The American Journal of Human Genetics, Volume *110*

Supplemental information

Genotype error due to low-coverage sequencing induces uncertainty in polygenic scoring

Ella Petter, Yi Ding, Kangcheng Hou, Arjun Bhattacharya, Alexander Gusev, Noah Zaitlen, and Bogdan Pasaniuc

Figure S1. Calibration of individual PGS credible intervals with respect to the matching array-PGS values

PGS CI are well calibrated with respect to their matching array-PGS, for different genetic architectures of the simulated effect sizes. The error bar represents one standard error of the mean empirical coverage calculated across 10 simulations.

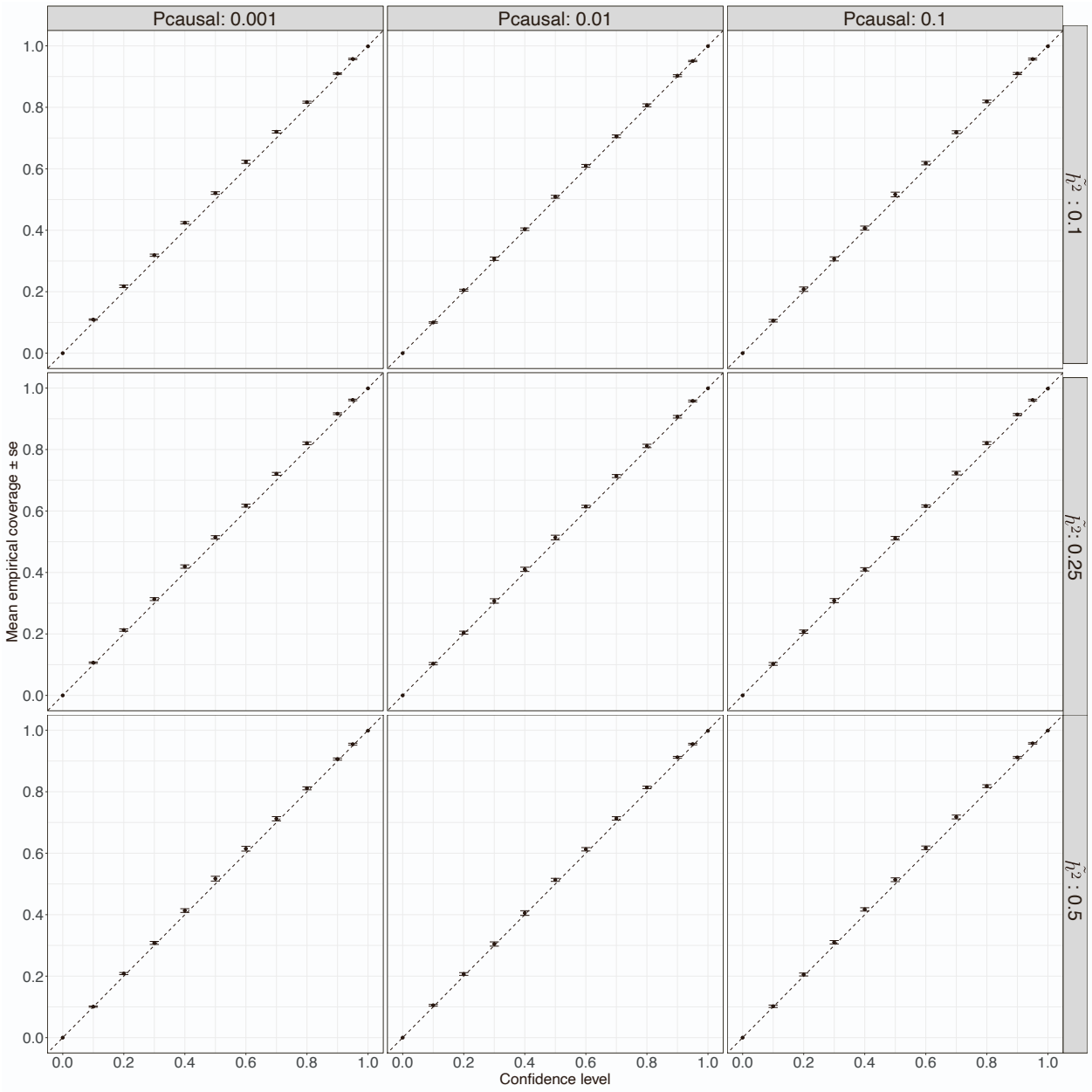

Figure S2. Calibration of individual PGS credible intervals with respect to the matching array-PGS values, under a minor allele frequency aware effect size simulation

PGS CI are well calibrated with respect to their matching array-PGS, for different genetic architectures of the simulated effect sizes. In this set of simulations, effect sizes are drawn dependent on minor allele frequency. The error bar represents one standard error of the mean empirical coverage calculated across 10 simulations.

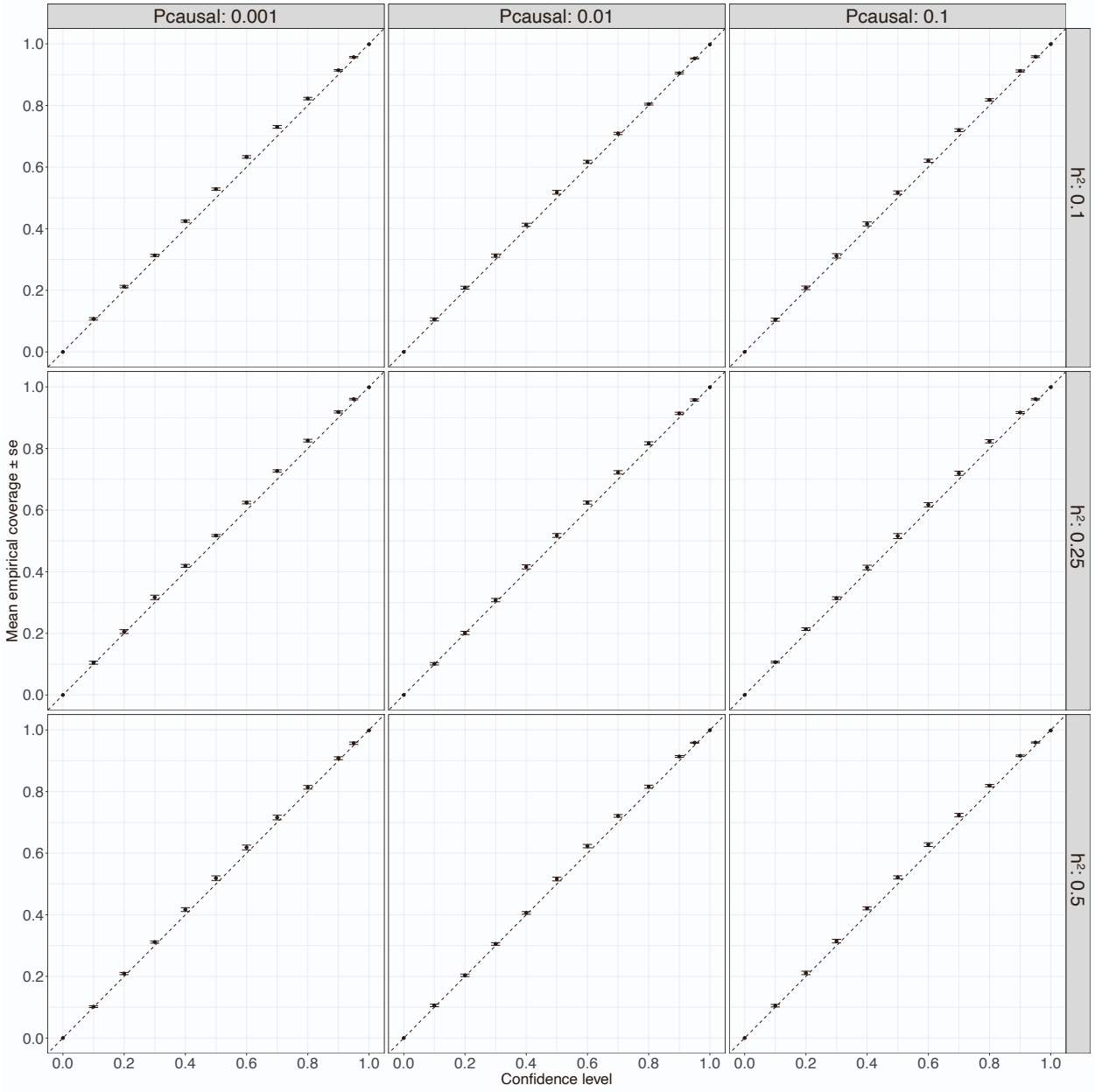

Figure S3. High LD between SNPs may induce over calibration of the PGS CI

We tested the calibration of the method for a simulated PGS with 1000 variants that are strongly correlated on chromosome 16, and another set of 1000 SNPs on chromosome 16 that have a low correlation. We found that for the high-LD simulation, the PGS CI is slightly over-calibrated.

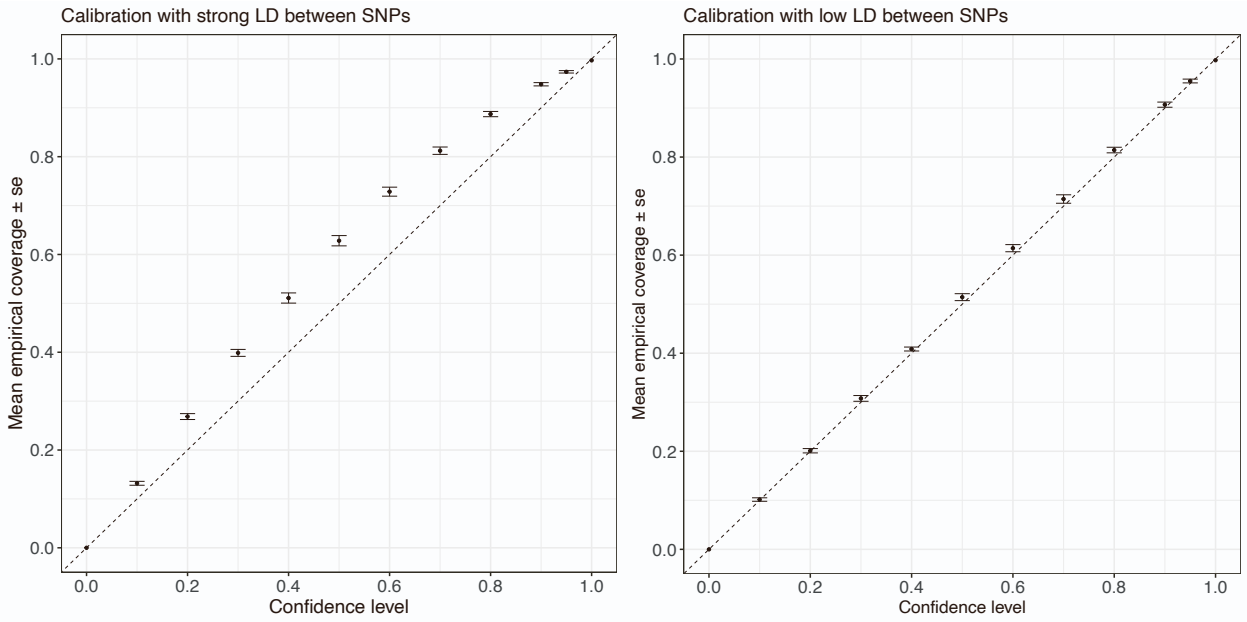

Figure S4. The impact of sequencing depth on PGS distribution variability and PGS error, for different levels of SNP imputation INFO scores

Using the effect sizes form the height PGS, we run three analyses: including all SNPs, including SNPs with INFO score > 0.4, and including SNPs with INFO score > 0.7. We scale the results to a unit equal to one standard deviation of the array PGS in each of the three groups. For high INFO scores, the individual PGS SD decreases in size, but the difference between sequencing depth groups remains significant (top). The error between the dosage-PGS and array-PGS also remains significantly different between the two sequencing depth groups across all INFO score levels, with no other noticeable difference between the three INFO score levels (bottom).

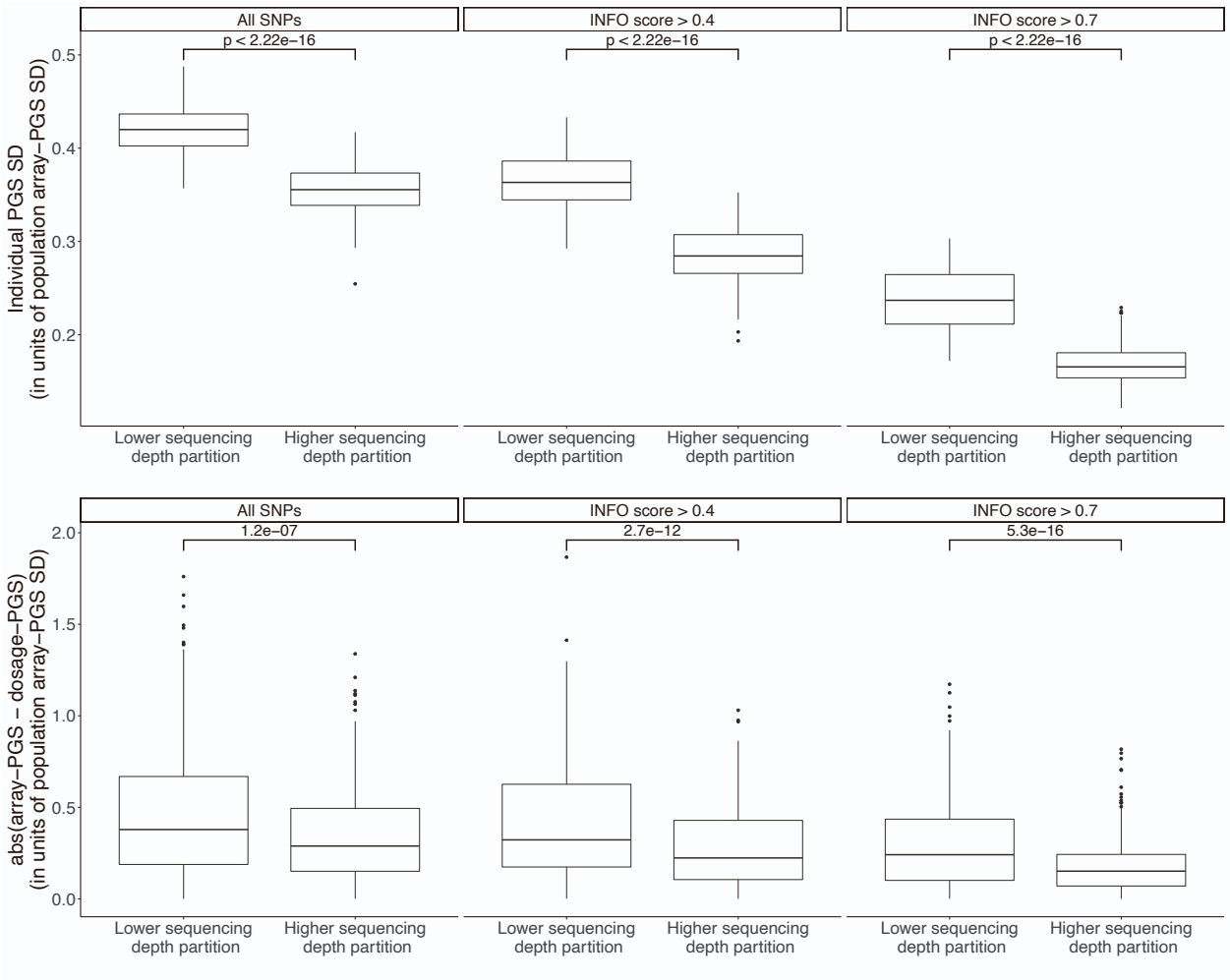

Figure S5. Stratification of individuals into groups based on classification certainty

Individuals are classified as high-risk (orange) or low-risk (blue) based on their dosage-PGS estimate, compared to the 90% threshold of the population's dosage-PGS. On average, only 20.4% of the high-risk individuals are certainly at high-risk when accounting for the overlap of their individual PGS CI with the threshold of interest. For the opposite direction, only an average of 71.35% of individuals classified as low-risk do not have their PGS CI overlap with the risk threshold.

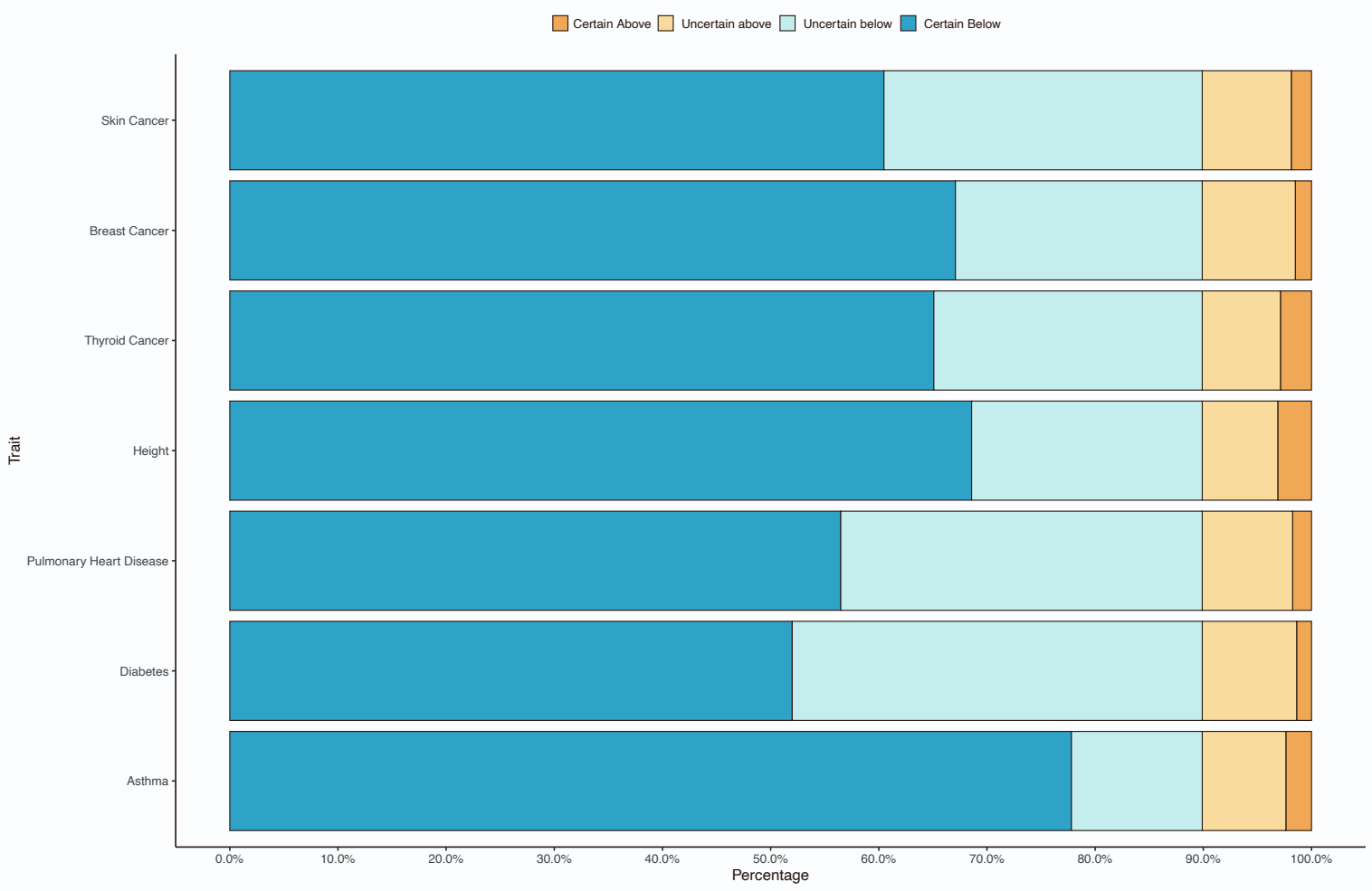

Figure S6. The effect of phenotype genetic architecture on risk-stratification

Using simulated effect sizes with varying levels of variance parameters and proportion of causal SNPs, we estimate how does the risk-stratification change by genetic architecture. We find that with higher variance parameters ($\widetilde{h^2}$ = 0.5) there is a significantly higher percentage of high-risk individuals reclassified as "uncertain above threshold". We did not find any difference between $\widetilde{h^2}$=0.25 and $\widetilde{h^2}$=0.1, nor did we find any difference between different causal SNP proportions.

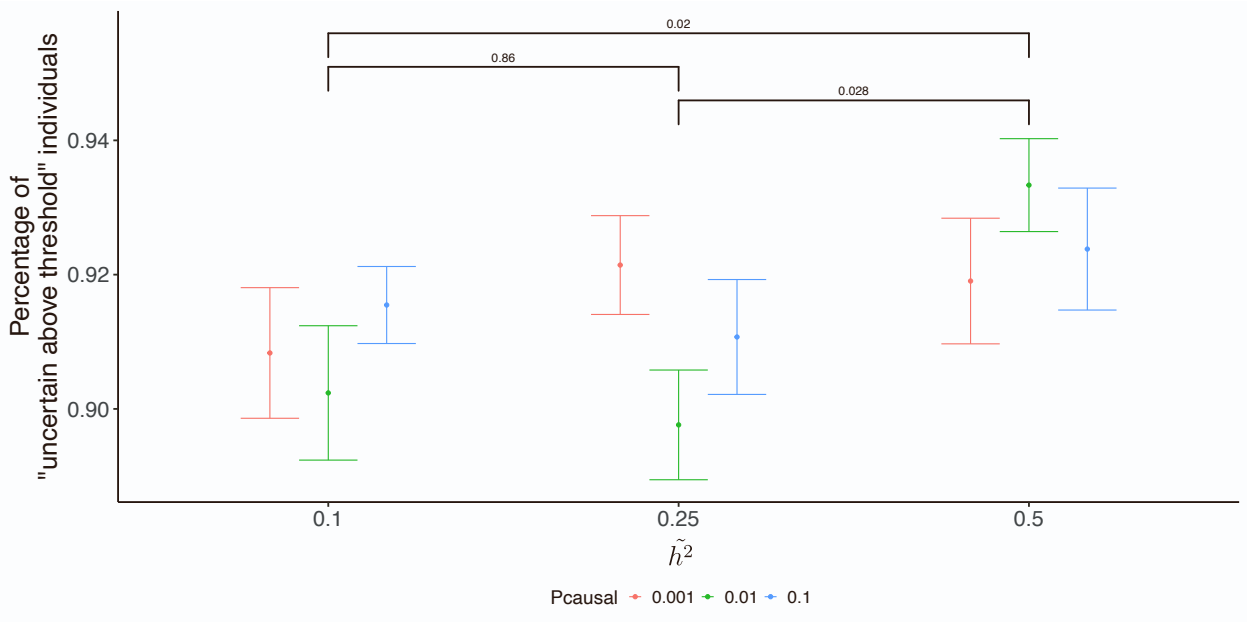

Figure S7. Higher coverage individuals tend to be more represented when selecting by $Pr(PGS_i > t)$

For the thyroid cancer PGS, we calculate the average proportion of each sequencing depth group out of all individuals selected at the top of the dosage-PGS and $Pr(PGS_i > t)$ distributions, across 20 iterations. We found that the 10x simulated coverage group is more represented than other groups when selecting by $Pr(PGS_i > t)$ (top right panel). When looking at the all individuals in the top two deciles of the array-PGS distribution, we find that the average $Pr(PGS_i > t)$ value is higher among the sub-group with higher sequencing depths, demonstrating the prioritization for higher certainty individuals.

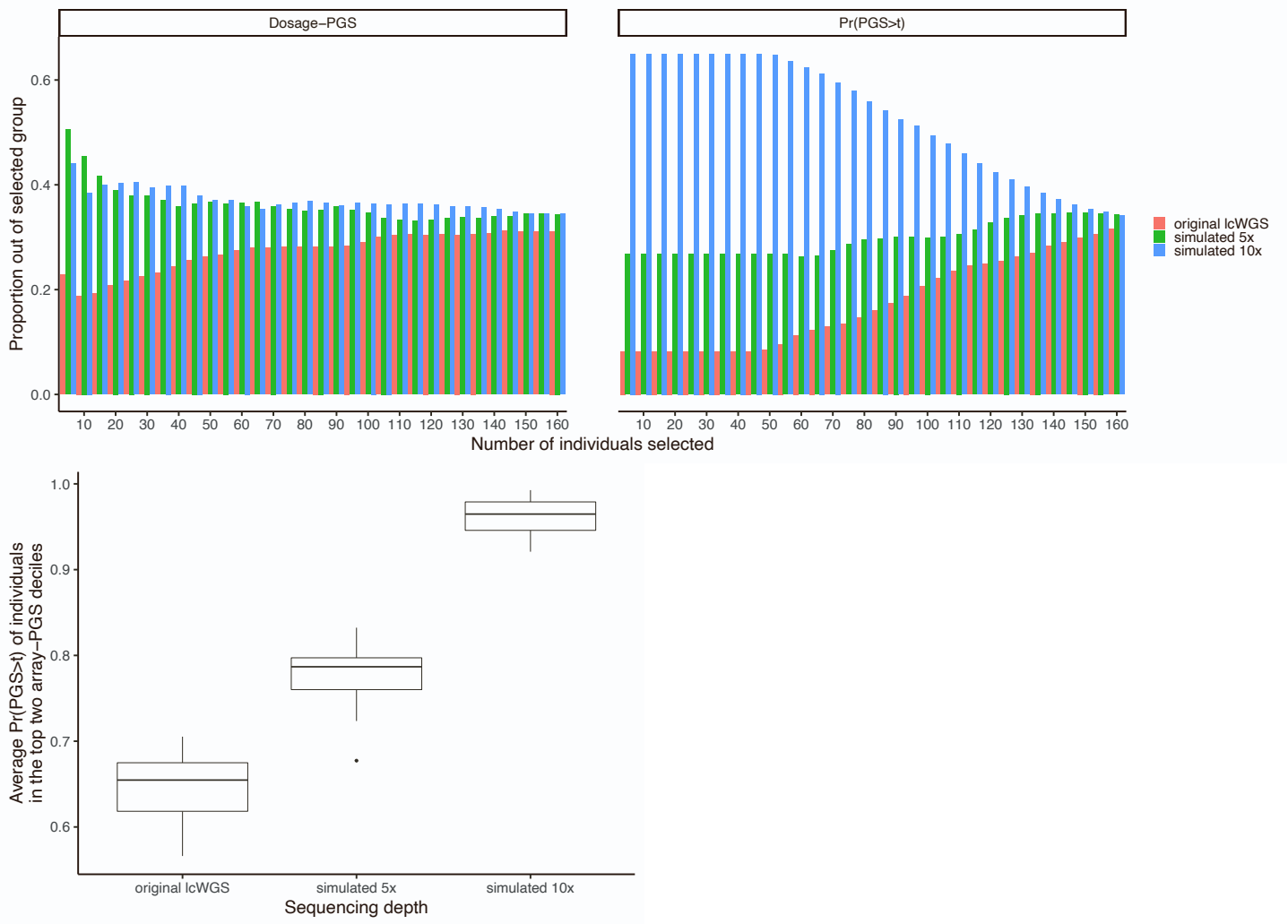

Figure S8. Risk-stratification using $Pr(PGS_i > t)$ does not yield improved precision in a homogenous sequencing depth cohort

True positive percentages of risk-stratification based on dosage-PGS (red) and $Pr(PGS_i > t)$ (blue) in real lcWGS data across 7 traits.

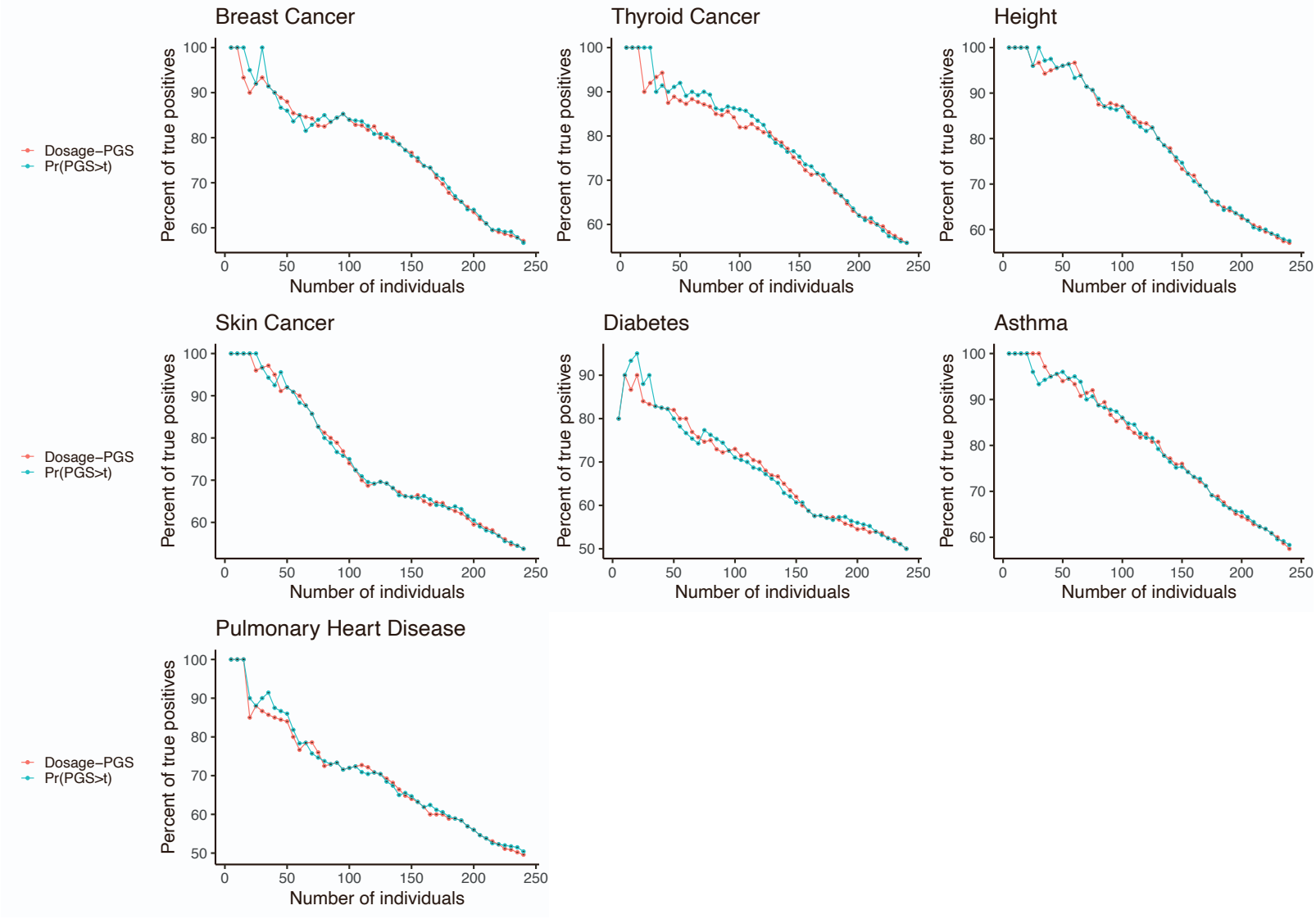

Figure S9. Changing the magnitude of effect-size error affects the balance between the contribution of each error type to the overall calibration

Empirical coverage was calculated for three types of credible intervals: accounting for genotype error alone, accounting for effect size error alone, accounting for both sources of error. Each empirical coverage calculation was repeated for different magnitudes of effect size error, controlled by the simulated GWAS sample size. We find that for higher GWAS sample sizes (darker colors), the genotype error only CIs (blue) tend to be more calibrated while the effect size only CIs (red) tend to be less calibrated.

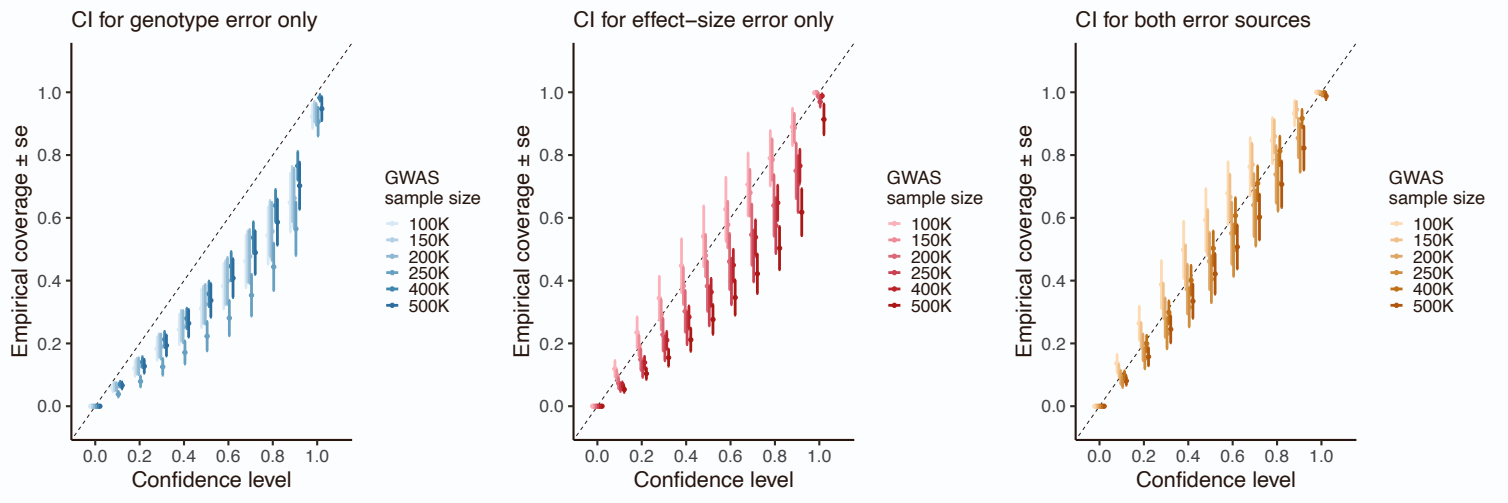

Table S1: The stratification of patients into risk groups changes when including genotype uncertainty

Using simulated effect sizes and real lcWGS genotypes we create a confusion matrix showing the average and standard errors of the number of individuals in our cohort that were classified to each risk groups when accounting for effect size uncertainty only (rows) and for both effect size and genotype uncertainty (columns). The experiment was ran using 10 different sets of random effect size vectors.

	Risk category based on PGS CI for effect size and genotype uncertainties				
Risk category based on PGS CI for effect size uncertainty only		Certain Below	Uncertain Below	Uncertain Above	Certain Above
	Certain Below	181.8 (18)	118.8 (5.3)	0	0
	Uncertain Below	0	448.4 (23.26)	0	0
	Uncertain Above	0	0	82.7 (0.33)	0
	Certain Above	0	0	0.9 (0.314)	0.4 (0.16)