

1 **Hidden ‘risk’ in polygenic scores: clinical use today could exacerbate health**
2 **disparities**

3

4 Alicia R. Martin^{1,2,3}, Masahiro Kanai^{1,2,3,4,5}, Yoichiro Kamatani^{5,6}, Yukinori Okada^{5,7,8},
5 Benjamin M. Neale^{1,2,3}, Mark J. Daly^{1,2,3,9}

6

7 ¹ Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston,
8 MA 02114, USA

9 ² Program in Medical and Population Genetics, Broad Institute of Harvard and MIT,
10 Cambridge, MA 02142, USA

11 ³ Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT,
12 Cambridge, MA 02142, USA

13 ⁴ Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02115,
14 USA

15 ⁵ Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences,
16 Yokohama 230-0045, Japan

17 ⁶ Kyoto-McGill International Collaborative School in Genomic Medicine, Graduate
18 School of Medicine, Kyoto University, Kyoto 606-8507, Japan

19 ⁷ Department of Statistical Genetics, Osaka University Graduate School of Medicine,
20 Suita 565-0871, Japan

21 ⁸ Laboratory of Statistical Immunology, Immunology Frontier Research Center (WPI-
22 IFReC), Osaka University, Suita 565-0871, Japan

23 ⁹ Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki,
24 Finland

25

26 **Abstract**

27

28 Polygenic risk scores have the potential to improve health outcomes for a variety of
29 complex diseases and are poised for clinical translation, driven by the low cost of
30 genotyping (<\$50 per person), the ability to predict genetic risk of many diseases with a
31 single test, and the dramatically increasing scale and power of genetic studies that aid
32 prediction accuracy. However, the major *ethical* and *scientific* challenge surrounding
33 clinical implementation is the observation that they are currently of far greater predictive
34 value in European ancestry individuals than others. The better performance of such risk
35 scores in European populations is an inescapable consequence of the heavily biased
36 makeup of genome-wide association studies, with an estimated 79% of participants in
37 all these existing studies being of European descent. Empirically, polygenic risk scores
38 perform far better in European populations, with prediction accuracy reduced by
39 approximately 2- to 5-fold in East Asian and African descent populations, respectively.
40 This highlights that—unlike specific clinical biomarkers and prescription drugs, which
41 may individually work better in some populations but do not ubiquitously perform far
42 better in European populations—clinical uses of prediction today would systematically
43 afford greater improvement to European descent populations. Early diversifying efforts,
44 however, show promise in levelling this vast imbalance, even when non-European
45 sample sizes are considerably smaller than the best-powered studies to date. Polygenic

46 risk scores provide a new opportunity to improve health outcomes for many diseases in
47 all populations, but to realize this full potential equitably, we must prioritize greater
48 inclusivity of diverse participants in genetic studies and open access to resulting
49 summary statistics to ensure that health disparities are not increased for those already
50 most underserved.

51

52 **Keywords:** health disparities, genetic risk prediction, polygenic risk scores, diversity,
53 population genetics, statistical genetics

54

55 Polygenic risk scores (PRS), which predict complex traits using genetic data, are of
56 burgeoning interest to the clinical community as researchers demonstrate their growing
57 power to improve clinical care, genetic studies of a wide range of phenotypes increase
58 in size and power, and genotyping costs plummet to less than US\$50. Many earlier
59 criticisms of limited prediction power are now recognized to have been chiefly an issue
60 of insufficient sample size, which is no longer the case for many outcomes ¹. For
61 example, polygenic risk scores alone already predict breast and prostate cancer risk in
62 European descent patients more accurately than current clinical models ²⁻⁴. Additionally,
63 integrated models of PRS together with other lifestyle and clinical factors have enabled
64 clinicians to more accurately quantify the risk of heart attack for patients; consequently,
65 they have more effectively targeted the reduction of LDL cholesterol and by extension
66 heart attack by prescribing statins to patients at the greatest overall risk of
67 cardiovascular disease ⁵⁻⁹. Promisingly, return of genetic risk of complex disease to at-
68 risk patients does not induce significant self-reported negative behavior or psychological

69 function, and some potentially positive behavioral changes have been detected ¹⁰. While
70 we share enthusiasm about the potential of PRS to improve health outcomes through
71 their eventual routine implementation as clinical biomarkers, we consider the consistent
72 observation that they are currently of far greater predictive value in individuals of recent
73 European descent than in others to be the major *ethical* and *scientific* challenge
74 surrounding clinical translation and, at present, the most critical limitation to genetics in
75 precision medicine. The scientific basis of this imbalance has been demonstrated in
76 population genetics simulations, theoretically, and empirically across many traits and
77 diseases ¹¹⁻²².

78
79 All studies to date using well-powered genome-wide association studies (GWAS) to
80 assess the predictive value of PRS in European and non-European descent populations
81 have made a consistent observation: PRS predict individual risk far more accurately in
82 Europeans than non-Europeans. In complex traits including height, body mass index
83 (BMI), educational attainment, schizophrenia, and major depression, existing PRS
84 computed with the largest available GWAS results predict outcomes far more accurately
85 in new samples of European-descent than they do in non-Europeans, with the clearest
86 study examples in East Asians and African Americans ^{15,16,18-24}. Rather than chance or
87 biology, this is a predictable consequence of the fact that the genetic discovery efforts to
88 date heavily underrepresent non-European populations globally. The correlation
89 between true and genetically predicted phenotypes decays with genetic divergence
90 from the makeup of the discovery GWAS, meaning that the accuracy of polygenic
91 scores in different populations is highly dependent on the study population

92 representation in the largest existing ‘training’ GWAS. Here, we document study biases
93 that underrepresent non-European populations in current GWAS, and explain the
94 fundamental concepts contributing to reduced phenotypic variance explained with
95 increasing genetic divergence from populations included in GWAS.

96

97 **Predictable basis of disparities in polygenic risk score accuracy**

98 Poor generalizability of genetic studies across populations arises from the
99 overwhelming abundance of European descent studies and dearth of well-powered
100 studies in globally diverse populations²⁵⁻²⁸. According to the GWAS catalog, ~79% of all
101 GWAS participants are of European descent despite making up only 16% of the global
102 population (**Figure 1**). This is especially problematic as previous studies have shown
103 that Hispanic/Latino and African American studies have contributed an outsized number
104 of associations relative to studies of similar sizes in Europeans²⁷. More concerning,ly,
105 the fraction of non-European individuals in GWAS has stagnated or declined since late
106 2014 (**Figure 1**), suggesting that we are not on a trajectory to correct this imbalance.
107 These numbers provide a composite metric of study availability, accessibility, and use—
108 i.e., cohorts that have been included in numerous GWAS are represented multiple
109 times, which may disproportionately include cohorts of European descent. Whereas the
110 average sample sizes of GWAS in Europeans continue to grow though, they have
111 stagnated and remain several-fold smaller in other populations (**Figure S1**).

112

113 The relative sample compositions of GWAS result in highly predictable disparities in
114 prediction accuracy; statistical and population genetics theory predicts that genetic risk

115 prediction accuracy will decay with increasing genetic divergence between the original
116 GWAS sample and target of prediction, a function of population history^{13,14}. This pattern
117 can be attributed to several statistical observations which we detail below: 1) GWAS
118 favor the discovery of genetic variants that are common in the study population; 2)
119 linkage disequilibrium (LD) differentiates marginal effect size estimates for highly
120 polygenic traits across populations, even when causal variants are the same; and 3)
121 environment and demography differ across populations, which may interact to drive
122 differential forces of natural selection that in turn drive differences in causal genetic
123 architecture*. Of note, the first two of these degrade prediction performance across
124 populations substantially even when there exist no biological, environmental, or
125 diagnostic differences.

126

127 *Common discoveries and low-hanging fruit*

128 First, the power to discover an association in a genetic study depends on the effect size
129 and frequency of the variant²⁹. This power dependence means that the most significant
130 associations tend to be more common in the populations in which they are discovered
131 than in other populations^{13,30}. For example, GWAS catalog variants are on average
132 more common in European populations compared to East Asian and African
133 populations (**Figure 2B**), an observation not representative of genomic variants at large.
134 Understudied populations offer low-hanging fruit for genetic discovery because variants
135 that are common in these groups but rare or absent in European populations could not

* We define the causal genetic architecture as the true effects (typically written as β in the GWAS literature) of variants that impact a phenotype that would be identified in a population of infinite sample size. Unlike effect size estimates (usually written as $\hat{\beta}$), true effects are typically modeled as invariant with respect to LD and allele frequency differences across populations.

136 be discovered even with very large European sample sizes. Some examples include
137 *SLC16A11* and *HNF1A* associations with type II diabetes in Latino populations, *APOL1*
138 associations with end-stage kidney disease, and associations with prostate cancer in
139 African descent populations³¹⁻³⁴. If we assume that causal genetic variants have an
140 equal effect across all populations—an assumption with some empirical support that
141 offers the best case scenario for transferability³⁵⁻⁴⁰—Eurocentric GWAS biases mean
142 that variants that are common in European populations are preferentially discovered
143 and associated with risk, and thus account for a larger fraction of the variance in
144 polygenic risk¹³. Furthermore, imputation reference panels share the same study
145 biases as in GWAS⁴¹, creating challenges for imputing sites that are rare in European
146 populations but common elsewhere when the catalog of non-European haplotypes is
147 substantially smaller. These issues are insurmountable through statistical methods
148 alone¹³, but rather motivate substantial investments in more diverse populations to
149 produce similar-sized GWAS of biomedical phenotypes as well as sequenced reference
150 panels in other populations.

151

152 *Linkage disequilibrium*

153 Second, the correlation structure of the human genome, i.e. LD, varies across
154 populations due to demographic history (**Figure 2A,C-E**). These LD differences in turn
155 drive differences in effect size estimates (i.e. predictors) from GWAS across
156 populations, even when causal effects are the same. (Mathematically, the marginal
157 GWAS estimate $\hat{\beta}_j = \sum_{k=1}^m r_{j,k} \beta_k + \epsilon_j$, where $\hat{\beta}_j$ are effect size estimates at SNP j , $r_{j,k}$ is
158 pairwise SNP LD between SNPs j and k , β_k is the causal SNP effect at nearby SNP k ,

159 and ϵ is residual error from bias or noise. More simply, when causal effects are the same
160 across populations, effect size *estimates* at SNPs tagging these causal variants from which we
161 construct predictors will differ across populations in proportion to LD between tagging and
162 causal SNP pairs.) While differences in effect size estimates due to LD differences may
163 typically be small for most regions of the genome (**Figure 2C-E**), PRS sum across these
164 effects, also aggregating these population differences. Statistical methods that account
165 for LD differences across populations may help improve risk prediction accuracy within
166 each population. While empirical studies suggest that causal effect sizes tend to be
167 shared³⁵⁻⁴⁰, it may not be feasible to fine-map most variants to a single locus to solve
168 issues of low generalizability (i.e. using causal rather than correlated effect size
169 estimates), even with very large GWAS. This is because complex traits are highly
170 polygenic, meaning most of our prediction power comes from small effects that do not
171 meet genome-wide significance and/or cannot be fine-mapped, even in many of the
172 best-powered GWAS to date⁴².

173

174 *History, selection, the environment, and complex interactions*

175 Lastly, other environmental, demographic, and cohort considerations may further
176 worsen prediction accuracy differences across populations in less predictable ways.
177 GWAS ancestry study biases and LD differences across populations are extremely
178 challenging to address, but these issues actually make many favorable assumptions
179 that all causal loci have the same impact and are under equivalent selective pressure in
180 all populations. In contrast, other effects on polygenic adaptation or risk scores such as
181 long-standing environmental differences across global populations that have resulted in
182 differing responses of natural selection can impact populations differently based on their

183 unique histories. Additionally, residual uncorrected population stratification may impact
184 risk prediction accuracy across populations, but the magnitude of its effect is currently
185 unclear. These effects are particularly challenging to disentangle, as has clearly been
186 demonstrated for height, where evidence of polygenic adaptation and/or its relative
187 magnitude is under question^{43,44}. Comparisons of geographically stratified phenotypes
188 like height across populations with highly divergent genetic backgrounds and mean
189 environmental differences, such as differences in resource abundance during
190 development across continents, are especially prone to confounding from correlated
191 environmental and genetic divergence^{43,44}. This residual stratification can lead to over-
192 predicted differences across geographical space⁴⁵.

193
194 Related to stratification, most polygenic scoring methods do not explicitly address
195 recent admixture and none consider recently admixed individuals' unique local mosaic
196 of ancestry—further methods development in this space is needed. Furthermore,
197 comparing PRS across environmentally stratified cohorts, such as in some biobanks
198 with healthy volunteer effects versus disease study datasets or hospital-based cohorts,
199 requires careful consideration of technical differences, collider bias, as well as variability
200 in baseline health status among studies. It is also important to consider differences in
201 clinical definition of the phenotypes and heterogeneous constitution of sub-phenotypes
202 among countries.

203
204 Another consideration is the fact that, like other existing biomarkers, the predictive utility
205 of PRS may change as a function of age, consistent with age-dependent heritability for

206 some traits⁴⁶. For example, increasing age is associated with higher risk of coronary
207 artery disease, and higher PRS accelerate this increased risk⁴⁷. Consequently, the age
208 of intervention e.g. with statins needs to be evaluated in aggregate with other clinical
209 risk factors that change over time. Autism spectrum disorder and schizophrenia also
210 have a genetic basis with differing developmental trajectories; their shared genetic
211 influences decrease with age, whereas the genetic overlap between schizophrenia and
212 social communication difficulties persists with age⁴⁸. Further work on how prediction
213 accuracy varies as a function of age across phenotypes is needed.

214
215 Differences in environmental exposure, gene × gene interactions, gene × environment
216 interactions, historical population size dynamics, and other factors will further limit
217 generalizability for genetic risk scores in an unpredictable, trait-specific fashion^{49,50}.

218 Complex traits do not behave in a genetically deterministic manner, with some
219 environmental factors dwarfing individual genetic effects, creating outsized issues of
220 comparability across globally diverse populations; among psychiatric disorders for
221 example, whereas schizophrenia has a nearly identical genetic basis across East
222 Asians and Europeans ($r_g=0.98$)⁴⁰, substantially different rates of alcohol use disorder
223 across populations is partially explained by differences in availability and genetic
224 differences impacting alcohol metabolism⁵¹. While non-linear genetic factors explain
225 little variation in complex traits beyond a purely additive model⁵², some unrecognized
226 nonlinearities and gene × gene interactions can also induce genetic risk prediction
227 challenges, as pairwise interactions are likely to vary more across populations than
228 individual SNPs. Mathematically, we can simplistically think of this in terms of a two-

229 SNP model, in which the sum of two SNP effects is likely to explain more phenotypic
230 variance than the product of the same SNPs. Some machine learning approaches may
231 thus modestly improve genetic prediction accuracy beyond current approaches for
232 some phenotypes⁵³, but these approaches are most likely to improve prediction
233 accuracy for atypical traits with simpler architectures, known interactions, and poor
234 prediction generalizability across populations, such as skin pigmentation⁵⁴.

235

236 **Limited generalizability of genetic prediction across diverse populations**

237 Previous work has assessed prediction accuracy across diverse populations in several
238 traits and diseases for which GWAS summary statistics are available. These
239 assessments are becoming increasingly feasible with the growth and public availability
240 of global biobanks for quantitative traits as well as diversifying priorities from funding
241 agencies^{55,56}. As of yet, multi-ethnic work has been slow in most disease areas⁵⁷,
242 limiting even the opportunity to assess prediction utility in non-European cohorts.
243 Nonetheless, we have assembled prediction accuracy statistics from several studies
244 using the largest European GWAS to predict several phenotypes in target European
245 and non-European cohorts. For example, multiple schizophrenia studies consistently
246 predicted risk on average 2.2-fold worse in East Asians relative to Europeans, (i.e.
247 $\mu=0.46$, $\sigma=0.06$), using summary statistics from a Eurocentric GWAS^{15,18} (**Figure S2**),
248 despite the fact that there is no significant genetic heterogeneity in schizophrenia
249 between the two populations⁴⁰. This finding is even more pronounced in African
250 Americans, consistent with higher genetic divergence from Europeans than between
251 Europeans and East Asians³⁰. Across several phenotypes with a range of genetic

252 architectures in which empirical evaluations were available, including BMI, educational
253 attainment, height, and schizophrenia, prediction accuracy using European GWAS
254 summary statistics was on average 4.5-fold less accurate in African Americans than in
255 Europeans (i.e. $\mu=0.22$, $\sigma=0.09$, **Figure S2**)^{15,16,19-22}. By extension, prediction accuracy
256 is expected to be even lower in African Americans with higher than average African
257 ancestry or among populations with greater divergence from Europeans (e.g. some
258 southern African populations). These enormous disparities are not simply
259 methodological issues, as various approaches (e.g. pruning and threshold versus
260 LDpred) and accuracy metrics (R^2 for quantitative traits and various pseudo- R^2 metrics
261 for binary traits) illustrate this consistently poorer performance in populations distinct
262 from the discovery sample across a range of polygenic traits (**Table S4**). We assessed
263 how prediction accuracy decayed across globally diverse populations for 17
264 anthropometric and blood panel traits in the UK Biobank (UKBB) when using European-
265 derived summary statistics (**Methods**); consistent with previous studies, we find that
266 relative to European prediction accuracy, genetic prediction accuracy was far lower in
267 other populations (**Figure 3**, notably 1.6-fold lower in Hispanic/Latino Americans, 1.7-
268 fold lower in South Asians, 2.5-fold lower in East Asians, and 4.9-fold lower in Africans
269 on average).

270

271 **Prioritizing diversity shows early promise for polygenic prediction**

272 Early diversifying GWAS efforts have been especially productive for informing on these
273 questions surrounding risk prediction. Rather than varying the prediction target dataset,
274 some GWAS in diverse populations have increased the scale of non-European

275 summary statistics and also varied the study dataset in multi-ethnic PRS studies. For
276 example, a BioBank Japan (BBJ) GWAS study (N=158,284) showed that compared to a
277 2× larger European GWAS (N=322,154), the variance in BMI explained in an
278 independent Japanese cohort with Japanese GWAS summary statistics was on
279 average 1.5-fold greater than with European GWAS summary statistics ($R^2=0.154$ vs
280 0.104 at $p < 0.05$, respectively)²³. Relatedly, an East Asian schizophrenia study
281 (N=22,778 cases and 35,362 controls) showed that compared to an effectively 3× larger
282 European study, prediction accuracy in East Asians was on average 1.3-fold higher
283 than with European summary statistics (liability $R^2=0.029$ vs 0.022 , respectively)⁴⁰.
284 Thus, even when studies in non-European populations are only a fraction the size of the
285 largest European study, they are likely to have disproportionate value for predicting
286 polygenic traits in other individuals of similar ancestry.

287

288 Given this background, we performed a systematic evaluation of polygenic prediction
289 accuracy across 17 quantitative anthropometric and blood panel traits and five disease
290 endpoints in British and Japanese individuals^{23,58,59} by performing GWAS with the exact
291 same sample sizes in each population. We symmetrically demonstrate that prediction
292 accuracy is consistently higher with GWAS summary statistics from ancestry-matched
293 summary statistics (**Figure 4, Figure S8–12**). Keeping in mind issues of comparability
294 described above, we note that BBJ is a hospital-based disease-ascertained cohort,
295 whereas UKBB is a healthier than average population-based cohort; thus, differences in
296 observed heritability among these cohorts (rather than among populations) due to
297 differences in phenotype precision likely explain lower prediction accuracy from the BBJ

298 GWAS summary statistics for anthropometric and blood panel traits, but higher
299 prediction accuracy for five ascertained diseases (**Table S5**). Indeed, other East Asian
300 studies have estimated higher heritability for some quantitative traits than BBJ using the
301 same methods, such as for height ($h^2 = 0.48 \pm 0.04$ in Chinese women)⁶⁰. Some
302 statistical fluctuations in the relative differences in prediction accuracy across
303 populations are likely driven by differences in heritability measured in each population
304 and/or trans-ethnic genetic correlation[†] (**Figure S4-7, Table S5-8**). These correlation
305 estimates indicate that effect sizes were mostly highly correlated across ancestries, with
306 a few traits that were somewhat lower than expected (e.g. height and BMI, with
307 $\rho_{ge} = 0.69$ and 0.75 , respectively); effect sizes can be different across populations due to
308 natural selection differences, gene-gene interactions, gene-environment interactions,
309 statistical noise, and/or other phenomena⁶¹. Prediction accuracy was far lower in
310 individuals of African descent in the UK Biobank (**Figure S10 and S13**) using GWAS
311 summary statistics from either European or Japanese ancestry individuals, consistent
312 with reduced prediction accuracy with increasing genetic divergence (**Figures 3 and 4**).
313 These population studies demonstrate the power and utility of increasingly diverse
314 GWAS for prediction, especially in populations of non-European descent.

315
316 While many other traits and diseases have been studied in multi-ethnic settings, few
317 have reported comparable metrics of prediction accuracy across populations.

318 Cardiovascular research, for example, has led the charge towards clinical translation of
319 PRS¹. This enthusiasm is driven by observations that a polygenic burden of LDL-

[†] Trans-ethnic genetic correlation compares the estimated correlation of common variant effect sizes at SNPs common in two populations, here using Popcorn⁶¹.

320 increasing SNPs can confer monogenic-equivalent risk of cardiovascular disease, with
321 polygenic scores improving clinical models for risk assessment and statin prescription
322 that can reduce coronary heart disease and improve healthcare delivery efficiency^{5,6,8}.
323 However, many of these studies have been conducted exclusively in European descent
324 populations, with few studies rigorously evaluating population-level applicability to non-
325 Europeans. Those existing findings indeed demonstrate a large reduction in prediction
326 utility in non-European populations¹¹, though often with comparisons of odds ratios
327 among arbitrary breakpoints in the risk distribution that make comparisons across
328 studies challenging. To better clarify how polygenic prediction will be deployed in a
329 clinical setting with diverse populations, more systematic and thorough evaluations of
330 the utility of PRS within and across populations for many complex traits are still needed.
331 These evaluations would benefit from rigorous polygenic prediction accuracy
332 evaluations, especially for diverse non-European patients⁶²⁻⁶⁴.

333

334 **Translational genetic prediction may uniquely exacerbate disparities**

335 Our impetus for raising these statistical issues limiting the generalizability of PRS across
336 population stems from our concern that, while they are legitimately clinically promising
337 for improving health outcomes for many biomedical phenotypes, they may have a larger
338 potential to raise health disparities than other clinical factors for several reasons. The
339 opportunities they provide for improving health outcomes means they inevitably will and
340 should be pursued in the near term, but we urge that a concerted prioritization to make
341 GWAS summary statistics easily accessible for diverse populations and a variety of
342 traits and diseases is imperative, even when they are a fraction the size of the largest

343 existing European datasets. *Individual* clinical tests, biomarkers, and prescription drug
344 efficacy may vary across populations in their utility, but are fundamentally informed by
345 the same underlying biology^{65,66}. Currently, guidelines state that as few as 120
346 individuals define reference intervals for clinical factors (though often smaller numbers
347 from only one subpopulation are used) and there is no clear definition of who is “normal”
348⁶⁵. Consequently, reference intervals for biomarkers can sometimes deviate
349 considerably by reported ethnicity⁶⁷⁻⁶⁹. Defining ethnicity-specific reference intervals is
350 clearly an important problem that can provide fundamental interpretability gains with
351 implications for some major health benefits (e.g. need for dialysis and development of
352 Type 2 diabetes based on ethnicity-specific serum creatinine and hemoglobin A1C
353 reference intervals, respectively)⁶⁸. Simply put, some biomarkers or clinical tests scale
354 directly with health outcomes independent of ancestry, and many others may have
355 distributional differences by ancestry but are equally valid after centering with respect to
356 a readily collected population reference.

357

358 In contrast, PRS are uniformly less useful in understudied populations due to
359 differences in genomic variation and population history^{13,14}. No analogous solution of
360 defining ethnicity-specific reference intervals would ameliorate health disparities
361 implications for PRS or fundamentally aid interpretability in non-European populations.
362 Rather, as we and others demonstrate, PRS are unique in that even with multi-ethnic
363 population references, these scores are fundamentally less informative in populations
364 more diverged from GWAS study cohorts.

365

366 The clinical use and deployment of genetic risk scores needs to be informed by the
367 issues surrounding tests that currently would unequivocally provide much greater
368 benefit to the subset of the world's population which is already on the positive end of
369 healthcare disparities[‡]. Conversely, African descent populations, which already endure
370 many of the largest health disparities globally, are often predicted marginally better, if at
371 all, compared to random (**Figure 4F**). They are therefore least likely to benefit from
372 improvements in precision healthcare delivery from genetic risk scores with existing
373 data due to human population history and study biases. This is a major concern globally
374 and especially in the U.S., which already leads other middle- and high-income countries
375 in both real and perceived healthcare disparities⁷⁰. Thus, we would strongly urge that
376 any discourse on clinical use of polygenic scores include a careful, quantitative
377 assessment of the economic and health disparities impacts on underrepresented
378 populations that might be unintentionally introduced by the use of PRS and raise
379 awareness about how to eliminate these disparities.

380

381 **How do we even the ledger?**

382 What can be done? The single most important step towards parity in PRS accuracy in
383 diverse populations is by vastly increasing the diversity of participants included and
384 analyzed in genetic studies. Regulatory protections against genetic discrimination are
385 necessary to accompany calls for more diverse studies; while some already exist in the
386 U.S., including for health insurance and employment opportunities via the Genetic
387 Information Nondiscrimination Act (GINA), stronger protections in these and other areas

[‡] To maximally benefit all populations, the largest existing GWAS results should be used. Downsampling European GWAS for the sake of parity results in worse predictors for all individuals.

388 globally will be particularly important for minorities and/or marginalized groups. An equal
389 investment in GWAS across all major ancestries and global populations is the most
390 obvious solution to generate a substrate for equally informative risk scores but is not
391 likely to occur any time soon absent a dramatic priority shift given the current imbalance
392 and stalled diversifying progress over the last five years (**Figure 1, Figure S1**). While it
393 may be challenging or in some cases infeasible to acquire sample sizes large enough
394 for PRS to be equally informative in all populations, some much-needed efforts towards
395 increasing diversity in genomics that support open sharing of GWAS summary data
396 from multiple ancestries are underway. Examples include the *All of Us* Research
397 Program, the Population Architecture using Genomics and Epidemiology (PAGE)
398 Consortium, as well as some disease-focused consortia, such as the T2D-genes and
399 Stanley Global initiatives on the genetics of type II diabetes and psychiatric disorders,
400 respectively[§]. Supporting data resources such as imputation panels, multi-ethnic
401 genotyping arrays, gene expression datasets from genetically diverse individuals, and
402 other tools are necessary to similarly empower these diverse studies to most effectively
403 improve predictive accuracy for all populations. The lack of supporting resources for
404 diverse ancestries creates financial challenges for association studies with limited
405 resources, e.g. raising questions about whether to genotype samples on GWAS arrays
406 that may favor European allele frequencies versus sequence samples, and how dense
407 of an array to choose or how deeply to sequence^{71,72}.

408

[§] For these diverse studies in the U.S. and globally, genetically determined ancestry and populations are important to delineate from self-identified race/ethnicity, as only controlling for the former can account for stratification of allele frequencies within a population.

409 Additional leading global efforts also provide easy unified access linking genetic, clinical
410 record, and national registry data in more homogeneous continental ancestries, such as
411 the UK Biobank, BioBank Japan, China Kadoorie Biobank, and Nordic efforts (e.g. in
412 Danish, Estonian, Finnish, and other integrated biobanks). Notably, some of these
413 biobanks such as UK Biobank have participants with considerable global genetic
414 diversity that enables multi-ethnic comparisons; although minorities from this cohort
415 provide the largest deeply phenotyped GWAS cohorts for several ancestries, these
416 individuals are often excluded in current statistical analyses in favor of single ancestries,
417 large sample sizes, and the simplicity afforded by genetic homogeneity. These
418 considerations notwithstanding, there are critical needs and challenges for expanding
419 the scale of genetic studies of heritable traits in diverse populations; this is especially
420 apparent in Africa where humans originated and retain the most genetic diversity, as
421 Africans are understudied but disproportionately informative for genetic analyses and
422 evolutionary history^{27,73}. The most notable investment here comes from the Human
423 Heredity and Health in Africa (H3Africa) Initiative, increasing genomics research
424 capacity in Africa through more than \$216 million in funding from the NIH (USA) and
425 Wellcome Trust (UK) for genetics research led by African investigators^{56,74}. The
426 increasing interest and scale of genetic studies in low- and middle-income countries
427 (LMICs) raises ethical and logistical considerations about data generation, access,
428 sharing, security, and analysis, as well as clinical implementation to ensure these
429 advances do not only benefit high-income countries. Frameworks such as the
430 H3ABioNet, a pan-African bioinformatics network designed to build capacity to enable

431 H3Africa researchers to analyze their data in Africa, provide cost-effective examples for
432 training local scientists in LMICs ⁷⁵.

433
434 The prerequisite data for dramatically increasing diversity also hypothetically exist in
435 several large-scale publicly funded datasets such as the Million Veterans Project and
436 Trans-Omics for Precision Medicine (TOPMed), but with problematic data access issues
437 in which even summary data from GWAS within and across populations are not publicly
438 shared. Existing GWAS consortia also need to carefully consider the granularity of
439 summary statistics they release, as finer scale continental ancestries and phenotypes in
440 large, multi-ethnic projects enable ancestry-matched analyses not possible with a single
441 set of summary statistics. While there is an understandable patient privacy balance to
442 strike when sharing individual-level data, GWAS summary statistics from all publicly
443 funded and as many privately funded projects as possible should be made easily and
444 publicly accessible to improve global health outcomes. Efforts to unify phenotype
445 definitions, normalization approaches, and GWAS methods among studies will also
446 improve comparability.

447
448 To enable progress towards parity, it will be critical that open data sharing standards be
449 adopted for all ancestries and for genetic studies of all sample sizes, not just the largest
450 European results. Locally appropriate and secure genetic data sharing techniques as
451 well as equitable technology availability will need to be adopted widely in Asia and
452 Africa as they are in Europe and North America, to ensure that maximum value is
453 achieved from existing and ongoing efforts that are being developed to help counter the

454 current imbalance. Simultaneously, ethical considerations require that research capacity
455 is increased in LMICs with simultaneous growth of diverse population studies to balance
456 the benefits of these studies to scientists and patients globally versus locally to ensure
457 that everyone benefits. Methodological improvements that better define risk scores by
458 accounting for population allele frequency, LD, and/or admixture differences
459 appropriately are underway and may help considerably but will not by themselves bring
460 equality. All of these efforts are important and should be prioritized not just for risk
461 prediction but more generally to maximize the use and applicability of genetics to inform
462 on the biology of disease. Given the acute recent attention on clinical use of PRS, we
463 believe it is paramount to recognize their potential to improve health outcomes for all
464 individuals and many complex diseases. Simultaneously, we as a field must address the
465 disparity in utility in an ethically thoughtful and scientifically rigorous fashion, lest we
466 inadvertently enable genetic technologies to contribute to, rather than reduce, existing
467 health disparities.

468

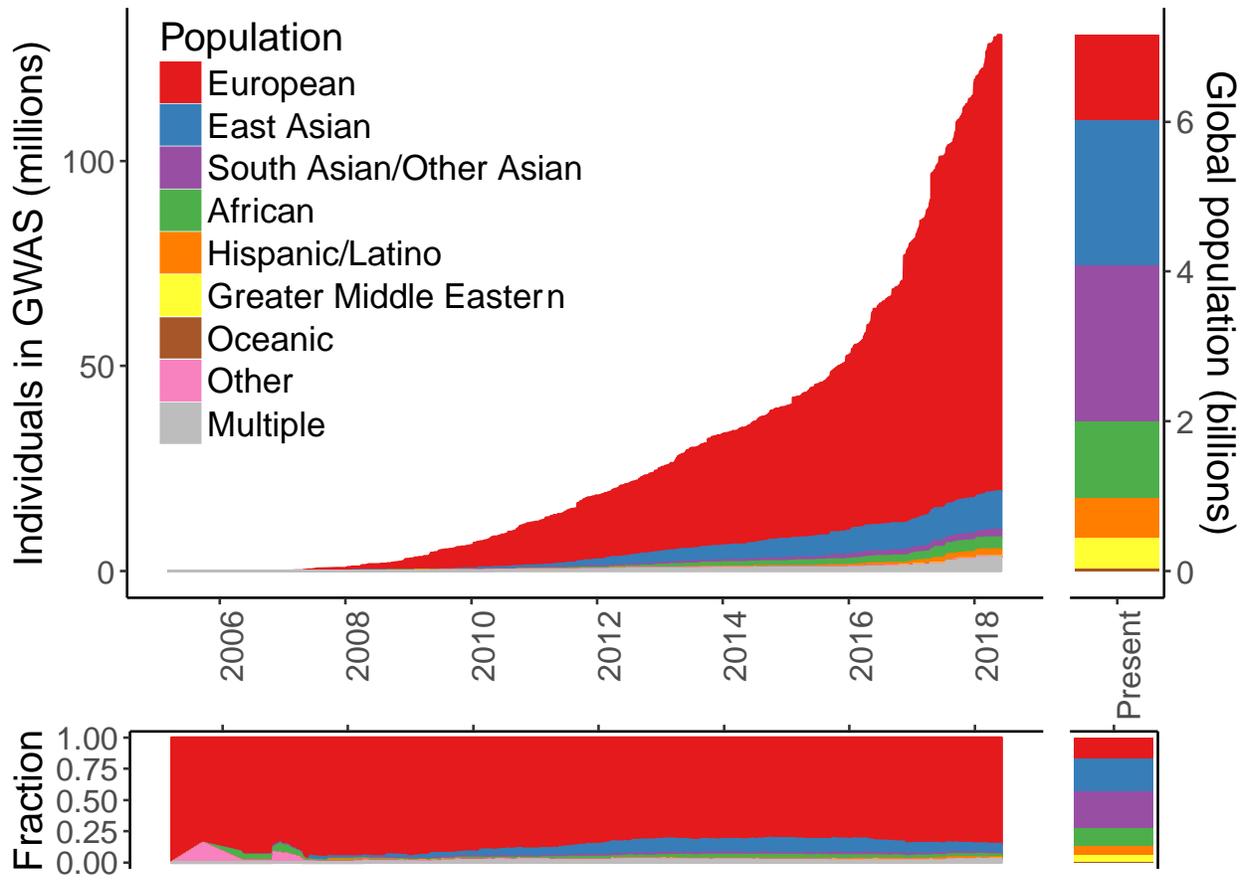
469 **Acknowledgments**

470 We thank Amit Khera for helpful discussions. We also thank Michiaki Kubo, Yoshinori
471 Murakami, Masato Akiyama, and Kazuyoshi Ishigaki for their support in the BioBank
472 Japan Project analysis. We are grateful to Steven Gazal for his help in calculating LD
473 scores. This work was supported by funding from the National Institutes of Health
474 (K99MH117229 to A.R.M.). UK Biobank analyses were conducted via application
475 31063. The BioBank Japan Project was supported by the Tailor-Made Medical
476 Treatment Program of the Ministry of Education, Culture, Sports, Science, and

477 Technology (MEXT) and the Japan Agency for Medical Research and Development
478 (AMED). M.K. was supported by a Nakajima Foundation Fellowship and the Masason
479 Foundation.

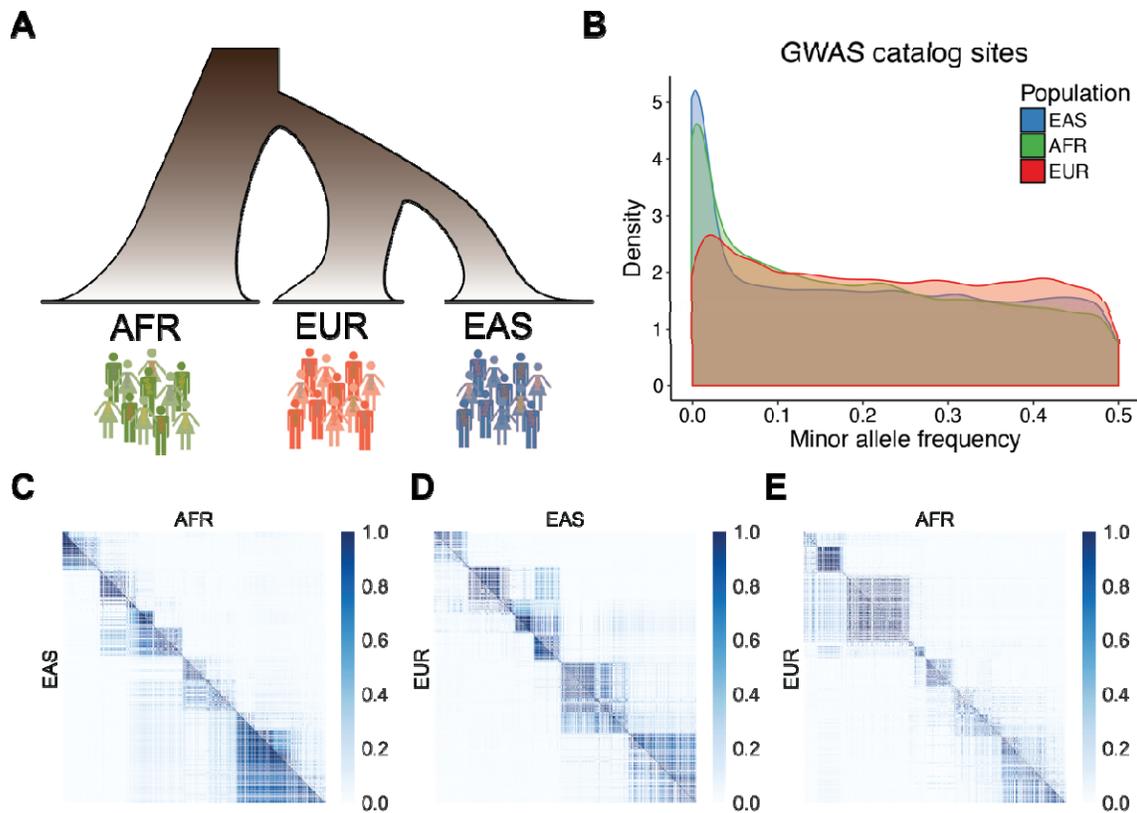
480

481 **Figures**



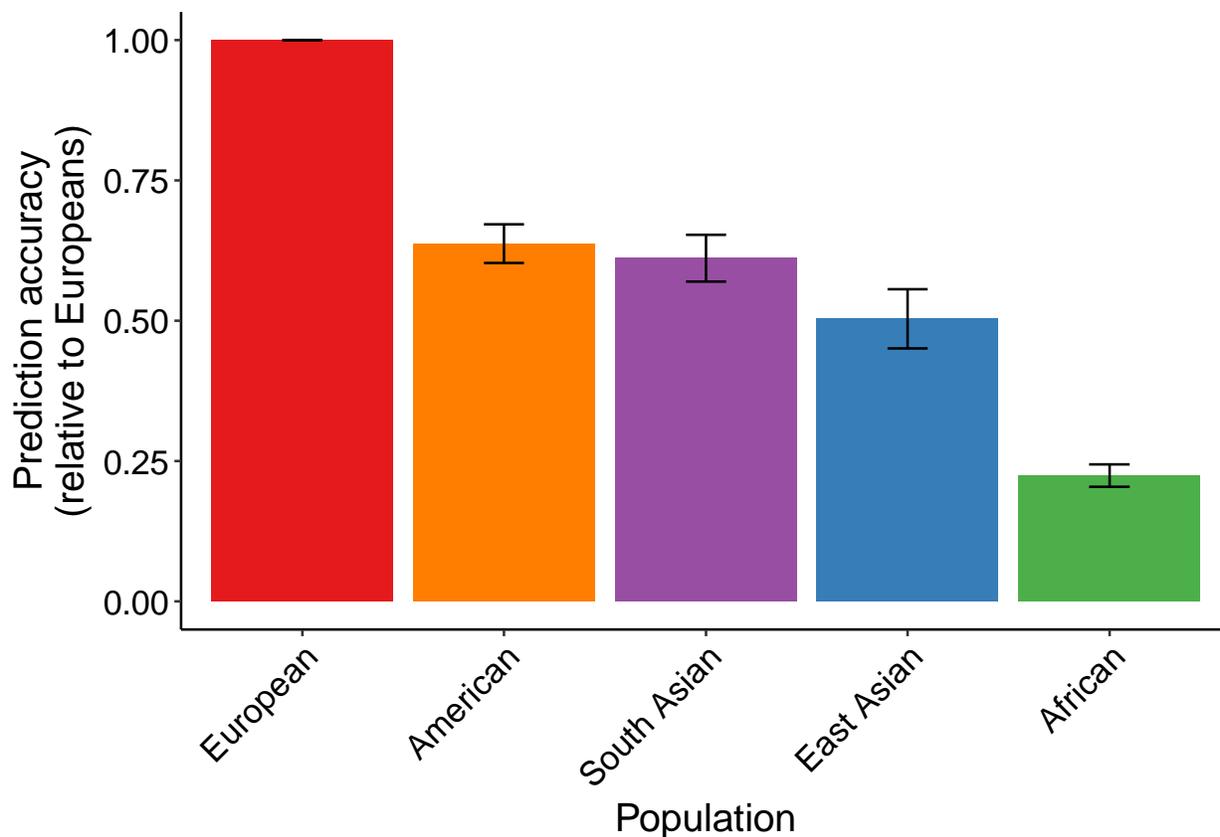
482

483 **Figure 1 – Ancestry of GWAS participants over time compared to the global**
484 **population.** Cumulative data as reported by the GWAS catalog²⁷. A notable caveat is
485 that because some cohorts are included in numerous studies, some individuals are
486 represented multiple times. This bias in multiple counting is especially likely for publicly
487 available cohorts, which are more likely to be of European or East Asian descent.
488 Individuals whose ancestry is “not reported” are not shown.

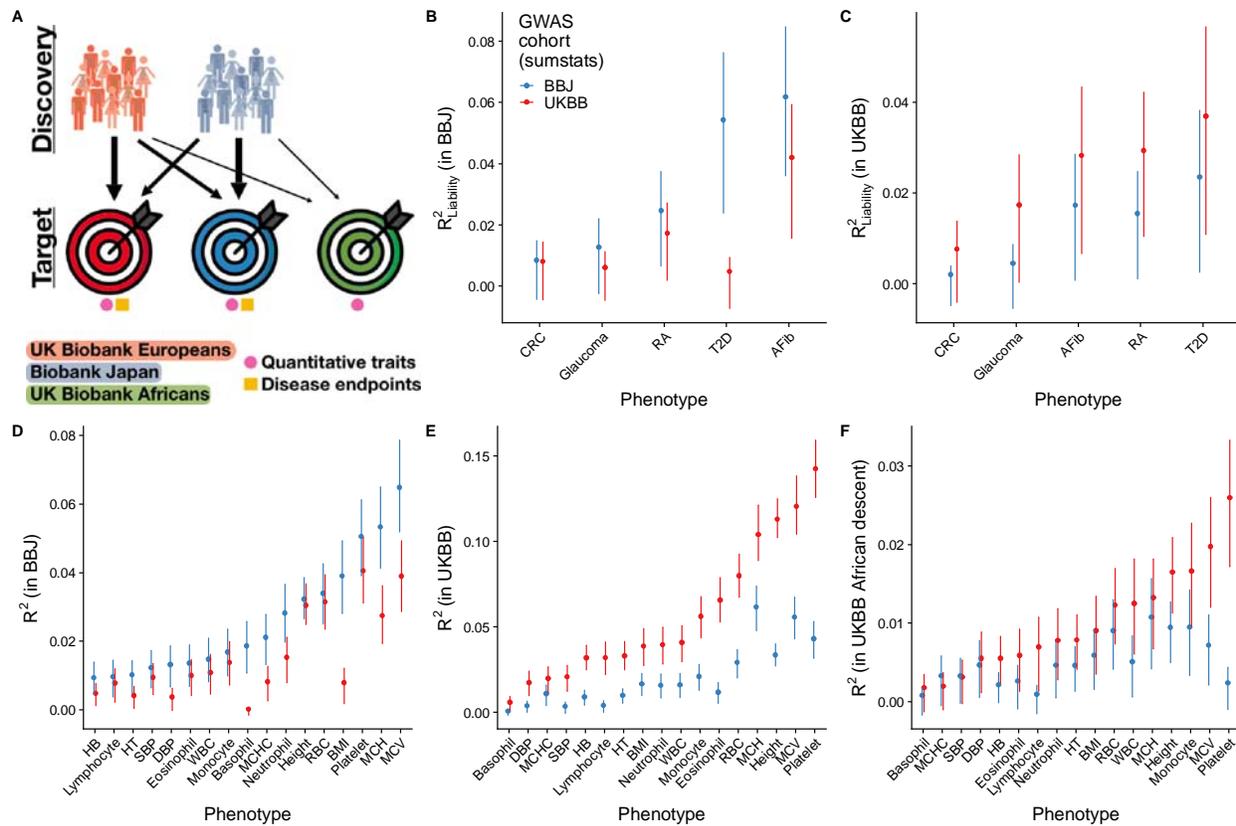


489
490 **Figure 2 – Demographic relationships, allele frequency differences, and local LD**
491 **patterns between population pairs.** Data analyzed from 1000 Genomes, in which
492 population labels are: AFR = continental African, EUR = European, and EAS = East
493 Asian. A) Cartoon relationships among AFR, EUR, and EAS populations. B) Allele
494 frequency distributions in AFR, EUR, and EAS populations of variants from the GWAS
495 catalog. C-E) Color axis shows LD scale (r^2). LD comparisons between pairs of
496 populations show the same region of the genome for each comparison (representative
497 region is chr1, 51572kb-52857kb) among pairs of SNPs polymorphic in both
498 populations, illustrating that different SNPs are polymorphic across some population
499 pairs, and that these SNPs have variable LD patterns across populations.

500



501
502 **Figure 3 – Prediction accuracy relative to European ancestry individuals across**
503 **17 quantitative traits and 5 continental populations in UKBB.** All phenotypes shown
504 here are quantitative anthropometric and blood panel traits, as described in **Table S1**,
505 including sample sizes. Bars show mean values, and error bars show standard errors of
506 the means.



508

509 **Figure 4 – Polygenic risk prediction accuracy in Japanese, British, and African**
 510 **descent individuals using independent GWAS of equal sample sizes in the**
 511 **BioBank Japan and UK Biobank.** All target prediction cohorts are withheld from the
 512 GWAS and thus independent. Sample sizes in each GWAS are identical by design
 513 between BBJ and UKBB (**Table S1 and S3**). To optimize signal to noise, each point
 514 shows the maximum R^2 (i.e. best predictor) across five p-value thresholds. R^2 values for
 515 all p-value thresholds are shown in **Figures S8-S12**. Prediction accuracy tends to be
 516 higher in the UK Biobank for quantitative traits than in BioBank Japan and vice versa for
 517 disease endpoints, likely because of concomitant phenotype precision and

518 consequently observed heritability for these classes of traits (**Table S5-S7**)^{**}.

519 Abbreviations are as follows: AFib = atrial fibrillation, BMI = body mass index, CRC =

520 colorectal cancer, DBP = diastolic blood pressure, Hb = hemoglobin, Ht = Hematocrit,

521 MCH = mean corpuscular hemoglobin, MCHC = mean corpuscular hemoglobin

522 concentration, MCV = mean corpuscular volume, RA = rheumatoid arthritis, RBC = red

523 blood cell count, SBP = systolic blood pressure, T2D = type 2 diabetes, WBC = white

524 blood cell count. A) Explanatory diagram showing the different discovery and target

525 cohorts/populations, and quantitative traits versus disease endpoints. B) Genetic

526 prediction accuracy for five diseases in Japanese individuals using summary statistics

527 from GWAS of independent BioBank Japan versus UK Biobank samples. C) Genetic

528 prediction accuracy for the same five diseases in independent British individuals using

529 summary statistics from GWAS of independent BioBank Japan versus the UK Biobank

530 samples. D) Genetic prediction accuracy for 17 anthropometric and blood panel traits in

531 Japanese individuals using summary statistics from GWAS of independent BioBank

532 Japan versus UK Biobank samples. E) Genetic prediction accuracy for the same 17

533 anthropometric and blood panel traits in independent British individuals using summary

534 statistics from GWAS of independent BioBank Japan versus the UK Biobank samples.

535 F) Genetic prediction accuracy for 17 anthropometric and blood panel traits in African

536 descent individuals in the UK Biobank using summary statistics from GWAS of

537 independent BioBank Japan versus UK Biobank samples.

^{**} Thalassemia and sickle cell disease are unlikely to explain a significant fraction of these prediction differences, as few individuals have been diagnosed with these disorders via ICD-10 codes (**Table S11**).

538 References

- 539 1. Knowles, J. W. & Ashley, E. A. Cardiovascular disease: The rise of the genetic
540 risk score. *PLoS Med* **15**, e1002546–7 (2018).
- 541 2. Maas, P. *et al.* Breast Cancer Risk From Modifiable and Nonmodifiable Risk
542 Factors Among White Women in the United States. *JAMA Oncol* **2**, 1295–8
543 (2016).
- 544 3. Mavaddat, N. *et al.* Polygenic Risk Scores for Prediction of Breast Cancer and
545 Breast Cancer Subtypes. *The American Journal of Human Genetics* 1–42 (2018).
546 doi:10.1016/j.ajhg.2018.11.002
- 547 4. Schumacher, F. R. *et al.* Association analyses of more than 140,000 men identify
548 63 new prostate cancer susceptibility loci. *Nat Genet* 1–13 (2018).
549 doi:10.1038/s41588-018-0142-8
- 550 5. Khera, A. V. *et al.* Genome-wide polygenic score to identify a monogenic risk-
551 equivalent for coronary disease. (2017). doi:10.1101/218388
- 552 6. Kullo, I. J. *et al.* Incorporating a Genetic Risk Score Into Coronary Heart Disease
553 Risk Estimates: Effect on Low-Density Lipoprotein Cholesterol Levels (the MI-
554 GENES Clinical Trial). *Circulation* **133**, 1181–1188 (2016).
- 555 7. Paquette, M. *et al.* Polygenic risk score predicts prevalence of cardiovascular
556 disease in patients with familial hypercholesterolemia. *Journal of Clinical*
557 *Lipidology* **11**, 725–732.e5 (2017).
- 558 8. Natarajan, P. *et al.* Polygenic Risk Score Identifies Subgroup With Higher Burden
559 of Atherosclerosis and Greater Relative Benefit From Statin Therapy in the
560 Primary Prevention Setting Clinical Perspective. *Circulation* **135**, 2091–2101
561 (2017).
- 562 9. Tikkanen, E., Havulinna, A. S., Palotie, A., Salomaa, V. & Ripatti, S. Genetic risk
563 prediction and a 2-stage risk screening strategy for coronary heart disease.
564 *Arterioscler. Thromb. Vasc. Biol.* **33**, 2261–2266 (2013).
- 565 10. Frieser, M. J., Wilson, S. & Vrieze, S. Behavioral Impact of Return of Genetic Test
566 Results for Complex Disease: Systematic Review and Meta-Analysis. *Health*
567 *Psychol* **37**, 1134–1144 (2018).
- 568 11. Khera, A. V. *et al.* Genetic Risk, Adherence to a Healthy Lifestyle, and Coronary
569 Disease. *N Engl J Med* **375**, 2349–2358 (2016).
- 570 12. Khera, A. V. & Kathiresan, S. Genetics of coronary artery disease: discovery,
571 biology and clinical translation. *Nat Rev Genet* **18**, 331–344 (2017).
- 572 13. Martin, A. R. *et al.* Human Demographic History Impacts Genetic Risk Prediction
573 across Diverse Populations. *Am. J. Hum. Genet.* **100**, 635–649 (2017).
- 574 14. Scutari, M., Mackay, I. & Balding, D. Using Genetic Distance to Infer the Accuracy
575 of Genomic Prediction. *PLoS Genet* **12**, e1006288 (2016).
- 576 15. Vilhjálmsson, B. J. *et al.* Modeling Linkage Disequilibrium Increases Accuracy of
577 Polygenic Risk Scores. *The American Journal of Human Genetics* **97**, 576–592
578 (2015).
- 579 16. Ware, E. B. *et al.* Heterogeneity in polygenic scores for common human traits. 1–
580 13 (2017). doi:10.1101/106062

- 581 17. Curtis, D. Polygenic risk score for schizophrenia is more strongly associated with
582 ancestry than with schizophrenia. *bioRxiv* (2018).
583 doi:<https://doi.org/10.1101/287136>
- 584 18. Ripke, S. *et al.* Biological insights from 108 schizophrenia-associated genetic loci.
585 *Nature* **511**, 421–427 (2014).
- 586 19. Belsky, D. W. *et al.* Development and Evaluation of a Genetic Risk Score for
587 Obesity. *Biodemography and Social Biology* **59**, 85–100 (2013).
- 588 20. Domingue, B. W., Belsky, D. W., Conley, D., Harris, K. M. & Boardman, J. D.
589 Polygenic Influence on Educational Attainment. *AERA Open* **1**,
590 233285841559997–13 (2015).
- 591 21. Wedow, R. *et al.* Gene discovery and polygenic prediction from a genome-wide
592 association study of educational attainment in 1.1 million individuals. *Nat Genet*
593 **92**, 109 (2018).
- 594 22. Vassos, E. *et al.* An Examination of Polygenic Score Risk Prediction in Individuals
595 With First-Episode Psychosis. *Biological Psychiatry* **81**, 470–477 (2017).
- 596 23. Akiyama, M. *et al.* Genome-wide association study identifies 112 new loci for
597 body mass index in the Japanese population. *Nat Genet* **49**, 1458–1467 (2017).
- 598 24. Li, Z. *et al.* Genome-wide association analysis identifies 30 new susceptibility loci
599 for schizophrenia. *Nature Publishing Group* **49**, 1576–1583 (2017).
- 600 25. Need, A. C. & Goldstein, D. B. Next generation disparities in human genomics:
601 concerns and remedies. *Trends in Genetics* **25**, 489–494 (2009).
- 602 26. Popejoy, A. B. & Fullerton, S. M. Genomics is failing on diversity. *Nature* **538**,
603 161–164 (2016).
- 604 27. Morales, J. *et al.* A standardized framework for representation of ancestry data in
605 genomics studies, with application to the NHGRI-EBI GWAS Catalog. 1–10
606 (2018). doi:10.1186/s13059-018-1396-2
- 607 28. Rosenberg, N. A. *et al.* Genome-wide association studies in diverse populations.
608 *Nat Rev Genet* **11**, 356–366 (2010).
- 609 29. Sham, P. C., Cherny, S. S., Purcell, S. & Hewitt, J. K. Power of linkage versus
610 association analysis of quantitative traits, by use of variance-components models,
611 for sibship data. *The American Journal of Human Genetics* **66**, 1616–1630
612 (2000).
- 613 30. 1000 Genomes Project Consortium *et al.* A global reference for human genetic
614 variation. *Nature* **526**, 68–74 (2015).
- 615 31. Consortium, T. S. 2. D. Sequence variants in SLC16A11 are a common risk factor
616 for type 2 diabetes in Mexico. *Nature* **506**, 97–101 (2014).
- 617 32. Estrada, K. *et al.* Association of a Low-Frequency Variant in HNF1A With Type 2
618 Diabetes in a Latino Population. *JAMA* **311**, 2305–2314 (2014).
- 619 33. Haiman, C. A. *et al.* Genome-wide association study of prostate cancer in men of
620 African ancestry identifies a susceptibility locus at 17q21. *Nat Genet* **43**, 570–573
621 (2011).
- 622 34. Genovese, G. *et al.* Association of trypanolytic ApoL1 variants with kidney
623 disease in African Americans. *Science* **329**, 841–845 (2010).
- 624 35. Liu, J. Z. *et al.* Association analyses identify 38 susceptibility loci for inflammatory
625 bowel disease and highlight shared genetic risk across populations. *Nat Genet*
626 **47**, 979–986 (2015).

- 627 36. Carlson, C. S. *et al.* Generalization and Dilution of Association Results from
628 European GWAS in Populations of Non-European Ancestry: The PAGE Study.
629 *PLoS Biol* **11**, e1001661 (2013).
- 630 37. Easton, D. F. *et al.* Genome-wide association study identifies novel breast cancer
631 susceptibility loci. *Nature* **447**, 1087–1093 (2007).
- 632 38. Mahajan, A. *et al.* Genome-wide trans-ancestry meta-analysis provides insight
633 into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet* **46**, 234–
634 244 (2014).
- 635 39. Waters, K. M. *et al.* Consistent Association of Type 2 Diabetes Risk Variants
636 Found in Europeans in Diverse Racial and Ethnic Groups. *PLoS Genet* **6**,
637 e1001078 (2010).
- 638 40. Lam, M. *et al.* Comparative genetic architectures of schizophrenia in East Asian
639 and European populations. 1–41 (2018). doi:10.1101/445874
- 640 41. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype
641 imputation. *Nat Genet* **48**, 1279–1283 (2016).
- 642 42. Huang, H. *et al.* Fine-mapping inflammatory bowel disease loci to single-variant
643 resolution. *Nature* **547**, 173–178 (2017).
- 644 43. Berg, J. J. *et al.* Reduced signal for polygenic adaptation of height in UK Biobank.
645 1–44 (2018). doi:10.1101/354951
- 646 44. Sohail, M. *et al.* Signals of polygenic adaptation on height have been
647 overestimated due to uncorrected population structure in genome-wide
648 association studies. 1–44 (2018). doi:10.1101/355057
- 649 45. Kerminen, S. *et al.* Geographic variation and bias in polygenic scores of complex
650 diseases and traits in Finland: (2018). doi:10.1101/485441
- 651 46. Ge, T., Chen, C.-Y., Neale, B. M., Sabuncu, M. R. & Smoller, J. W. Phenome-
652 wide heritability analysis of the UK Biobank. *PLoS Genet* **13**, e1006711–21
653 (2017).
- 654 47. Inouye, M. *et al.* Genomic Risk Prediction of Coronary Artery Disease in 480,000
655 Adults: Implications for Primary Prevention. *J. Am. Coll. Cardiol.* **72**, 1883–1893
656 (2018).
- 657 48. Pourcain, B. S. *et al.* ASD and schizophrenia show distinct developmental profiles
658 in common genetic overlap with population-based social communication
659 difficulties. *Molecular Psychiatry* 1–8 (2016). doi:10.1038/mp.2016.198
- 660 49. Novembre, J. & Barton, N. H. Tread Lightly Interpreting Polygenic Tests of
661 Selection. *Genetics* **208**, 1351–1355 (2018).
- 662 50. Henn, B. M., Botigué, L. R., Bustamante, C. D., Clark, A. G. & Gravel, S.
663 Estimating the mutation load in human genomes. *Nat Rev Genet* **16**, 333–343
664 (2015).
- 665 51. Li, D., Zhao, H. & Gelernter, J. Strong protective effect of the aldehyde
666 dehydrogenase gene (ALDH2) 504lys (*2) allele against alcoholism and alcohol-
667 induced medical diseases in Asians. *Human Genetics* **131**, 725–737 (2011).
- 668 52. Zhu, Z. *et al.* Dominance Genetic Variation Contributes Little to the Missing
669 Heritability for Human Complex Traits. *The American Journal of Human Genetics*
670 **96**, 377–385 (2015).
- 671 53. Paré, G., Mao, S. & Deng, W. Q. A machine-learning heuristic to improve gene
672 score prediction of polygenic traits. *Sci Rep* **7**, 12665 (2017).

- 673 54. Martin, A. R. *et al.* An Unexpectedly Complex Architecture for Skin Pigmentation
674 in Africans. (2017). doi:10.1101/200139
- 675 55. Hindorff, L. A. *et al.* Prioritizing diversity in human genomics research. *Nature*
676 *Publishing Group* 1–11 (2017). doi:10.1038/nrg.2017.89
- 677 56. H3Africa Consortium *et al.* Research capacity. Enabling the genomic revolution in
678 Africa. *Science* **344**, 1346–1348 (2014).
- 679 57. Duncan, L. E. *et al.* Largest GWAS of PTSD (N=20,070) yields genetic overlap
680 with schizophrenia and sex differences in heritability. *Molecular Psychiatry* (2017).
681 doi:10.1038/mp.2017.77
- 682 58. Howrigan, D. Details and Considerations of the Uk Biobank GWAS.
683 [http://www.nealelab.is/blog/2017/9/11/details-and-considerations-of-the-uk-](http://www.nealelab.is/blog/2017/9/11/details-and-considerations-of-the-uk-biobank-gwas)
684 [biobank-gwas](http://www.nealelab.is/blog/2017/9/11/details-and-considerations-of-the-uk-biobank-gwas) (2017). Available at: (Accessed: 9 November 2017)
- 685 59. Kanai, M. *et al.* Genetic analysis of quantitative traits in the Japanese population
686 links cell types to complex human diseases. *Nat Genet* 1–16 (2018).
687 doi:10.1038/s41588-018-0047-6
- 688 60. Liu, S. *et al.* Genomic Analyses from Non-invasive Prenatal Testing Reveal
689 Genetic Associations, Patterns of Viral Infections, and Chinese Population
690 History. *Cell* **175**, 347–359.e14 (2018).
- 691 61. Brown, B. C., Ye, C. J., Price, A. L. & Zaitlen, N. Transethnic Genetic-Correlation
692 Estimates from Summary Statistics. *The American Journal of Human Genetics*
693 **99**, 76–88 (2016).
- 694 62. Wray, N. R. *et al.* Pitfalls of predicting complex traits from SNPs. *Nature*
695 *Publishing Group* **14**, 507–515 (2013).
- 696 63. Wray, N. R. *et al.* Research Review: Polygenic methods and their application to
697 psychiatric traits. *J Child Psychol Psychiatr* **55**, 1068–1087 (2014).
- 698 64. Torkamani, A., Wineinger, N. E. & Topol, E. J. The personal and clinical utility of
699 polygenic risk scores. *Nat Rev Genet* 1–10 (2018). doi:10.1038/s41576-018-
700 0018-x
- 701 65. Manrai, A. K., Patel, C. J. & Ioannidis, J. P. A. In the Era of Precision Medicine
702 and Big Data, Who Is Normal? *JAMA* **319**, 1981–1982 (2018).
- 703 66. Plenge, R. M., Scolnick, E. M. & Altshuler, D. Validating therapeutic targets
704 through human genetics. *Nature Publishing Group* **12**, 581–594 (2013).
- 705 67. Carroll, M. D., Kit, B. K., Lacher, D. A., Shero, S. T. & Mussolino, M. E. Trends in
706 lipids and lipoproteins in US adults, 1988-2010. *JAMA* **308**, 1545–1554 (2012).
- 707 68. Rappoport, N. *et al.* Creating ethnicity-specific reference intervals for lab tests
708 from EHR data. *bioRxiv* (2017). doi:10.1101/213892
- 709 69. Lim, E., Miyamura, J. & Chen, J. J. Racial/Ethnic-Specific Reference Intervals for
710 Common Laboratory Tests: A Comparison among Asians, Blacks, Hispanics, and
711 White. *Hawaii J Med Public Health* **74**, 302–310 (2015).
- 712 70. Hero, J. O., Zaslavsky, A. M. & Blendon, R. J. The United States Leads Other
713 Nations In Differences By Income In Perceptions Of Health And Health Care.
714 *Health Aff* **36**, 1032–1040 (2017).
- 715 71. Gilly, A. *et al.* Very low depth whole genome sequencing in complex trait
716 association studies. 1–24 (2018). doi:10.1101/169789
- 717 72. Pasaniuc, B. *et al.* Extremely low-coverage sequencing and imputation increases
718 power for genome-wide association studies. *Nat Genet* **44**, 631–635 (2012).

- 719 73. Martin, A. R., Teferra, S., Möller, M., Hoal, E. G. & Daly, M. J. The critical needs
720 and challenges for genetic architecture studies in Africa. *Current Opinion in*
721 *Genetics & Development* **53**, 113–120 (2018).
- 722 74. Coles, E. & Mensah, G. A. Geography of Genetics and Genomics Research
723 Funding in Africa. *Global Heart* **12**, 173–176 (2017).
- 724 75. Mulder, N. J. *et al.* Development of Bioinformatics Infrastructure for Genomics
725 Research. *Global Heart* **12**, 91–98 (2017).
- 726