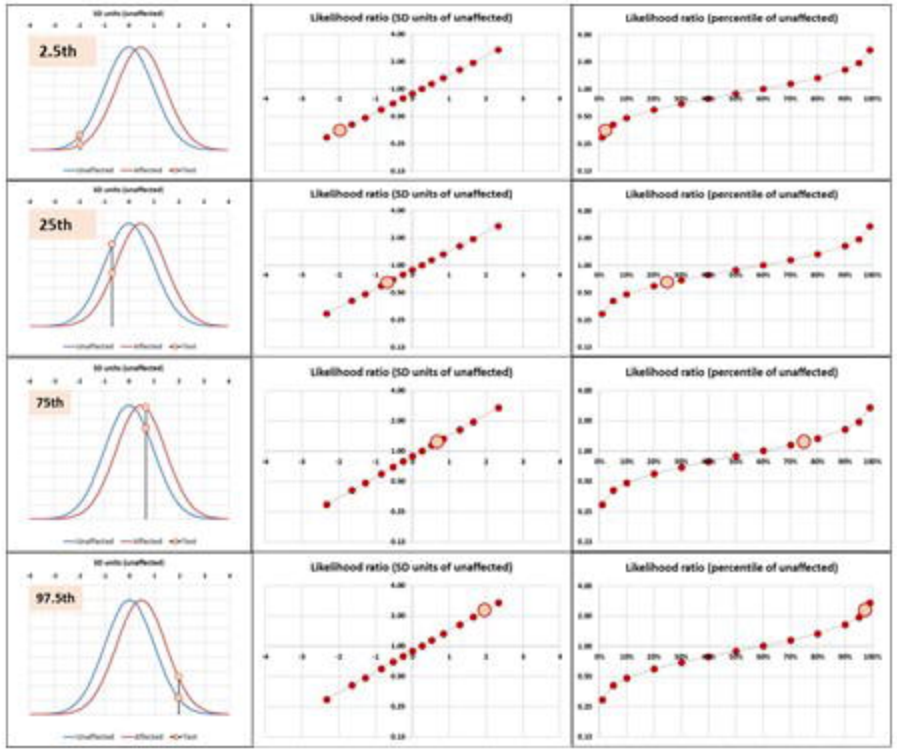
OR Top vs Bottom 1%	34
OR Top vs Bottom 5%	15
OR Top vs Bottom 10%	10
OR Top vs Bottom 20%	6
OR Top vs Bottom 25%	5
OR Top vs Bottom 33%	4
Detection rate	0.16
False positive rate	0.05

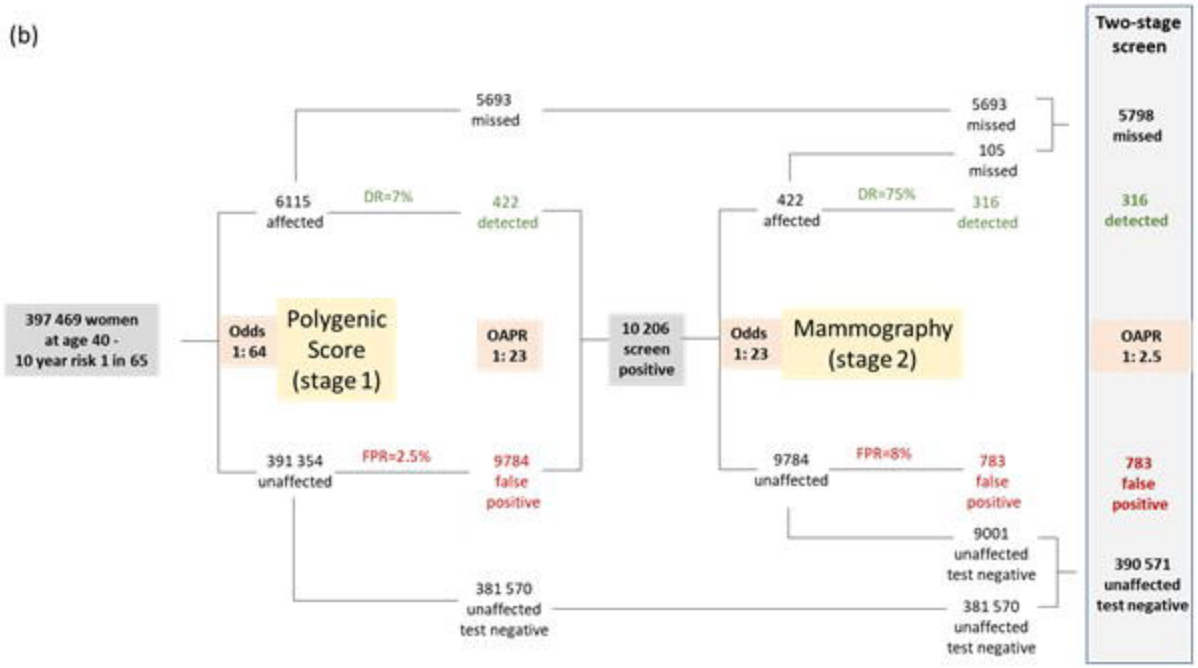


OR Top vs Bottom 1%	8
OR Top vs Bottom 5%	5
OR Top vs Bottom 10%	4
OR Top vs Bottom 20%	3
OR Top vs Bottom 25%	3
OR Top vs Bottom 33%	2
Detection rate	0.11
False positive rate	0.05


(a)
CRF Model
10-year risk >10%
3214
missed
7997
affected
DR=60%
4783
detected
RRR=20%
957
events
prevented
100,000
(age 40-75)
(10-year risk
8%)
Odds
1: 12.5
OAPR
1:5
26 722
offered
statins
92,003
unaffected
FPR=24%
21,939
false
positive
70 064
unaffected
test negative
70,577
2701
Diabetes
LDL-C >5mmol/L
209 cases
2492 non-cases
CRF Model
10-year risk 5-10%
1143
missed
3005
affected
DR=62%
1862
detected
Odds
1:22
OAPR
1:12
67,572
unaffected
FPR=33%
22,111
false
positive
45 461
unaffected
test negative
23,976
PRS
10-year risk >10%
1505
missed
1862
affected
DR=19%
357
detected
RRR=20%
72
events
prevented
OAPR
1:8
3115
offered
statins
22,111
unaffected
FPR=13%
2758
false
positive
19 353
unaffected
test negative
(b)
CRF+PRS Model
10-year risk >10%
3127
missed
7997
affected
DR=61%
4870
detected
RRR=20%
974
events
prevented
100,000
(age 40-75)
(10-year risk
8%)
Odds
1: 12.5
OAPR
1:4
26 445
offered
statins
92,003
unaffected
FPR=23%
21,575
false
positive
70 428
unaffected
test negative
Population prevention
(c)
7997
affected
RRR=20%
1600
events
prevented
100,000
(age 40-75)
(10-year risk
8%)
Odds
1: 12.5
100 000
offered
statins
92,003
unaffected

(a)



LR	Odds
0.4	1:160
0.7	1:91
	1:64
1.2	1:53
2.3	1:28

(b)