Supplementary Information

Table of Contents	
Supplementary Notes	1
Supplementary Note 1. PGS Catalog Scoring File Specifications.	1
Supplementary Note 2. Inclusion Criteria for the PGS Catalog.	3
A newly developed PGS	3
An evaluation of a previously developed PGS	4
Supplementary Note 3. PGS Catalog Data Acquisition and Curation Processes.	4
Supplementary Note 4. PGS Catalog Data Access and Implementation.	5
Supplementary Note 5. Colorectal cancer benchmarking methods.	6
Supplementary Figures	7
Supplementary Figure 1. Examples of PGS Catalog Publication and Trait website pages.	7
Supplementary Figure 2. Examples of PGS Catalog search results.	8
Supplementary Figure 3. Performance metrics for colorectal cancer PGS in UKB.	9
Supplementary Tables	11
Supplementary Table 1. FAIR indicators of PGS Catalog.	11
This table describes details of how the current PGS Catalog conforms to FAIR data principles. For the purposes of this table the Score constitutes the data (e.g. varian effect weights and alleles), and is linked to metadata (Samples, Performance Metric Publications) describing it.	ts,
Supplementary Table 2. PGS Catalog reporting items.	13
Supplementary Table 3. UKB Benchmarking cohort description and results.	17
Supplementary References	18

Supplementary Notes

Supplementary Note 1. PGS Catalog Scoring File Specifications.

The PGS Catalog's Scoring File format is described on our website: <u>https://www.pgscatalog.org/downloads/</u>. Each scoring file (variant information, effect alleles/weights) is formatted to be a gzipped tab-delimited text file, labelled by its PGS Catalog Score ID (e.g. PGS000001.txt.gz). We developed the scoring file format to closely resemble existing formats used to calculate scores in common software (e.g. PLINK) so that users could easily apply these scores within existing pipelines.

Scores are extracted from the relevant publication, and a consistent header (lines starting with #) has been added to each file listing relevant information about the PGS with links to the original publication and Catalog identifier:

```
### PGS CATALOG SCORING FILE - see www.pgscatalog.org/downloads/#dl_ftp for
additional information
## POLYGENIC SCORE (PGS) INFORMATION
# PGS ID = PGS identifier, e.g. 'PGS000001'
# PGS Name = PGS name, e.g. 'PRS77_BC' - optional
# Reported Trait = trait, e.g. 'Breast Cancer'
# Original Genome Build = Genome build/assembly, e.g. 'GRCh38'
# Number of Variants = Number of variants listed in the PGS
## SOURCE INFORMATION
# PGP ID = PGS publication identifier, e.g. 'PGP000001'
# Citation = Information about the publication
# LICENSE = License and terms of PGS use/distribution - refers to the EMBL-EBI
Terms of Use by default
rsID chr_name chr_position effect_allele reference_allele...
```

PGS scoring files are re-formatted to have consistent column headings based on the following schema:

Column Header	Field Name	Field Description	Mandatory?
rsID	dbSNP Accession ID (rsID)	The SNP's rs ID	
chr_name	Location - Chromosome	Chromosome name/number associated with the variant	Yes - Each PGS Scoring file must have
chr_position	Location - Base pair position within the Chromosome	Chromosomal position associated with the variant	either an <i>rsID</i> column or both a <i>chr_name</i> and <i>chr_position</i> column to identify the variant.
effect_allele	Effect Allele	The allele that's dosage is counted (e.g. {0, 1, 2}) and multiplied by the variant's weight ('effect_weight') when calculating score. The effect allele is also known as the 'risk allele'.	Yes
reference_allele	Reference Allele	The other allele(s) at the loci	Suggested - most software requires this for the calculation of scores and matching of the variants to existing genotype data,
effect_weight	Variant Weight	Value of the effect that is multiplied by the dosage of the effect allele ('effect_allele') when calculating the score.	Yes
locus_name	Locus Name	This is kept in for loci where the variant may be referenced by the gene (APOE e4). It is also common (usually in smaller PGS) to see the variants named according to the genes they impact.	Optional
weight_type	Type of Weight	Whether the author supplied Variant Weight is a: beta (effect size), or a log(OR/HR (odds/hazard ratio))	Optional

allelefrequency_ effect	Effect Allele Frequency	Reported effect allele frequency, if the associated locus is a haplotype then haplotype frequency will be extracted.	Optional
is_interaction	FLAG: Interaction	This is a TRUE/FALSE variable that flags whether the weight should be multiplied with the dosage of more than one variant. Interactions are demarcated with a $_x_$ between entries for each of the variants present in the interaction.	Optional
is_recessive	FLAG: Recessive Inheritance Model	This is a TRUE/FALSE variable that flags whether the weight should be added to the PGS sum only if there are 2 copies of the effect allele (e.g. it is a recessive allele).	Optional
is_haplotype		This is a TRUE/FALSE variable that flags	
is_diplotype	FLAG: Haplotype or Diplotype	whether the effect allele is a haplotype/diplotype rather than a single SNP. Constituent SNPs in the haplotype are semi-colon separated.	Optional
imputation_metho d	Imputation Method	This describes whether the variant was specifically called with a specific imputation or variant calling method. This is mostly kept to describe HLA-genotyping methods (e.g. flag SNP2HLA, HLA*IMP) that gives alleles that are not referenced by genomic position.	Optional
variant_descript ion	Variant Description	This field describes any extra information about the variant (e.g. how it is genotyped or scored) that cannot be captured by the other fields.	Optional
inclusion_criter ia	Score Inclusion Criteria	Explanation of when this variant is included into the PGS (e.g. if it depends on the results from other variants).	Optional

Supplementary Note 2. Inclusion Criteria for the PGS Catalog.

For the current PGS Catalog inclusion criteria see: <u>https://www.pgscatalog.org/about</u>. For a publication's data to be included in the PGS Catalog, it must fulfil the following criteria for either a newly developed polygenic score or an evaluation of an existing score(s):

A newly developed PGS

This includes the following information about the score and its predictive ability (evaluated on samples not used in training):

- Variant information necessary to apply the PGS to new samples (variant rsID and/or genomic position, weights/effect sizes, effect allele, genome build).
- Information about how the PGS was developed (computational method, variant selection, relevant parameters).

- Descriptions of the samples used for training (e.g. discovery of the variant associations [these can usually be extracted directly from the GWAS Catalog using GCST IDs], as well as fitting the PGS) and external evaluation.
- Establishment of the PGS' analytic validity, and a description of its predictive performance (e.g. effect sizes [beta, OR, HR, etc.], classification accuracy, proportion of the variance explained (R²), and/or covariates evaluated in the PGS prediction).

An evaluation of a previously developed PGS

This would include the evaluation of PGS already present in the Catalog (or one that meets the inclusion criteria specified above), on samples not used for PGS training. The requirements for description would be the same as for the evaluation of a new PGS.

Supplementary Note 3. PGS Catalog Data Acquisition and Curation Processes.

The current PGS Catalog employs a manual search process to identify publications that may be eligible for inclusion in the PGS Catalog. Papers are identified on Google Scholar, PubMed and Twitter using common keywords: "genetic risk", "polygenic risk", "polygenic risk score", "polygenic score", and "genetic risk score". A curator then scans the abstract and methods/results to identify whether the paper develops and/or validates a PGS (inclusion criteria #1), and adds it to our curation queue if the paper appears eligible. Subsequently, the paper is checked for the inclusion of PGS information (e.g. variants, effect alleles/weights), often sourced from supplementary excel spreadsheets within the paper, or in many cases extracted from external websites, figshare accessions, or Google drives linked within the paper. PGS information was determined to be available for inclusion in the Catalog provided no terms or restrictions on the data are imposed for download or resharing. If the data is unavailable, or sufficient information is not provided, the paper is marked as currently ineligible and a data-request email is sent to the corresponding authors if the paper is prioritized for curation. Papers suggested by users or communicated to us by authors are also checked according to this process and added to the queue.

Papers for full inclusion in the Catalog are selected from the list of eligible papers, prioritizing the papers that have been submitted to us by authors/users and based on data availability, citations, and our efforts to make the Catalog more comprehensive with respect to the diversity of traits included as well as ancestral diversity of populations represented in score development and evaluation. Full curation involves filling out a curation template (current version: www.pgscatalog.org/template/current) and formatting the variant information to have column headings consistent with our PGS Scoring File specification. Guidelines for filling out the curation template and extracting relevant data are provided online (current version: http://www.pgscatalog.org/docs/curation), and were developed in collaboration with experienced curators from the NHGRI-EBI GWAS Catalog. The curation guidelines describe the aspects of PGS study design captured in the Catalog, and each of the extracted data fields at each stage. Curation templates and scoring files were completed by expert curators, according to the information provided in the publication, or submitted by authors. All completed templates were validated by a second curator to ensure consistency before being

uploaded to the database. The most up to date description of the process for users to submit PGS data is provided at <u>www.pgscatalog.org/submit</u>.

The PGS Catalog data is released as available, but will move to a more regular release schedule in the future as data input increases. Individual PGS metadata or scores are versioned by date and provided as *archived_versions* on our FTP site (<u>http://ftp.ebi.ac.uk/pub/databases/spot/pgs/</u>) along with *previous_releases* of the complete Catalog metadata if any changes are made.

Supplementary Note 4. PGS Catalog Data Access and Implementation.

Data in the PGS Catalog is provided under EMBL-EBI's standard terms of use (<u>https://www.ebi.ac.uk/about/terms-of-use/</u>). The data in the Catalog can be currently accessed in the following three ways:

- **Bulk download** of the entire PGS Catalog's metadata, describing all PGS in terms of their publication source, samples used for development/evaluation, and related performance metrics (details and links: www.pgscatalog.org/downloads/).
- The PGS Catalog FTP server (available at: <u>https://ftp.ebi.ac.uk/pub/databases/spot/pgs/</u>) is indexed by Polygenic Score (PGS) ID to allow programmatic access to the Scoring Files and metadata for each PGS, archived versions of the scoring files and metadata are also stored for reference (additional details: <u>www.pgscatalog.org/downloads/</u>).
- A REST API is also provided to allow programmatic access and querying of the PGS Catalog, better enabling other applications to be built on top of the resource. Endpoints to retrieve all or individual PGS Catalog data objects (Publications, Scores, Samples, Traits, Performance Metrics) are available (details at: <u>https://www.pgscatalog.org/rest/</u>).

The PGS Catalog is also is indexed on <u>FAIRsharing.org</u> (ref: <u>bsg-d001448</u>), and polygenic score identifiers (e.g. PGS000018) can be externally resolved via <u>IDENTIFIERS.org</u> (ref: <u>pgs</u>). A description of the FAIR indicators for the PGS Catalog are provided in <u>Supplemental Table 1</u>.

Additional bibliographic information for PGS Catalog **Publication** objects are retrieved from EuropePMC (e.g. title, authors, journal, publication dates)¹. Additional information for each ontology term (e.g. synonyms, parent/child relationships, and mapped terms from other ontologies and disease coding resources [e.g. ICD/READ/SNOMED]) from the EFO ² are obtained using the EMBL-EBI Ontology Lookup Service (OLS)³.

The PGS Catalog website and database are developed using the Django framework (version 3.1; <u>https://djangoproject.com</u>) in Python (version 3.8; <u>https://www.python.org</u>) with a PostgreSQL database (version 12; <u>https://www.postgresql.org/</u>). The search functionality is built using ElasticSearch (v7.8; <u>https://www.elastic.co</u>).The website, database, and search index are all deployed on the Google Cloud (<u>https://cloud.google.com/</u>). The codebase for the Catalog can be viewed within our public GitHub repository (<u>https://github.com/PGScatalog</u>), currently provided under an <u>Apache 2.0 License</u>.

Supplementary Note 5. Colorectal cancer benchmarking methods.

To evaluate the predictive ability of PGS for colorectal cancer in the Catalog we used data from the UK Biobank (UKB), a cohort of ~500,000 participants from three countries (England, Wales, Scotland) of the United Kingdom⁴. Our analysis included 421,332 participants with genetic and phenotypic data (<u>Supplemental Table 3</u>), corresponding to 409,253 participants of European ancestry (UKB "White British" subset), 6,086 South Asian ancestry, and 5,984 African ancestry participants. South Asian (self-identifying as: Indian, Pakistani, or Bangladeshi) and African ancestry (self-identifying as: Caribbean, African, or Any other black background) participants were defined using an identical process to the White British participants, using principal components of genetic ancestry to identify a homogenous subset of self-identifying individuals by clustering⁴.

Diagnosis of colorectal cancer was performed using data linkage to the UK's national cancer and death registries. Cases of colorectal cancer were identified using previously used ICD codes in UKB 5 :

ICD9: 153.0 - 153.9, 154.0, 154.1, 154.8 ICD10: C18.0 - C18.9, C19, C20, C21.8

For each colorectal cancer diagnosis or death we recorded the date and age of the event. colorectal cancer events were defined as the first event of colorectal cancer, and participants were censored after the last cancer registry linkage date (2016-03-31). We excluded 449 participants who had self-reported history of colorectal cancer at recruitment and no linked cancer registry data.

PGS files were downloaded from the PGS Catalog and scores for each participant were calculated using PLINK⁶. Scores were standardised within each ancestry; the mean and standard deviation for colorectal cancer cases and controls are reported by ancestry group (<u>Supplemental Table 3</u>).

Each score's predictive ability is measured in terms of classification of individuals diagnosed with colorectal cancer versus those without, via the standardised effect size of the PGS (OR/HR per standard deviation increase of PGS) and classification accuracy (AUROC and concordance statistic [C-index]). We measured the HR and C-index using a Cox Proportional Hazards model with age-as-timescale, adjusting for sex, age at recruitment, country of recruitment, genotyping array, and 10 PCs of genetic ancestry. We measured the OR and AUROC using a logistic regression model adjusting for the sex, age at recruitment, country of recruitment, genotyping array, and 10 PCs of genetic ancestry. The effect sizes are reported with the 95% confidence interval for each PGS (<u>Supplemental Table 3</u>). Statistical analyses were performed in python: the Cox model was implemented using the *lifelines* package⁷, and logistic regression was performed using the *statsmodels* package⁸.

Supplementary Figures

			Publication In	formation (Euro	pePMC)					
Title	Genomic R	isk Prediction of C	Coronary Artery I	Disease in 480,0	00 Adults: Implic	ations for I	Nimary Pres	ertion.		
PubMed ID		(Europe PMC ?	,							
doi		cc.2018.07.079 (#								
Publication Date										
	J Am Coll C	tandiof braham G, Nelson	CR Wood MA Se	meters M1. Dock	ridne E I al EV Ka	ctopa G. Ba	omencies M 5	V Shown		
Poleased in PGS			01,11000788,08		- togo 1, car 1 1, rea	940 9 0 0. 04	erjeona ne, r			
Associated			e(s)							
GS Develo										
do bereio	peaby	This F doile	auon							H
									Search	
Polygenic Score	(PG5) ID *	PGS Name	PGS Publicat	tion (PGP) ID	Reported Trait		ped Trait(s)	(Ontology)	Number of Varianta	PGS Scoring File (FTP Lin
PG5000018		metuGR5_CAD	> Incorpo M et al.	J Am Coll Candiol	Coronary artery disease	coro	nary arlery di	sease	1,745,180	B
External PG		PGS Name	S Publicatio		Reported Trait	. Here	ed Trait(s) (C	-	Search Number of Variants	III • ▲ • PGS Scoring File (FTP Link
	(PGS) 10 *		PGS Publicatio	si (PGP) ID	Coronary arkery					
PG5000011		GR\$50	> Tada H et al. Eur	Heart J (2015)	dsease	corors	ary artery dis	8330	50	2
PG5000012		GR540K	PGP000005 > Absahan G et al.	Eur Heart J	Coronary artery disease	corora	ary artery dis	8850	49,310	e
PGS Perfc	formance me	trics are displayed	pone as reported by the stalog Sample Set	e source studies. (PSS) ID), phere	It is important to n Styping, and statist	ole that me ical modelli	trics are not r rg. Please re	recessarily o for to the so	omparable with each other. F aree publication for additional Search	or example, motrics depend on guidance on performance.
Naclaimer: The per	formance me ristics (descri	trics are displayed bed by the PGS Ca ample Set ID	as reported by the	e source studies. (PSS) ID), phon PGS Effect Si (per SD chare	otyping, and statist	ole Etat me ical modelli sasiFication	ng. Please re	fer to the so	ree publication for additional	guidance on performance.
Neclaimer: The per he sample characte	formance me ristics (descri PGS &	trics are displayed bed by the PSS Ca ample Set ID 3)	as reported by the stalog Sample Set	PGS Effect Si	zes pr) PGS Cb AUROC	ical modelli assification	ng, Please re n Metrics ()	for to the sol	ste publication for additional Search	guidance on performance.
Naclainer: The per he sample characte Evaluated Score PGS00018	PGS 5 (PSS 10	trice are displayed bed by the PGS Ca ample Set ID 3)	as reported by the statog Sample Set Trail: Besotted Trail: Incident coconary	PGS Effect Si (per SD chare HR:	2019 2019 2019 2019 PGS Cb 2019 PGS Cb 2019 2019 2019 2019 2019 2019 2019 2019	ical modelli sssification : 0.79	ng, Please re n Metrics ()	Covariati sex, gene amity	see publication for additional Search na included in the Model	guidance on performance.
Evaluated Score PS800018 (neta095_CAD) PS800012	formance me ristics (descri	trics are displayed bed by the PGS Ca ample Set ID 3) 2018	as reported by the statog Sample Set Trait Besorted Trait Incident coronary artery disease Besorted Trait Incident coronary	PGS Effect Si (per 5D chare HR: 1.706 (1.662, 1 HR:	2019 PGS CL 2019 PGS CL 2019 C-index 2019 -	ical modelli sssification : 0.79	ng, Please re n Metrics ()	Covariate Sex, gene amay bex, gene amay	Search Search es Included in the Model 60 PCs (1-10), genetyping	guidance on performance.
Naclement: The port the sample character Evaluated Score PGS000018 (entroPRS_CAD) PGS000012 (pRS500011 (pRS50)	POS 8 POS 8 PSS00 PSS00 PSS00	trics are displayed bed by the PGS Ca ample Set ID 3) 2018	as reported by the nation Sample Set Trait Besoned, Trait Incident coronary antery disease Besoned, Trait Incident Besoned, Trait Incident December 2015	PGS (ID), phone PGS Effect Si (per 5D chare) HR: 1.706 (1.662, 1 HR: 1.524 (1.495, 1 HR:	2019 PGS CL 2019 PGS CL 2019 C-index 2019 -	ical modelli sssification : 0.79	ng, Please re n Metrics ()	Covariate Sex, gene amay bex, gene amay	Search Search in Included in the Medel dic PCs (1-10), genetyping dic PCs (1-10), genetyping	guidance on performance.
Naclement: The port the sample character Evaluated Score PGS000018 (entroPRS_CAD) PGS000012 (pRS500011 (pRS50)	POS 8 POS 8 PSS00 PSS00 PSS00	trics are displayed bed by the PGS Ca ample Set ID 3) 2018	as reported by the nation Sample Set Trait Besoned, Trait Incident coronary antery disease Besoned, Trait Incident Besoned, Trait Incident December 2015	PGS (ID), phone PGS Effect Si (per 5D chare) HR: 1.706 (1.662, 1 HR: 1.524 (1.495, 1 HR:	2019 PGS CL 2019 PGS CL 2019 C-index 2019 -	ical modelli sssification : 0.79	ng, Please re n Metrics ()	Covariate Sex, gene amay bex, gene amay	Search Search in Included in the Medel dic PCs (1-10), genetyping dic PCs (1-10), genetyping	guidance on performance.
Neckelmer: The per he sample characte Evaluated Score PGS000018 (metac0RS_CAD) PGS000012 (0RS480) PGS000011	PSS00 PSS00 PSS00 PSS00 PSS00 amples	trics are displayed bed by the PGS Ca ample Set ID 3) 2018	as reported by the trateg Sample Set Trail Besoned Date incident occuracy anthry disease Besoned Tast incident coronary anthry disease	PGS (ID), phone PGS Effect Si (per 5D chare) HR: 1.706 (1.662, 1 HR: 1.524 (1.495, 1 HR:	2019 PGS CL 2019 PGS CL 2019 C-index 2019 -	ical modelli mailfication : 0.79 : 0.623 (3.6	ng, Please re n Metrics ()	Covariate sex, pene amay sex, pene amay sex, pene amay	Search Search es Included in the Model de PCs (1-16), genotyping de PCs (1-16), genotyping de PCs (1-16), genotyping	guidence or performance. If the A + PGS Performance: Other age-as time solite Car regin age-as time solite Car regin

| Trait: brea

 | |
 | |
 | | | |
 | | | |

--|--
--
---|--|---
--|---|---|---
---|--|---|
| Martine 11

 | Experime |
 | logy (EPO) Information |
 | | | |
 | | | |
|

 | |
 | helial cells of the breast (N | IONDO: DesignP
 | attern) | | |
 | | | |
| Trait category

 | Learcinoma that arises from ophihalial cells of the breast (MONDO: DesignPattern) |
 | |
 | | | |
 | | | |
| Synonyms 17

 | 7 synonyms O |
 | |
 | | | |
 | | | |
| Mapped term(s) 11

 | | ns O
 | |
 | | | |
 | | | |
| Child trait(s) 6 c

 | |
 | |
 | | | |
 | | | |
| Associated I

 | |
 | (S) |
 | | | |
 | | | |
| Note: This table sho

 | vs all PGS to | or "breast cardino
raits
 | ma' and any child terms o | d this trait in the E
 | 3PO hiera | rchy by delault. | |
 | | | |
| Polygenic Score (PC

 | 000 | BOS Name
 | PGS Publication (PG | 60
 | Reports | d Task | _ | Mapped Trail(
 | Sear | _ | III • |
| PG500001

 | | PRS77_BC
 | PGP000001
+ Neveddel N et el 2 Nel C |
 | Breast C | | | breast carcinor
 | | ogy | 77 |
| PG500002

 | | PRS77_ERpos
 | PGP000001 |
 | ER-posi | ive Breast Cancer | | estrogen recep
cancer
 | tor positi | ve breast | 77 |
| P\$500003

 | | PRS77_ERveq
 | Mevaddet N et al. J Nat C PGP000001 Mevaddet N et al. J Nat C |
 | | tive Breast Caroo | , |
 | tor negat | | 77 |
| PG500004

 | | PR5313_BC
 | > Mevadat N et al. J Nat C
PGP000002
> Mevadat N et al. Am J H |
 | Breast C | | | cancer
breast carcinor
 | | | 313 |
|

 | |
 | PGP00002 |
 | | | |
 | | ve breast | 313 |
| PG5000005

 | | PRS313_ERpos
 | |
 | | ive Breast Cancer | | estrogen-recep
cancer
 | | | |
| PG500006

 | | PR5313_ERnog
 | PGP000002
+ Neveddet N et al. Am J H | um Genet (2018)
 | | tive Breast Canor | r | estrogen-recep
cancer
 | | | 313 |
| P3500007

 | | PR53820_BC
 | PGP00002 >Messibilit N et al. Am J H |
 | Breast C | | | breast carcinor
 | | | 3,820 |
| PG500008

 | | PR53820_ERpos
 | PGP00002
> Mevadole N et al. Am J H |
 | | ive Breast Cancer | | estrogen-recep
cancer
 | | | 3,820 |
| PG500009

 | | PR\$3620_ERneg
 | PGP000002
In Manadole IN et al. Am J H |
 | ER-regi | stive Breast Cance | r | estrogen-recep
cancer
 | tor negat | ive breast | 3.820 |
| PG500015

 | | GPS_BC
 | PGP000006
+ Khers All et al. Nat Gene | (22119)
 | Breast c | ancer | | breast carcinor
 | na | | 5,218 |
| PG5000028

 | | PRS
 | PGP000018 > Shish Y et al. Broast Can | or Pen Treat
 | Breast c | ancer | | breast carcinor
 | na | | 83 |
| PGS Perform

 | |
 | d as reported by the source
the PGS Catalog Sample | a studies. It is imp
Set (PISS) ID), pr
 | ortant to
tenotypin | note that metrics a
g, and statistical m | re not ne
odeling. | cessarily compan
Please refer to th
 | able with
e source
Sear | each other. Fo
publication to
th | |
|

 | formance me
le characteris | etics are displayed
stics (described by
 | d as reported by the source
the PGS Catalog Sample
Performance Source | Trait
 | PG5 | nde fhat metrica a
g, and statistical m
Effect Sizes
SD change) | | cessarily company
Piesse refer to the
Dessification Me
 | Sear | ch | |
| Disclaimer: The per
depend on the sampli
performance.

 | formance me
le characteris | etrics are displayed
alics (described by
sple Set ID
 | |
 | PG5 | | PGS |
 | Sour | ch | |
| Dischlerer: The per
depend on the samp
performance.
Evaluated Score
PG5000001
(PRS77_BC)

 | PGS Sen
(PSS ID) | etics are displayed
stics (described by
sple Set ID
01
 | Performance Source | Trait
Beconted Trait
All breast
carbor
 | PGS
(per | Effect Sizes
SD change) | PGS | lassification Me
 | Sour | ch | or example, met
additional guid
IIII T |
| Disclaiver: The peri
depend on the samp
performance.
Evaluated Boore
PG5000001
(PHS77_BC)
PG5000000
(PHS77_BRing)

 | POS Sen
(PSS Sen
(PSS D)
PSS0000 | ables are displayed
ratios (described by
spile Set ID 1
101 1
102 1
 | Performance Source | Trait
Beconted Traits
All breast
cancer
Beconted Traits
EPI regative
breast cancer
 | PG8
(per
0R:
0R: | Effect Sizes
SD change)
1.55 (1.52, 1.58)
1.45 (1.4, 1.48) | PGS | lassification Me
 | Sour | ch | |
| Dischlerer: The per
depend on the samp
performance.
Evaluated Score
PG5000001
(PRS77_BC)

 | POS Sen
(PSS Sen
(PSS ID)
PSS0000 | ables are displayed
ratios (described by
spile Set ID 1
101 1
102 1
 | Performance Source | Trait
Beconted Toats
Al breast
cancer
Beconted Toats
ER-negative
breast cancer
Beconted Toats
ER-positive
breast cancer
 | | Effect Sizes
SD change)
1.55 (1.52, 1.54) | PGS | lassification Me
 | Sour | ch | |
| Disclaiver: The peri
depend on the samp
performance.
Evaluated Boore
PG5000001
(PIB577_BC)
PG5000000
(PIB577_BRing)

 | POS Sen
(PSS Sen
(PSS D)
PSS0000 | etica are displayed by a second by a secon
 | Performance Source | Trait
Beconted Traits
All breast
cancer
Beconted Traits
EPI regative
breast cancer |
 | Effect Sizes
SD change)
1.55 (1.52, 1.58)
1.45 (1.4, 1.48) | PGS | EaseHisation Mar
g: 0.622 (0.619, 0 | Sour
 | ch | Isoluded in th |
| Decision: The performance.
International control of the sample
performance.
Resource of the sample
Resource of the sample
 | POS Sen
(PSS Sen
(PSS Social
PSS0000
PSS0000
PSS0000
 | etica are displayed
rece (described by
reple Set 10 1 1
101 1
100 100
 | Performance Source PGP000001 Vervalue N al (2015) PGP00001 PGP00001 Vervalue N al (2015) PGP00002 PGP00002 Vervalue N al (2015) PGP00002 Vervalue N al (2015) | Trait
Recorded Task
All breast
cancer
Besoned Task
EFA regular
breast concer
Besoned Task
bryasive breast
cancer
Besoned Task
bryasive breast
cancer
Besoned Task | PG8
(per
0R:
0R:
0R:
0R:
0R:
0R:
0R:
0R:
0R:
 | Effect Sizes
SD change)
1.55 (1.62, 1.64)
1.45 (1.4, 1.48)
1.63 (1.6, 1.47) | PGS C
C-inde
-
AURO | EaseHisation Mar
g: 0.622 (0.619, 0 | Sour
 | ch
Covariates
-
- | Iscluded in th |
| Decisitive: The period depend on the sample
performance. The sample
performance. Prakwated Score Prakwated Score POS000001 POS000001 POS000002 POS000002
 | POS Sam
(1955 80)
PS50000
PS50000
PS50000
PS50000
 | teles are displayed by
refer (described by
npite Set ID
 | Performance Source DrDP00001 Amounter N et al (2016) Pr0P00001 Amounter N et al (2016) Pr0P00002 Pr0P0000 Pr0P0000 Pr0P0000 Pr0P0000 Pr0P000 Pr0P000 Pr0P000 Pr0P0000 Pr0P0000 Pr0P000 Pr0P00 Pr0P00 Pr0P00 Pr0P00 Pr0P00 Pr0P00 Pr0P00 Pr0P0 Pr0P00 Pr0P00 Pr0P0 Pr0P00 Pr0P0 | Trait
Beconst Tast
Al broast
Encoded Tast
Propositive
trast cancer
Beconst Tast
Propositive
trast cancer
Beconst Tast
Prosolutions
Concer
Beconst Tast
Concer
Beconst Tast
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Concer
Conc | POS (per : OR: : OR: : OR: : OR: : OR: :
OR: | Effect Sizes
5D chaspe)
1.65 [1.02, 1.66]
1.65 [1.0, 1.66]
1.65 [1.0, 1.67]
1.63 [1.0, 1.67]
1.61 [1.07, 1.65]
1.66 [1.61, 1.7] | PGS C
C-inte
-
AURO | Sassification Me
& 0.622 (0.619, 0
C: 0.63
C: 0.636 | Sour
 | Coveriates shudy, gener | tic PCs 1-15 |
| Decision: The performance.
International control of the sample
performance.
Resource of the sample
Resource of the sample
 | POS San
(PS San
(PS San
(PS Soco
PS Soco
PS Soco
PS Soco
PS Soco
 | teles are displayed by
refer (described by
npite Set ID
 | Performance Source PGP000001 Vervalue N al (2015) PGP00001 PGP00001 Vervalue N al (2015) PGP00002 PGP00002 Vervalue N al (2015) PGP00002 Vervalue N al (2015) | Trait
Resorted Tast
Al Dreat
cancer
Resorted Tast
Sit result
cancer
Resorted Tast
Dit positive treast cancer
Resorted Tast
Provide treast
cancer
Resorted Tast
Concent Tast
C | POE (per (per <
 | Effect Sizes
50 change)
1.65 [1.42, 1.66]
1.65 [1.4, 1.48]
1.63 [1.4, 1.47]
1.63 [1.57, 1.65] | PGS C
C-inte
-
AURO | Classification Me
© 0.402 (0.419, 0
© 0.403 | Sour | Covariates study, genet
 | tic PCs 1-15 |
| Decisitive: The period depend on the sample
performance. The sample
performance. Prakwated Score Prakwated Score POS000001 POS000001 POS000002 POS000002
 | POS Sam
(1955 80)
PS50000
PS50000
PS50000
PS50000
 | Notes are displayed by an end of the section of the
 | Performance Source DrDP00001 Amounter N et al (2016) Pr0P00001 Amounter N et al (2016) Pr0P00002 Pr0P0000 Pr0P0000 Pr0P0000 Pr0P0000 Pr0P000 Pr0P000 Pr0P000 Pr0P0000 Pr0P0000 Pr0P000 Pr0P00 Pr0P00 Pr0P00 Pr0P00 Pr0P00 Pr0P00 Pr0P00 Pr0P0 Pr0P00 Pr0P00 Pr0P0 Pr0P00 Pr0P0 | Trait
Records Tast
Al broat
cancer
Becond Tast
Broageline
trast cancer
Record Tast
Diponite
trast cancer
Record Tast
broative
trast cancer
Record Tast
broative
trast cancer
Record Tast
trastic
cancer
Record Tast
trastic
cancer
Record Tast
trastic
cancer
Record Tast
trastic
cancer
Record Tast
trastic
cancer
Record Tast
trastic
cancer
Record Tast
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic
trastic |
PGE
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(pe))
(per
)
(per
)
(per
))
(per
)
(per
))
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
)
(per
))
(per
)
(per
))
(per
)
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
(per
))
()
)
()
)
(per
))
())
()
))
()
())
())
(| Effect Sizes
5D chaspe)
1.65 [1.02, 1.66]
1.65 [1.0, 1.66]
1.65 [1.0, 1.67]
1.63 [1.0, 1.67]
1.61 [1.07, 1.65]
1.66 [1.61, 1.7] | Pas a
C-inde
 | Sassification Me
& 0.622 (0.619, 0
C: 0.63
C: 0.636 | Sour
 | Coveriates shudy, gener | tic PCs 1-15
tic PCs 1-15 |
| Disclation: The period depend on the sample interview. The second period Enaluted Score The second period Passoccore The seco
 | P350000
P350000
P350000
P350000
P350000
P350000
 | tota are displayed by the set ID * 1 nybe Set ID * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 004 * 1 1 004 * 1 1 004 * 1 1 004 * 1 1 005 * 1 1
 | Performance Source Performance Source Performance Source Performance Performan | Trait Bacons Tast Al breat casoor Bacons Tast acasor Bacons Tast Bacons Tast Propulse trast cancer Propulse trast cancer Bacons Tast Proster Trait Proster Trait Proster Tast | POS
(per
COR:
COR:
COR:
COR:
COR:
COR:
COR:
COR:
 | Effect Sizes
30 chaspe)
1.05 [1.0; 1.0;
1.45 [1.4; 1.46]
1.63 [1.4; 1.47]
1.63 [1.4; 1.47]
1.61 [1.57; 1.66]
1.66 [1.41; 1.7]
1.66 [1.42; 1.66] | PGS C
C-intel

AURO
AURO
AURO | Classification Mer
(c. 0.622 (p.em, o)
(c. 0.63
(c. 0.636
(c. 0.636 | Sour | ch
Covariates
-
-
study, genet
study, genet
 | 10 Included in the
Included in the
Included In the
Included In the
Included Included Included
Included Included Included
Included Included Included Included
Included Included |
| Decisitive: The performance. depend on the sample performance. Persistent Boose RVANUERD Boose Persistent Boose RVSD000000000000000000000000000000000000

 | formance me
PGS Same
PGS Same
PS50000
PS50000
PS50000
PS50000
PS50000
PS50000
PS50000
PS50000
PS50000
PS50000 | tota are displayed by the set ID * 1 nybe Set ID * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 004 * 1 1 004 * 1 1 004 * 1 1 004 * 1 1 005 * 1 1
 | Performance Source of Performance Source Source (Performance Source Source (Performance Source Source (Performance Source Source (Performance Source | Trait Resofted Tast All breat assort Resofted Tast All breat assort Resofted Tast Fit positive breast cancer Resofted Tast breast cancer Resofted Tast breast brea | POS
(per
COR:
COR:
COR:
COR:
COR:
COR:
COR:
COR:
 | Effect files
60 chaspe)
1.05 [1.0; 1.9]
1.45 [1.4; 1.46]
1.45 [1.4; 1.46]
1.63 [1.4; 1.47]
1.64 [1.47; 1.66]
1.66 [1.41; 1.7]
1.66 [1.42; 1.46]
1.68 [1.42; 1.46] | PGS C
C-intel

AURO
AURO
AURO | 2assification Me
© 0.622 (p.ens, o
C. 0.63
© 0.636
© 0.645
© 0.641 | Sour
 | Covariates Covariates study: gener etudy: gener etudy: gener etudy: gener etudy: gener | 10 Included in the
Included in the
Included In the
Included In the
Included Included Included
Included Included Included
Included Included Included Included
Included Included |
| Disclaimer: The peri-
depend on the samp
production constraints Exeluted Boost
(PRS27, EPupel) Postcocce
(PRS27, EPupel) Postcocce
(PRS22, EPupel) Postcocce
(PRS22, EPupel)
 | formance me
PGS Same
PGS
Same
PS50000
PS50000
PS50000
PS50000
PS50000
PS50000
PS50000
PS50000
PS50000
PS50000 | tota are displayed by the set ID * 1 nybe Set ID * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 004 * 1 1 004 * 1 1 004 * 1 1 004 * 1 1 005 * 1 1
 | Performance Source Performance Source Performance Source Performance Performan | Trait Bacons Tast Al breat casoor Bacons Tast acasor Bacons Tast Bacons Tast Propulse trast cancer Propulse trast cancer Bacons Tast Proster Trait Proster Trait Proster Tast | POS
(per
COR:
COR:
COR:
COR:
COR:
COR:
COR:
COR:
 | Effect files
60 chaspe)
1.05 [1.0; 1.9]
1.45 [1.4; 1.46]
1.45 [1.4; 1.46]
1.63 [1.4; 1.47]
1.64 [1.47; 1.66]
1.66 [1.41; 1.7]
1.66 [1.42; 1.46]
1.68 [1.42; 1.46] | PGS C
C-intel

AURO
AURO
AURO | 2assification Me
© 0.622 (p.ens, o
C. 0.63
© 0.636
© 0.645
© 0.641 | Sour
 | Coversities Covers | 10 Included in the
Included in the
Ris PCs 1-15
Ris PCs 1-15
Ris PCs 1-15 |
| Disclaimer: The peri-
depend on the samp
production constraints Exeluted Boost
(PRS27, EPupel) Postcocce
(PRS27, EPupel) Postcocce
(PRS22, EPupel) Postcocce
(PRS22, EPupel)
 | forsance in a local strategy of the strategy o
 | tota are displayed by the set ID * 1 nybe Set ID * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 00 * 1 1 004 * 1 1 004 * 1 1 004 * 1 1 004 * 1 1 005 * 1 1
 | References Board
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDEDIT
ISANDED
ISANDEDIT
ISANDEDIT
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDED
ISANDE | Trait Bacons Tast Al breat casoor Bacons Tast acasor Bacons Tast Bacons Tast Propulse trast cancer Propulse trast cancer Bacons Tast Proster Trait Proster Trait Proster Tast | P08
jer
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
 | Effect files
60 chaspe)
1.05 [1.0; 1.9]
1.45 [1.4; 1.46]
1.45 [1.4; 1.46]
1.63 [1.4; 1.47]
1.64 [1.47; 1.66]
1.66 [1.41; 1.7]
1.66 [1.42; 1.46]
1.68 [1.42; 1.46] | POS 0
C-intel
-
AURO
AURO
AURO | 2assification Me
© 0.622 (p.ens, o
C. 0.63
© 0.636
© 0.645
© 0.641 | Source |
Coversities
Coversities
-
-
study, genet
study, genet | 10 Included in the
Included in the
Ris PCs 1-15
Ris PCs 1-15
Ris PCs 1-15 |
Disclaimer, The performance Disclaimer, The performance Producted Boost Robinson Robinson <t< td=""><td>formance in a feature of the deviation o</td><td>and digitapped bit is a a splat bit is a a a cl cl a a cl cl a a cl cl cl <td< td=""><td>References Board ISANDEDIT ISANDED ISANDEDIT ISANDEDIT ISANDED ISANDE</td><td>Trail Tenend Tarl Second Tarl Al Dead Al Dead Descend Tarl Respective Encode Land Descend Tarl Respective Descend Tarl Novals Head Descend Tarl Novals Head Descend Tarl Novals Head Descend Tarl Novals Head Descend Tarl Novals Head Descend Tarl Descend Tarl Desc</td><td>PG5 (jeer 2 OE 2 OE 2 OE 2 OE 2 OE 2 OE 2 OE 2 OE</td><td>Lifect lifes 100 margin 146 (1.4.14) 146 (1.4.14) 149 (1.</td><td>POS 0 C-intel - AURO AURO AURO</td><td>2 0.42 p.m. 0 2 0.42 p.m. 0 2 0.43 2 0.44 2 0.44</td><td>Sear</td><td>Coversiences Coversiences Cover</td><td>10 Included in the Included in the Ris PCs 1-15 Ris PCs 1-15 Ris PCs 1-15</td></td<></td></t<>	formance in a feature of the deviation o	and digitapped bit is a a splat bit is a a a cl cl a a cl cl a a cl cl cl a cl cl cl <td< td=""><td>References Board ISANDEDIT ISANDED ISANDEDIT ISANDEDIT ISANDED ISANDE</td><td>Trail Tenend Tarl Second Tarl Al Dead Al Dead Descend Tarl Respective Encode Land Descend Tarl Respective Descend Tarl Novals Head Descend Tarl Novals Head Descend Tarl Novals Head Descend Tarl Novals Head Descend Tarl Novals Head Descend Tarl Descend Tarl Desc</td><td>PG5 (jeer 2 OE 2 OE 2 OE 2 OE 2 OE 2 OE 2 OE 2 OE</td><td>Lifect lifes 100 margin 146 (1.4.14) 146 (1.4.14) 149 (1.</td><td>POS 0 C-intel - AURO AURO AURO</td><td>2 0.42 p.m. 0 2 0.42 p.m. 0 2 0.43 2 0.44 2 0.44</td><td>Sear</td><td>Coversiences Coversiences Cover</td><td>10 Included in the Included in the Ris PCs 1-15 Ris PCs 1-15 Ris PCs 1-15</td></td<>	References Board ISANDEDIT ISANDED ISANDEDIT ISANDEDIT ISANDED ISANDE	Trail Tenend Tarl Second Tarl Al Dead Al Dead Descend Tarl Respective Encode Land Descend Tarl Respective Descend Tarl Novals Head Descend Tarl Novals Head Descend Tarl Novals Head Descend Tarl Novals Head Descend Tarl Novals Head Descend Tarl Descend Tarl Desc	PG5 (jeer 2 OE 2 OE 2 OE 2 OE 2 OE 2 OE 2 OE 2 OE	Lifect lifes 100 margin 146 (1.4.14) 146 (1.4.14) 149 (1.	POS 0 C-intel - AURO AURO AURO	2 0.42 p.m. 0 2 0.42 p.m. 0 2 0.43 2 0.44 2 0.44	Sear	Coversiences Cover	10 Included in the Included in the Ris PCs 1-15 Ris PCs 1-15 Ris PCs 1-15
Declaration: The performance of the second period performance of the second performance of	tornance no los based of the deviation o	and during filter bit i in during filter bit i and during filter bit i in during filter bit i and during filter bit i in during filter bit i	Partementer Reve a la construir de la construi	Trail Tessential	Possibility	Effect lines 1 1-56 (1,00, 1,00) 1 <	POS 0 C-intel - AURO AURO AURO	Control 0 20 0427	Sear	Cheverlates Coverlates - - - study, genet study, gen	10 Included in the Included in the Ris PCs 1-15 Ris PCs 1-15 Ris PCs 1-15
Declaration: The performance. The performance. Persistent Control Persistent Control Passoner Persistent Control	ternance ne le characteris le characteris pessoon pess	and digraphic law digraphi law digraphic law digraphic law digraphic law digraphic law digr	Performance Reveals Performance Reveals Performance Pe	Test Alternative Alternative Consort Carlo Alternative Consort	POS (per (per <	Lifect lifes 100 margin 146 (1.4.1.46) 146 (1.4.1.47) 149 (1.4.1.4	POS 0 C-intel - AURO AURO AURO	Lassification Me 2 0.422 perc. 0 0 0.03 0 0.423 0 0.433 0 0.447 0 0.427 0 0.477 0 0.477 0 0.477 0 0.477 0 0.477 0 0.4	Sear Nice 6 8271 Sear Address 100055 100055	ch Coversites - - study, gener study, g	10 Included in the Included in the Ris PCs 1-15 Ris PCs 1-15 Ris PCs 1-15
Batter to applications Annual to applications Character to applications Annual to applications <tr< td=""><td>tornavce ne lo deverteres lo deve lo deverteres lo deverte</td><td>htto are digrifysperiod in the provided in the</td><td></td><td>Table Descend Table Al Investor Insected Table Descend Tab</td><td>Post juer control con</td><td>Hitel Bise Hitel Bise H</td><td>POS 0 C-intel - AURO AURO AURO</td><td>Conserve) C 0.653 C 0.654 C 0.655 C 0.</td><td>Sear NGS 1 RE71 Sear AddB 10005 10005</td><td>ch Covertaines - - study, ganal study, ganal study,</td><td>Re PCa 1-15 Re PCa 1-15</td></tr<>	tornavce ne lo deverteres lo deve lo deverteres lo deverte	htto are digrifysperiod in the provided in the		Table Descend Table Al Investor Insected Table Descend Tab	Post juer control con	Hitel Bise H	POS 0 C-intel - AURO AURO AURO	Conserve) C 0.653 C 0.654 C 0.655 C 0.	Sear NGS 1 RE71 Sear AddB 10005 10005	ch Covertaines - - study, ganal study,	Re PCa 1-15
Declaration: The performance. The performance. Persistent Control Persistent Control Passoner Persistent Control	tornavce ne lo deverteres lo deve lo deverteres lo deverte	and digraphic law digraphi law digraphic law digraphic law digraphic law digraphic law digr		Yead Descend Tart Al beam Proceedings and only the Proceedings and and only the Proceedings and	POE guer guer guer guer guer guer guer guer	Effect lines 1 1-56 (1,00, 1,00) 1 <	POS 0 C-intel - AURO AURO AURO	Lassification Me 2 0.422 perc. 0 0 0.03 0 0.423 0 0.433 0 0.447 0 0.427 0 0.477 0 0.477 0 0.477 0 0.477 0 0.477 0 0.4	Sear NGS 1 8271 Sear AddB 10005 10005	ch Coversites - - study, gener study, g	Re PCa 1-15
Batter to applications Annual to applications Character to applications Annual to applications <tr< td=""><td>tornavce ne lo deverteres lo deve lo deverteres lo deverte</td><td>white are displayed by a set of an and an an an an an and an an</td><td></td><td>Trait Description Description</td><td>Post geer geer control</td><td>Hitel Bise Hitel Bise H</td><td>POS 0 C-intel - AURO AURO AURO</td><td>Conserve) C 0.653 C 0.654 C 0.655 C 0.</td><td>Sear NGS 1 8271 Sear AddB 10005 10005</td><td>ch Coversities - - study, genet study, genet</td><td>Excluded in the Excluded in the E</td></tr<>	tornavce ne lo deverteres lo deve lo deverteres lo deverte	white are displayed by a set of an and an an an an an and an		Trait Description	Post geer geer control	Hitel Bise H	POS 0 C-intel - AURO AURO AURO	Conserve) C 0.653 C 0.654 C 0.655 C 0.	Sear NGS 1 8271 Sear AddB 10005 10005	ch Coversities - - study, genet study, genet	Excluded in the E
Restances in your of the second secon	Presented in the characteristic of the chara	white are displayed by a set of an and an an and an an and an An and an and an and an and an and an and an an and an and an	References Courses C	Tast Research Tast Annotation of the second tast Research and the second tast Research and tast Research T	Pott guers guers Pott guers official guers official construction official construction off	Effect lines Effect lines Effect lines Effect lines Effect lines Effect lines Effect Eff	POS 0 C-intel - AURO AURO AURO	Construction Me C 0.043 C 0.043 C 0.044 C 0.0447 C 0.0477 C 0.04777 C 0.04777 C 0.047777 C 0.0477777777777777777777777777777777777	Soar Nice 427] Soar Add5 10005 10005 10005	ch Coversities - - study, genet study, genet	RE PCa 1-15 Historie de la filia Historie PCa 1-15 Historie Ca 1-15 Historie Ca 1-15 Catalant Information
Restance in your of the second	Production of the characteristic of the char		References Research Characteristics Ch	Trait Research Lind waves Research Carlos Research Ca	POG geer Const	Bitel Rises Bitel Ris Bitel Ris Bitel Rises Bitel Rises Bitel Rises	POS 0 C-intel - AURO AURO AURO	Construction Me C 0.02 0.03 C 0.04 C 0.04	Soar Hiss 827) Soar Addis 10005 1000	ch Coversities - - study genet study genet	Resoluted in the Resolution Resolution Resolution Resolution Resolution Resolution Resolution

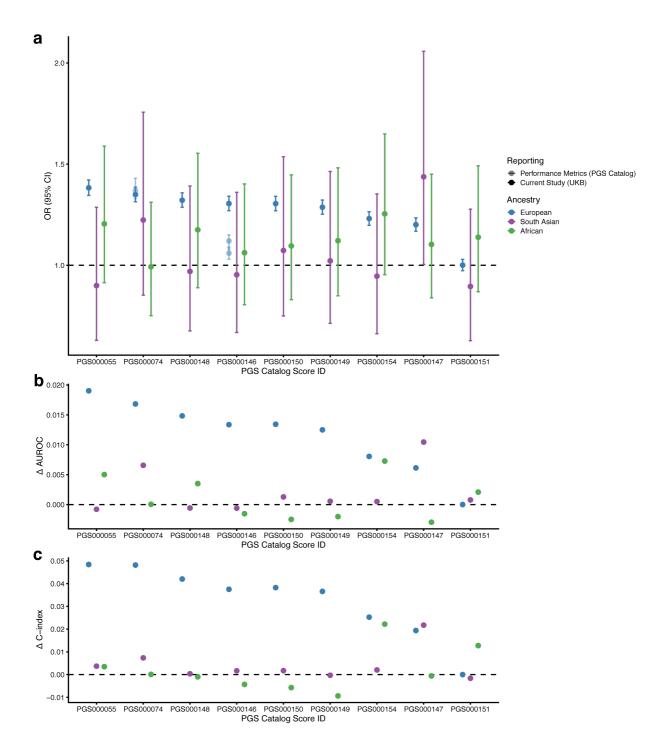
Supplementary Figure 1. Examples of PGS Catalog Publication and Trait website pages.

(A) Example of how each Publication and its related metadata (links to publication, EuropePMC, and PGS that were developed and evaluated within the paper) are displayed on <u>PGSCatalog.org</u> (example publication PGP00007⁹). (B) Example of how each Trait (ontology term, description, synonyms, mapped terms [e.g. ICD/SNOMED], and child ontology terms/sub-traits extracted from EFO^{2,3}) and its related metadata (PGS that have predicted the current trait, and subsequent evaluation of those scores) are displayed on <u>PGSCatalog.org</u> (example trait: breast carcinoma, EFO_000305). Sub-traits from the ontology (in this example breast cancer subtypes) are displayed by default, but can be removed by de-selecting the "Include PGS Score(s) for child traits" button. Sections of each webpage are highlighted with coloured bars corresponding to the data objects they display in **Figure 1A**.

PGS Catalog / Search / breast cancer	Search results for "diabetes"
Search results for "breast cancer"	O O O Publications 1
0 9 9 Al results (6) 0 9 Traits (7) 0 2 Addicators (7)	1 diabetes mellitus (Metadole disorder) (EFO_0000000)
Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. Manadatal N et al. (2018) - Am J Hum Genet PMID:30554720 doi:10.1016/j.a/pg.2018.11.002 Associated PGS © 0	A netabolic disorder characterized by abnormally high blood sugar levels due to diminished production of insulin or insulin resistance/desamititation. [IV.CIT: P378] < Show iess P050 P05 News Reported Trait P0500001 (0°5, Tio) Typ 2 abanes P050001 (0°5, Tio) Typ 2 abanes P050001 (0°5, Tio) Typ 2 abanes P050001 (0°51001) Typ 2 abanes P05001 (0°51001) Typ 2 abane
Prediction of breast cancer risk based on profiling with common genetic variants. Mavadat N et al. (2015) - J Nati Cancer Int PMID:25855707 doi:10.1093/jncidjv036 Associated POS O	Pd5600021 CRG1 Type 1 diabates Pd5500022 TD, GR6 Type 1 diabates Pd5500021 GR6 Type 1 diabates Pd5500021 GR6 Type 1 diabates
Breast cancer risk prediction using a clinical risk model and polygenic risk score. Shieh Y et al. (2016)- Breast Cancer Res Treat PMID 27555998 doi:10.1007/s10549-016-3953-2 Associated POS Concerning Polymony	PSS50001 CR081 Type 2 database, based on DMPs involved in § -64 farction) PSS500011 CR016 Type 2 database, based on DMPs involved in § -64 farctiona) PSS500011 CR016 Type 2 database PSS50011 CR016 Type 2 database PSS50011 CR016 Type 2 database PSS50011 CR016 Type 2 database
Prediction of breast cancer risk based on common genetic variants in women of East Asian ancestry. Wen W et al. (2016) - Breast Cancer Res PMID 27531260 dei:10.1186/s13058-016-0786-1 Associated POS ① 0	type I diabetes mellitus (Metabolic disorder) (Immune system disorder) (Digestive system disorder) (EFO_0001300 A chronic condition characterized by minimal or absent production of insulin by the pancreas. [NCIT Show more > Associated PGS © 0
C Evaluation of Polygenic Risk Scores for Breast and Ovarian Cancer Risk Prediction in BRCA1 and BRCA2 Mutation Carriers. Kuchenbaseker KB et al. (2017) - J Nati Cancer Inst PMID:28376175 doi:10.1093/prci/djw302 Associated POS ● 9	type II diabetes mellitus Dependent storder Metabolic disorder (EFO_0001366) Alype of diabetes mellitus that is characterized by insulin resistance or desensitization and incre Show more > Associated POS O
Prediction of Breast and Prostate Cancer Risks in Male BRCA1 and BRCA2 Mutation Carriers Using Polygenic Risk Scores. Lexemperiter J et al. (2017) - J Clin Chicol PMID:28448241 doi:10.1200UC0.2016.69.4395 Associated POS ① 9	diabetes mellitus biomarker (The measurement (FF0,0006842) Associated POS O FOS Man Reported Trait Possent12 GS-EXA Herrogian Atc Possent12 GS-EXA Herrogian Atc Possent12 GS-EXA Herrogian Atc Possent12

Supplementary Figure 2. Examples of PGS Catalog search results.

Traits and publications are indexed, and can be queried through the search box at the top right corner of each page's header (**A**). By default the search returns both trait and publication results, but results can be faceted to either. (**A**) Example of search results for publications related to "*breast cancer*". (**B**) Example of search results for traits (ontology terms) related to "*diabetes*". The results include a higher-level "diabetes mellitus" trait (which includes both type 1 and 2), the specific subtypes, and polygenic scores for the related HbA1c measurements under the diabetes mellitus biomarker trait.



Supplementary Figure 3. Performance metrics for colorectal cancer PGS in UKB.

Each PRS was evaluated within a logistic regression model for predicting colorectal cancer status for participants in UKB (**A-B**), and a separate Cox proportional hazards regression model (age-as-timescale) (**Figure 2, C**). (**A**) Standardised effect size (Odds Ratio; OR) describing the odds of having colorectal cancer per unit increase in each PGS. Previously reported effect sizes that were recorded in the Catalog are also plotted for PGS000074 and PGS000146. (**B**) Change in model classification accuracy (Area Under the Receiver Operating Characteristic Curve; Δ AUROC) when the PGS is added to a logistic regression model including the existing covariates (age at recruitment, sex, recruitment country, genotyping array, and 10 PCs of genetic ancestry). (**C**) Change in model classification

accuracy (concordance statistic; Δ C-index) when the PGS is added to a risk model including the existing covariates (sex, age at recruitment, recruitment country, genotyping array, and 10 principal components [PCs] of genetic ancestry).

Supplementary Tables

Supplementary Table 1. FAIR indicators of PGS Catalog.

This table describes details of how the current PGS Catalog conforms to FAIR data principles. For the purposes of this table the Score constitutes the data (e.g. variants, effect weights and alleles), and is linked to metadata (Samples, Performance Metrics, Publications) describing it.

Core FAIR	FAIR principle	PGS Catalog indicator
principle		
Findable	F1. (meta)data are assigned a globally unique and persistent identifier F2. data are described with rich	Each polygenic score is assigned a unique identifier (e.g. PGS000018) that is linked to all relevant metadata and publication sources in the Catalog. The PGS identifier can be resolved externally through <u>IDENTIFIERS.org</u> (prefix: <i>pgs</i>) Polygenic scores included in the database are
	metadata (defined by R1 below)	well-described, both in terms of their provenance and ability to be applied. Details in <u>Supplemental</u> <u>Table 2</u> and on our website at: <u>http://www.pgscatalog.org/docs/</u>
	F3. metadata clearly and explicitly include the identifier of the data it describes	All metadata is linked to either a Polygenic Score (PGS), Sample Set (PSS), Performance Metric (PPM), or Publication (PGP) ID within the database. Ontology terms are described using the identifiers from the Experimental Factor Ontology. Publication sources are described using DOI and PMID.
		Scoring files for each PGS are labelled with their PGS ID, and finable with the metadata on our FTP (<u>http://ftp.ebi.ac.uk/pub/databases/spot/pgs/</u>) described here: <u>http://www.pgscatalog.org/downloads/</u>
	F4. (meta)data are registered or indexed in a searchable resource	The PGS Catalog is indexed at <u>FAIRsharing.org</u> (ID: <i>bsg-d001448</i>) and indexed by Google Search.
Accessible	A1. (meta)data are retrievable by their identifier using a standardized communications protocol	Metadata can be easily viewed on our web interface (<u>www.pgscatalog.org</u>) with visible download links for each Score.
		Scoring files and metadata can also be browsed and downloaded from our FTP site by PGS ID.
		The full Catalog can also be accessed using our REST API: https://www.pgscatalog.org/rest/.
	A1.1 the protocol is open, free, and universally implementable	Yes, the <u>www.pgscatalog.org</u> website is freely accessible to all.
	A1.2 the protocol allows for an authentication and authorization procedure, where necessary	Not applicable

	A2. metadata are accessible, even	Archived versions of the scoring files and
	when the data are no longer	metadata are stored for the complete database
	available	as well as individual scores on our FTP
		(http://ftp.ebi.ac.uk/pub/databases/spot/pgs/)
Interoperable	I1. (meta)data use a formal,	PGS metadata is distributed from our API using
	accessible, shared, and broadly	JSON formats, the REST API is documented
	applicable language for knowledge	using the OpenAPI Specification (OAS3;
	representation.	https://github.com/OAI/OpenAPI-
		Specification/blob/master/versions/3.0.2.md).
	I2: (Meta)data use vocabularies that	The PGS identifier can be resolved externally
	follow the FAIR principles	through IDENTIFIERS.org (prefix: pgs). The
		traits are consistently described with EFO terms.
	I3. (meta)data include qualified	Traits are represented using (represented using
	references to other (meta)data	ontology terms) associated with PGS are linked
		to the Experimental Factor Ontology (EFO)
		terms and include links to the EFO.
Reusable	R1. meta(data) are richly described	Polygenic scores included in the database are
	with a plurality of accurate and	well-described, both in terms of their provenance
	relevant attributes	and ability to be applied. Details in Supplemental
		Table 2 and on our website at:
		http://www.pgscatalog.org/docs/
	R1.1. (meta)data are released with a	All data are made available through EMBL-EBI's
	clear and accessible data usage	standard terms of use
	license	(https://www.ebi.ac.uk/about/terms-of-use/).
		PGS with different licenses and terms are
		shared openly, but clearly marked with their
		terms in the metadata and scoring file
		downloads, as well as beside any download
		links.
	R1.2. (meta)data are associated with	Each PGS and Performance Metric is linked to a
	detailed provenance	source Publication that can be accessed by
		either a digital object identifier (DOI) or PubMed
		ID (PMID).
	R1.3. (meta)data meet domain-	The PGS Catalog is consistent with Polygenic
	relevant community standards	Risk Score Reporting Standards (PRS-RS) ¹⁰

Supplementary Table 2. PGS Catalog reporting items.

This table describes the reporting items that can be captured for each of the data objects in the PGS Catalog.

PGS Catalog Data Objects	Reporting Item	Description	Comments
Publication (Identified by	PubMed ID (PMID)	PubMed Identification number	
PGP ID)	Digital Object Identifier (DOI)	The DOI of each publication is curated in addition to the PMID to allow unpublished work (e.g. pre-prints) to be added to the Catalog.	This information is extracted and annotated according to EuropePMC ¹ .
	Title	Title of the publication or preprint	
	Author(s)	List of publication authors, the first author is also extracted for a shorter display.	Publications are flagged if they are preprints (e.g. not undergone peer
	Journal	The name of the publication source.	review).
	Publication Date	Date of publication (with respect to the PMID or DOI upon DB upload).	
	Release Date	Date the publication was added to the PGS Catalog.	
Score (Identified by PGS ID)	Reported Trait	The author-reported trait (<i>e.g.</i> body mass index [BMI], or coronary artery disease) that the PGS has been developed to predict.	
	Mapped Trait(s)	The <u>Reported Trait</u> is mapped to Experimental Factor Ontology (EFO) terms and their respective identifiers by PGS Catalog curators. For more information about the ontology traits see the Trait object.	Linked to Ontology Term(s).
	PGS Name	This may be the name that the authors use to refer to the PGS, or a name that a curator has assigned to identify the score during the curation process (before a PGS ID has been given).	
	Original Genome Build	The version of the genome that the variants present in the PGS are associated with. Listed as NR (Not Reported) if unknown.	
	Number of Variants	Number of variants used to calculate the PGS. In the future this will include a more detailed description of the types of variants present.	
	Number of Variant Interaction Terms	Number of higher-order variant interactions included in the PGS.	
	PGS Development Method	The name or description of the method or computational algorithm used to develop the PGS.	
	PGS Development Details/Relevant Parameters	A description of the relevant inputs and parameters relevant to the PGS development method/process.	
	Contributing Samples: Source of Variant Associations (GWAS)	Samples used to define the variant associations/effect-sizes used in the PGS. These data are extracted from and linked to the NHGRI-EBI GWAS Catalog when a GWAS study ID (GCST) is provided.	Linked as a Sample object(s).

	Contributing Samples:	Samples used to develop or train the score			
	Score	(e.g. not used for variant discovery, and non-	Linked as a Sample		
	Development/Training	overlapping with the samples used to	object(s).		
		evaluate the PGS predictive ability).	,		
	Publication/Citation	A PGP ID links the PGS to the publication in	Linked as a Publication		
		which it was described.	object.		
	Terms and Licenses	The PGS Catalog distributes its data			
		according to EBI's standard Terms of use.			
		Some PGS have specific terms, licenses, or			
		restrictions (e.g. non-commercial use) that we			
		highlight in this field, if known			
	Release Date	Date the score was added to the PGS			
		Catalog.			
Ontology	Name	The trait label from the ontology.			
Term	Identifier	The Experimental Factor Ontology ID			
(Mapped traits	Identillei	(EFO ID) identifier to consistently refer to	This information is		
are identified		traits using the EFO, and to other resources	extracted and annotated		
by an EFO ID)		like the NHGRI-EBI GWAS Catalog.	according to		
	Description	Detailed description of the trait from EFO.	Experimental Factor		
			Ontology (EFO) ² using		
	Synonyms	Other names for the trait.	the Ontology Lookup		
	Mapped Term(s)	Includes references to terms in other	Service (OLS) ³ .		
	Mapped Term(s)	databases and ontologies (<i>e.g.</i> ICD9/ICD10,			
		MONDO, SNOMEDCT, <i>etc.</i>).			
Sample	Number	Number of individuals included in the sample			
Campie	of Individuals	Number of individuals included in the sample			
(Groups of	Number of Cases	Number of individuals with the phenotype of			
samples used		interest (<i>if dichotomous</i>).			
	Number of Controls	Number of individuals without the phenotype			
are given a		of interest (<i>if dichotomous</i>).			
Sample Set	Percent of participants	Percent individuals in the sample that are			
[PSS ID])	who are Male	identified as male.			
	Age of Study	A summary of the age	1		
	Participants	distribution(mean/median, range/confidence	Similar to the GWAS		
		intervals) of study participants.	Catalog sample		
	Broad Ancestral	Author reported ancestry is mapped to the	descriptions, and directly		
	Category	best matching ancestry category from the	extracted from the		
		NHGRI-EBI GWAS Catalog framework (Table	GWAS Catalog for		
		1, Morales et al. (2018)).	samples with a GCST		
	Ancestry	A more detailed description of sample	ID.		
		ancestry that usually matches the most			
		specific description described by the authors			
		(e.g. French, Chinese).			
	Country of recruitment	Author reported countries of recruitment (if			
		available).			
	Additional American	Any additional description asttime-life			
	Additional Ancestry	Any additional description not captured in the			
	Additional Ancestry Description	structured data (e.g. founder or genetically			
		structured data (e.g. founder or genetically isolated populations, or further description of			
	Description	structured data (e.g. founder or genetically isolated populations, or further description of admixed samples).			
	Description Age of Study	structured data (e.g. founder or genetically isolated populations, or further description of admixed samples). A summary (mean/median, range/confidence			
	Description Age of Study Participants	structured data (e.g. founder or genetically isolated populations, or further description of admixed samples). A summary (mean/median, range/confidence intervals) of study participants ages.			
	Description Age of Study Participants Participant Follow-up	structured data (e.g. founder or genetically isolated populations, or further description of admixed samples). A summary (mean/median, range/confidence intervals) of study participants ages. A summary of the follow-up time			
	Description Age of Study Participants	structured data (e.g. founder or genetically isolated populations, or further description of admixed samples). A summary (mean/median, range/confidence intervals) of study participants ages. A summary of the follow-up time (mean/median, range/confidence intervals)			
	Description Age of Study Participants Participant Follow-up	structured data (e.g. founder or genetically isolated populations, or further description of admixed samples). A summary (mean/median, range/confidence intervals) of study participants ages. A summary of the follow-up time			

	Phenotype Definitions	A description of how the phenotype was	
	and Methods	measured or defined (e.g. ICD codes used to	
		identify cases/phenotypes in EHR data).	
	Cohort(s)	A list of cohorts that collected the samples.	The initial list of common cohorts used in genetics studies that seeded these annotations is from Mills & Rahal. Communications Biology (2019) ¹¹
	Additional Sample/Cohort Information	Any additional description about the samples and what they were used for that is not captured by the structured categories (e.g. sub-cohort information).	
Performance Metrics	Evaluated Score		Linked as a Score object
(Identified by a PPM ID)	Evaluated Samples	ID that links to the samples the displayed PGS evaluated.	Linked as a Sample object(s).
			Samples used in evaluations are given a Sample Set (PSS ID) so that PGS evaluated on the exact same samples can be extracted from the Catalog.
	Trait	This field displays both the <u>Reported</u> and <u>Mapped Traits</u> . The reported trait often corresponds to the test set names reported in the publication, or more specific aspects of the phenotype being tested (e.g. if the disease cases are incident vs. recurrent events).	Can be linked to a Trait object.
	Reported Metric: PGS Effect Size Reported Metric: PGS Classification Metrics Reported Metric: Other	Standardised effect sizes, per standard deviation [SD] change in PGS. Examples include regression coefficients (betas) for continuous traits, Odds ratios (OR) and/or Hazard ratios (HR) for dichotomous traits depending on the availability of time-to-event data. Examples include the Area under the Receiver Operating Characteristic (AUROC) or Harrell's C-index (Concordance statistic). Metrics that do not fit into the structured	The reported values of the performance metrics are all reported similarly (e.g. the estimate is recorded along with the 95% confidence interval
		categories. Examples include: R2 (proportion of the variance explained), reclassification metrics, p-values from association tests, binned comparisons of PGS risk (e.g. odds ratio of disease risk in the top vs. bottom decile of score).	(if supplied)
	Covariates Included in PGS Model	List of covariates used in the prediction model to evaluate the PGS. Examples include: age, sex, smoking habits, etc.	
	Other Relevant Information	Any other information relevant to the understanding of the performance metrics.	Linkad as a Dat V. C.
	Source	ID that links to the publication where the performance metrics were reported.	Linked as a Publication object.

Supplementary Table 3. UKB Benchmarking cohort description and results.

Cohort age and sex demographics broken down by colorectal cancer case/control status and participant ancestry. The distribution (mean and standard deviation [SD]) of each standardised PGS in colorectal cancer cases is also given, along with its effect size (Hazard Ratio; HR), citation and number of variants included in the PGS; the distribution of each PGS in controls is zero-mean and unit-variance.

		Europe	ean	South	Asian	African	Ancestry			
		Cases	Controls	Cases	Controls	Cases	Controls			
	Cohort Demographics									
N		5188 (1.28%)	404065	31 (0.51%)	6055	51 (0.86%)	5933			
N (Female	e)	2213	218990	18	2751	30	3503			
N (Male))	2975	185075	13	3304	21	2430			
Mean age at recrui	tment (SD)	61.97 (6.15)	57.35 (8.00)	57.87 (7.93)	53.63 (8.45)	58.34 (8.35)	52.87 (8.06)			
Mean event/censori	ng age (SD)	61.47 (8.66)	64.51 (7.98)	58.38 (8.15)	60.43 (8.42)	56.88 (9.96)	59.57 (8.07)			
		PGS dis	tribution and	effect size						
PGS0000	<u>55</u>	Case PGS Di			Distribution =	Case PGS Distribution				
Schmit SL et al. J Natl Cancer Inst (2019) ¹²	76	0.32 (1 HR = 1.38 [1		-0.10 (0.85) HR = 0.89 [0.63 - 1.28]		0.17 (1.07) HR = 1.2 [0.91 - 1.58]				
PGS0000	74	Case PGS Distribution = 0.30 (1.01)		Case PGS Distribution = 0.18 (0.71)		Case PGS Distribution -0.02 (0.96)				
Graff RE et al. bioRxiv (2020) ¹³	103	HR = 1.35 [1			[0.85 - 1.74]		0.75 - 1.31]			
PGS00014	<u>46</u>	Case PGS Distribution = 0.26 (1.00) HR = 1.3 [1.27 - 1.34]		Case PGS Distribution =		Case PGS Distribution				
Hsu L et al. Gastroenterology (2015) ¹⁴	27			-0.06 (0.88) HR = 0.95 [0.67 - 1.35]		0.06 (1.01) HR = 1.06 [0.8 - 1.4]				
PGS00014	<u>47</u>	Case PGS Dis		Case PGS Distribution =			Distribution =			
Ibáñez-Sanz G et al. Sci Rep (2017) ¹⁵	21	0.18 (1 HR = 1.2 [1.			(0.97) [1.02 - 2.08]		(0.89)).84 - 1.44]			

PGS000148		Case PGS Distribution =		Case PGS Distribution =
Jeon J et al. Gastroenterology (2018) ¹⁶	63	0.28 (1.00) HR = 1.32 [1.28 - 1.35]	-0.03 (0.84) HR = 0.96 [0.67 - 1.38]	0.15 (1.03) HR = 1.17 [0.89 - 1.54]
PGS000149		Case PGS Distribution =		Case PGS Distribution =
Smith T et al. Br J Cancer (2018) ¹⁷	41	0.25 (1.00) HR = 1.28 [1.25 - 1.32]	0.01 (0.95) HR = 1.02 [0.71 - 1.47]	0.10 (1.10) HR = 1.12 [0.85 - 1.48]
PGS000150		Case PGS Distribution =		Case PGS Distribution =
Weigl K et al. Gastroenterology (2018) ¹⁸	48	0.26 (1.00) HR = 1.3 [1.27 - 1.34]	0.05 (0.91) HR = 1.07 [0.75 - 1.53]	0.08 (0.95) HR = 1.09 [0.83 - 1.44]
PGS000151		Case PGS Distribution = 0.00 (1.02)	Case PGS Distribution = -0.10 (0.88)	Case PGS Distribution = 0.12 (1.07)
Xin J et al. Gene (2018) ¹⁹	14	HR = 1 [0.97 - 1.03]	HR = 0.9 [0.63 - 1.28]	HR = 1.13 [0.87 - 1.48]
PGS000154		Case PGS Distribution =		Case PGS Distribution =
Shi Z et al. Cancer Med (2019) ²⁰	30	0.20 (1.00) HR = 1.23 [1.2 - 1.26]	-0.06 (0.93) HR = 0.94 [0.66 - 1.34]	0.21 (1.05) HR = 1.25 [0.95 - 1.64]

Supplementary References

- Levchenko, M. *et al.* Europe PMC in 2017. *Nucleic Acids Res.* 46, D1254–D1260 (2018).
- Malone, J. *et al.* Modeling sample variables with an Experimental Factor Ontology. *Bioinformatics* 26, 1112–1118 (2010).
- Jupp, S., Burdett, T., Leroy, C. & Parkinson, H. E. A new Ontology Lookup Service at EMBL-EBI. in *Proceedings of the 8th Semantic Web Applications and Tools for Life Sciences International Conference, Cambridge UK, December 7-10, 2015* (eds. Malone, J., Stevens, R., Forsberg, K. & Splendiani, A.) **1546,** 118–119 (CEUR-WS.org, 2015).
- 4. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data.

Nature 562, 203–209 (2018).

- Saunders, C. L. *et al.* External validation of risk prediction models incorporating common genetic variants for incident colorectal cancer using UK Biobank. *Cancer Prev Res (Phila Pa)* (2020). doi:10.1158/1940-6207.CAPR-19-0521
- 6. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
- 7. Davidson-Pilon, C. lifelines: survival analysis in Python. JOSS 4, 1317 (2019).
- Seabold, S. & Perktold, J. Statsmodels: Econometric and Statistical Modeling with Python. in *Proceedings of the 9th Python in Science Conference* 92–96 (SciPy, 2010). doi:10.25080/Majora-92bf1922-011
- 9. Inouye, M. *et al.* Genomic risk prediction of coronary artery disease in 480,000 adults: implications for primary prevention. *J. Am. Coll. Cardiol.* **72**, 1883–1893 (2018).
- Wand, H. *et al.* Improving reporting standards for polygenic scores in risk prediction studies. *medRxiv* (2020). doi:10.1101/2020.04.23.20077099
- Mills, M. C. & Rahal, C. A scientometric review of genome-wide association studies. *Commun. Biol.* 2, 9 (2019).
- Schmit, S. L. *et al.* Novel common genetic susceptibility loci for colorectal cancer. *J Natl Cancer Inst* **111**, 146–157 (2019).
- Graff, R. E. *et al.* Cross-Cancer Evaluation of Polygenic Risk Scores for 17 Cancer
 Types in Two Large Cohorts. *BioRxiv* (2020). doi:10.1101/2020.01.18.911578
- Hsu, L. *et al.* A model to determine colorectal cancer risk using common genetic susceptibility loci. *Gastroenterology* **148**, 1330–9.e14 (2015).
- Ibáñez-Sanz, G. *et al.* Risk Model for Colorectal Cancer in Spanish Population Using Environmental and Genetic Factors: Results from the MCC-Spain study. *Sci. Rep.* 7, 43263 (2017).
- Jeon, J. *et al.* Determining risk of colorectal cancer and starting age of screening based on lifestyle, environmental, and genetic factors. *Gastroenterology* **154**, 2152-2164.e19 (2018).

- Smith, T., Gunter, M. J., Tzoulaki, I. & Muller, D. C. The added value of genetic information in colorectal cancer risk prediction models: development and evaluation in the UK Biobank prospective cohort study. *Br. J. Cancer* **119**, 1036–1039 (2018).
- 18. Weigl, K. *et al.* Genetic risk score is associated with prevalence of advanced neoplasms in a colorectal cancer screening population. *Gastroenterology* **155**, 88-98.e10 (2018).
- 19. Xin, J. *et al.* Evaluating the effect of multiple genetic risk score models on colorectal cancer risk prediction. *Gene* **673**, 174–180 (2018).
- Shi, Z. *et al.* Systematic evaluation of cancer-specific genetic risk score for 11 types of cancer in The Cancer Genome Atlas and Electronic Medical Records and Genomics cohorts. *Cancer Med.* 8, 3196–3205 (2019).