



Published in final edited form as:

Mol Psychiatry. 2019 June ; 24(6): 819–827. doi:10.1038/s41380-019-0394-4.

Genomic prediction of cognitive traits in childhood and adolescence

A.G. Allegrini^{1,*}, S. Selzam¹, K. Rimfeld¹, S. von Stumm², J.B. Pingault³, R. Plomin¹

¹King's College London, Social, Genetic and Developmental Psychiatry Centre, London, United Kingdom

²London School of Economics and Political Science, Department of Psychological and Behavioural Science, London, United Kingdom

³University College London, Clinical Educational and Health Psychology, Division of Psychology and Language Sciences, London, United Kingdom

Abstract

Recent advances in genomics are producing powerful DNA predictors of complex traits, especially cognitive abilities. Here, we leveraged summary statistics from the most recent genome-wide association studies of intelligence and educational attainment, with highly genetically correlated traits, to build prediction models of general cognitive ability and educational achievement. To this end, we compared the performances of multi-trait genomic and polygenic scoring methods. In a representative UK sample of 7,026 children at ages 12 and 16, we show that we can now predict up to 11 percent of the variance in intelligence and 16 percent in educational achievement. We also show that predictive power increases from age 12 to age 16 and that genomic predictions do not differ for girls and boys. We found that multi-trait genomic methods were effective in boosting predictive power. Prediction accuracy varied across polygenic score approaches, however results were similar for different multi-trait and polygenic score methods. We discuss general caveats of multi-trait methods and polygenic score prediction, and conclude that polygenic scores for educational attainment and intelligence are currently the most powerful predictors in the behavioural sciences.

Introduction

Ever increasing sample sizes and methodological advances in polygenic methods have made it possible to powerfully predict complex traits such as cognitive abilities without knowing anything about the causal chain between genes and behaviour. Progress in predicting cognitive traits from inherited DNA variants has been rapid in the past five years and

*corresponding author: Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, Denmark Hill, London, United Kingdom, SE5 8AF, andrea.allegrini@kcl.ac.uk.

Author contributions

AGA and RP conceived and designed the study. AGA analysed and interpreted the data. SS performed quality control of genotype data. AGA and RP wrote the manuscript. SS, KR, SvS and JBP contributed to and critically reviewed the manuscript.

Competing interests

The authors declare no conflict of interest.

Supplementary information is available at MP's website

especially in the past year¹. Three methodological advances have mainly been responsible for this progress: increasingly large genome-wide association (GWA) studies, genome-wide polygenic scores (GPS) and multivariate analytic tools. The key has been the recognition that the largest associations are extremely small, accounting for less than 0.05% of the variance². To achieve sufficient power to detect such small effect sizes, samples in the hundreds of thousands are needed before GWA studies can begin to detect these tiny effects. Because the largest associations are so small, useful predictions of individual differences can only be made by aggregating the effects of thousands of DNA variants in GPS³. The third advance is the development of genomic methods that leverage genetic correlations between traits to boost power for variant discovery⁴ and polygenic risk prediction⁵.

Together, these three advances have greatly increased the ability to predict intelligence, educational attainment (years of schooling), and educational achievement (tested performance). For example, for intelligence, until 2017, no replicable associations were found in seven GWA studies^{6–12}, which we refer to collectively as ‘IQ1’. These studies had sample sizes from 18,000 to 54,000, which seemed large at the time but were not sufficiently powered to detect effect sizes of 0.05%. GPS derived from these IQ1 GWA studies at most accounted for 1% of the variance in independent samples. Increasing GWA sample sizes to 78,000 (IQ2¹³) and then to 280,000 (IQ3¹⁴) paid off in increasing predictive power of GPS from 1% to 3% to 4%. Here we present results for IQ3.

Educational attainment has led the way in terms of increasing GWA sample size, from 125,000 in 2013 (EA1¹⁵) to 294,000 in 2016 (EA2¹⁶) to 1.1 million in 2018 (EA3¹⁷). The growing sample sizes increased the predictive power of GPS from 2% to 3% to 12% of the variance in educational attainment¹. Similarly, in previous work we showed that EA GPS predicted an increasingly substantial amount of variance in tested educational achievement as sample size from replications of the EA GWAS increased over the years. EA1 predicted 3% of the variance in educational achievement at age 16¹⁸ and EA2 predicted 9% of the variance for overall educational achievement at age 16¹⁹.

Because ‘years of education’ is obtained as a demographic marker in most GWA studies, it was possible to accumulate samples sizes with the necessary power to detect very small effect sizes. It is more difficult to obtain very large sample sizes for intelligence, which needs to be assessed with a psychometric test administered to each individual, whereas years of education can be captured with a single self-reported item. Because of the large sample size available for EA GWA studies and the substantial genetic correlation between EA and intelligence, EA GPS predicted as much or more variance in intelligence than did GPS derived from GWAS of the target trait of intelligence itself. EA1 predicted 1% of the variance in intelligence^{18, 20} and EA2 predicted 4% of the variance¹⁶. Here we present results for EA3.

Finding that EA GPS predict educational achievement and intelligence better than do GWA of the target traits themselves suggests the usefulness of multivariate approaches. In a previous study, a multivariate GPS approach involving regularized regression was applied to show that with EA2 and 80 other GPS 11% of the variance in educational achievement at age 16 and 5% of the variance in intelligence at age 12 could be predicted²¹. Although

adding 1–2% to the predictive power of GPS might not seem like much, it should be noted that five years ago the total variance that could be predicted in either trait was statistically indistinguishable from zero.

The aim of the present study is to estimate how much variance in intelligence and educational achievement can be predicted by applying several state-of-the-art multi-trait genomic approaches and leveraging highly powered GWA summary statistics. First we compare three polygenic score methods (PRSice²², LDpred²³, and Lassosum²⁴) and test how much variance the new IQ3 and EA3 GPS maximally predict. We then jointly analyse IQ3 and EA3 with three highly (genetically) correlated traits (Income²⁵, Age when completed full time education²⁶, Time spent using computer²⁶) to boost predictive power and compare performance of three multi-trait methods (Genomic SEM²⁷, MTAG⁴ and SMTpred⁵) using predictive power as our criterion.

We conducted these analyses in a sample of 7,026 unrelated individuals from the Twins Early Development study, which is representative of the UK population²⁸. We analysed intelligence and educational achievement at the end of compulsory schooling in the UK at age 16; we also investigated developmental trends in genomic prediction from age 12 to 16. Based on previous research¹⁹, we expected genomic predictions to increase from 12 to 16.

Materials and Methods

Sample

The sample was drawn from the Twins Early Development Study (TEDS²⁹), an ongoing population-based longitudinal study. It consists of twins born in England and Wales between 1994 and 1996, who have been assessed on a variety of psychological domains. More than 10,000 twin pairs representative of the general UK population²⁸ remain actively involved in the study to date. Ethical approval for TEDS has been provided by the King's College London Ethics Committee (reference: PNM/09/10–104). Parental consent was obtained before data collection. Genotypes for 10,346 individuals (including 3,320 DZ twin pairs) were processed with stringent quality control procedures followed by SNP imputation using the Haplotype Reference Consortium (release 1.1) reference panels. Current analyses were limited to the genotyped and imputed sample of 7,026 unrelated individuals. Following imputation, we excluded variants with minor allele frequency < 0.5%, Hardy-Weinberg equilibrium p-values of < 1×10^{-5} . To ease computational demands, we selected variants with an info score of 1, resulting in 515,000 SNPs used for analysis (see Supplementary Methods S1 for a full description of quality control and imputation procedures).

Outcome variables

The outcome variables were intelligence and educational achievement at ages 12 and 16. Intelligence was assessed as a composite of verbal and nonverbal web-based tests. Educational achievement was indexed by a mean of scores on the compulsory subjects of English, mathematics and science obtained from the UK National Pupil Database. A more detailed description of outcome variables is provided in the Supplementary Methods S2. Supplementary Table S1 includes descriptive statistics for the outcomes variables and

Supplementary Figure S1 shows phenotypic correlations. Phenotypes and polygenic scores were corrected for age, sex and 10 genetic principal components. The obtained standardised residuals were used in all subsequent analyses.

Discovery GWA summary statistics

We based our prediction models on beta weights derived from large, publicly available, GWA summary statistics. Of central importance for our analyses were the most recent GWA studies of educational attainment (EA3¹⁷) and intelligence (IQ3¹⁴). Because the original IQ GWA meta-analysis included TEDS as one of its samples, to avoid bias due to sample overlap with our target sample we used summary statistics from new GWA analyses that excluded TEDS. The EA3 summary statistics employed here do not include 23andMe data (~300k individuals) due to their data availability policy.

Polygenic score approaches

We used IQ3 and EA3 summary statistics to construct genome-wide polygenic scores (GPS) comparing three distinct approaches: PRSice²², a clumping/pruning + P-value thresholding (P+T) approach, with an in-built high-resolution option that returns the best-fit GPS for the trait of interest; LDpred²³, a Bayesian approach that uses a prior on the expected polygenicity of a trait (assumed fraction of non-zero effect markers) and adjusts for linkage disequilibrium based on a reference panel to compute SNPs weights; and Lassosum²⁴, a machine-learning approach which uses penalized regression on GWA summary statistics to produce more accurate beta weights.

A detailed description of the construction of these polygenic scores is included in Supplementary Methods S3.

Multi-trait approaches

In order to boost power of IQ3 (N = 266,453) and EA3 (N = 766,345) GWA results and thus precision of beta weights to construct more predictive IQ3 and EA3 polygenic scores, we jointly analysed these summary statistics with three cognitive and educationally relevant traits: “Income”²⁵ (N = 96,900), “Age when completed full time education”²⁶ (N = 226,899) and “Time spent using computer”²⁹ (N = 261,987). The choice of these traits is consistent with a multi-trait framework, as these traits show the highest genetic correlations with IQ and educational attainment among publicly available GWA summary statistics, with pairwise-genetic correlations ranging from ~.5 to ~.9 (see Supplementary Figure S2). Summary statistics from these GWA studies are reported in Supplementary Table S2.

We used three recently developed multi-trait methods, one of which is specifically designed to boost polygenic score prediction: SMTpred⁵, and two of which are strictly speaking multivariate GWA approaches, designed to boost power for discovery, but which have been shown to increase predictive power of polygenic scores created from multi-trait reweighted summary statistics: MTAG⁴ and Genomic SEM²⁷. Details about these methods are provided in Supplementary Methods S4. Briefly, SMTpred⁵ is a multi-trait extension of the random effects model approach, which can be used to create multivariate best linear unbiased predictors based on summary statistics (wMT-SBLUP). MTAG is a generalization of

inverse-variance weighted meta-analysis, which jointly analyses univariate GWA summary statistics. It boosts power for discovery for each trait conditional on the effect size estimates of other traits and outputs trait-specific summary statistics. Genomic SEM is a two-stage structural equation modelling approach that can be applied in the context of multivariate GWA. In the form employed here (common factor GWAS), it directly tests effect of SNPs on a latent genetic factor defined by several indicators (i.e. traits) and outputs summary statistics for the common factor. We also compared these new multivariate approaches to a simple multiple regression on intelligence and on educational achievement using five GPS, each derived from the univariate GWA summary statistics used in multi-trait analyses.

Analyses

Univariate analyses

We first calculated polygenic scores for the IQ3 and EA3 GWA summary statistics using PRSice, LDpred and Lassosum. This was done to compare current state-of-the-art polygenic scores approaches and in order to obtain a benchmark against which to compare improvements in prediction accuracy due to multivariate GWA analyses. For each phenotype (i.e. intelligence and educational achievement at ages 12 and 16), we randomly split the sample into training and test sets (~50% training, ~50% test). Supplementary Table S1 shows descriptive statistics for each set. In the training sets, parameter optimization of GPS was performed, in which each GPS instrument (or p-value threshold in the case of PRSice, fraction of markers with nonzero effect in the case of LDpred, and tuning parameters in the case of Lassosum) was tested on each of the four phenotypes and the best instrument was selected with respect to prediction accuracy (as indexed by R^2). Performance of the optimized GPS instrument retained from the validation was then assessed in the test sample in order to evaluate how well the chosen predictors would perform in independent samples. We then proceeded to perform the multi-trait analyses.

Multi-trait analyses

We performed a multi-trait reweighting in SMTpred after transforming the ordinary least square betas from GWA studies of 'IQ', 'EA', 'Income', 'Age completed full time education' and 'Time spent using computer' in approximate Best Linear Unbiased Predictors (BLUP) using GCTA-Cojo³⁰. We then used LDSC to calculate SNP h^2 and genetic correlations between traits and proceeded to the multivariate weighting of traits as described in (Meier et al., 2018) to obtain multi-trait summary statistics BLUP (wMT-SBLUP; see also Supplementary Methods S4).

MTAG was run on the five GWA summary statistics (IQ, EA, Income, Age completed full time Education, Income) using standard settings. Because MTAG combines differently powered summary statistics (as indexed by the GWAS mean χ^2 ; see Supplementary Methods S4), as well as differing degrees of genetic overlap between traits, it can lead to an increased rate of false positives Type I error⁴. However, this is not an issue in the present study, which focuses on prediction accuracy rather than variant discovery. It has been shown⁴ that MTAG estimates consistently have a lower genome-wide mean-squared error compared to single-trait GWA estimates, and, therefore, polygenic scores created from

MTAG perform better than those created at the univariate level. However, in order to control for type I error inflation, we used the recommended⁴ false discovery rate (FDR) calculations (see Supplementary Methods S4).

The same five summary statistics were analysed using Genomic SEM. First a common factor model with the five summary statistics as indicators was fitted using a weighted least-square (WLS) estimator (default setting in Genomic SEM). Then a common factor GWA analysis with a WLS estimator was run, testing effects of single SNPs on the common factor. The WLS estimator was expected to yield lower standard errors and possibly increased prediction accuracy of GPS³⁰.

We then created polygenic scores from the MTAG EA, MTAG IQ and common factor GWA summary statistics across the three polygenic scores approaches, after splitting the sample into a training set to tune parameters and a testing set to assess model performance. In the case of SMTpred, polygenic scores for IQ3 and EA3 converted and reweighted indices (wMT-SBLUP) were calculated using PLINK³¹. These multi-trait predictors were then directly tested for model performance in the test set, as with the other GPS approaches. Based on previous power analyses for polygenic score prediction in the TEDS sample^{19, 32} we did not expect any power issues for the current analysis plan.

For prediction estimates derived from both univariate and multi-trait models, we calculated bootstrapped confidence intervals with 1000 replications. Furthermore, we performed a comparison of R² estimates between models, by calculating bootstrapped confidence intervals for the R² pairwise mean differences. As such, for each model, bootstrap samples were generated by sampling with replacement from the data 1000 times. Each row of data for resampling consisted of all polygenic scores and phenotypes examined herein. This procedure yielded an R² distribution for each method tested. The R² difference between methods was then calculated for each bootstrap iteration. This generated a distribution of R² differences, from which we calculated 95% confidence intervals.

Results

Polygenic score prediction of IQ and EA across GPS methods

Figure 1 shows variance in intelligence and educational achievement predicted by IQ3 GPS and EA3 GPS calculated following three polygenic score methods (PRSice, LDpred and Lassosum). Supplementary Table S3 presents associations in the training and test sets across all models.

For intelligence, IQ3 GPS predicted a maximum of 5.3% ($\beta = 0.221$, $se = 0.023$, $p < .0001$) of the variance at age 12 and 6.7% ($\beta = 0.266$, $se = 0.032$, $p < .0001$) at age 16. For educational achievement, EA3 GPS predicted a maximum of 6.6% ($\beta = 0.259$, $se = 0.020$, $p < .0001$) of the variance at age 12 and 14.8% ($\beta = 0.389$, $se = 0.019$, $p < .0001$) at age 16. EA3 GPS was also a powerful predictor of intelligence, predicting 7.2% ($\beta = 0.265$, $se = 0.024$, $p < .0001$) of the variance in intelligence at age 12 and 9.9% ($\beta = 0.321$, $se = 0.031$, $p < .0001$) at age 16.

Generally, Lassosum was the most powerful approach, predicting up to 1% more of the variance compared to LDpred and up to 2% more compared to PRSice. Supplementary Figure S4 shows a comparison of prediction estimates for each pair of approaches tested. Bootstrapped confidence levels calculated for pairwise comparisons indicated significant differences in prediction accuracy of IQ3 GPS within-trait between LDpred and PRSice at age 12 (MeanDiff = -0.014, 95% CIs [-0.024; -0.005]), and cross-trait at age 12 and 16. However, no significant differences were found for PRSice- vs LDpred-based EA3 GPS. Similarly, significant differences were also found for IQ3 GPS between Lassosum and PRSice at age 12 within trait (MeanDiff = -0.010, 95% CIs [-0.021; -0.001]), and at age 12 and 16 cross-trait. Lassosum-based EA3 GPS performed better within trait at age 16 (MeanDiff = -0.020, 95% CIs [-0.031; -0.009]).

No differences were found in prediction accuracy between LDpred- vs Lassosum-based IQ3 GPS or EA3 GPS, within or cross-trait. Supplementary Table S3a reports mean differences and CIs for these comparisons.

Multi-trait polygenic score prediction

Results of multivariate GWA analyses are reported in Supplementary Methods S4 and Supplementary Tables S5 and S6. Here we report results of polygenic score associations for our best predictive polygenic models after multi-trait approaches were applied to GWA summary statistics (Figure 3 and Figure S5). Figure S6 shows a comparison of variance predicted in intelligence and educational achievement at ages 12 and 16 in the test samples across polygenic score methods after multi-trait analyses. Supplementary Table S4 reports details of these results.

Figure 2 presents variance predicted in intelligence and educational achievement at age 16 by polygenic scores derived from multi-trait methods. For intelligence, variance predicted by IQ3 GPS increased from 6.7% (Figure 1) to a maximum of 10.0% ($\beta = 0.327$, $se = 0.032$, $p < .0001$) at age 16. For educational achievement, variance predicted by EA3 GPS increased from 14.8% to a maximum of 15.9% ($\beta = 0.403$, $se = 0.018$, $p < .0001$) at age 16. Again, EA3 GPS was generally the best performing predictor across phenotypes, predicting a maximum of 10.6% ($\beta = 0.332$, $se = 0.031$, $p < .0001$) in intelligence. Similar improvements in prediction were observed at age 12 (see Supplementary Table S4 and supplementary figure S4).

Supplementary Figure S6 shows a test of the differences in predictive performance of Lassosum-based scores between multi-traits methods tested at age 12 and 16. There were no significant differences between multi-trait methods for both IQ3 and EA3 GPS across all phenotypes. The only exceptions were the SMTpred IQ3 score, which tended to perform better than MTAG at age 16 cross-trait (MeanDiff = -0.011, 95% CIs [-0.022; -0.001]), and the MTAG EA3 score which tended to perform better than Genomic SEM at age 16 within trait (MeanDiff = -0.0077, 95% CIs [-0.0143; -0.002]). Supplementary Table S4 reports mean differences and CIs for these comparisons.

Polygenic scores quantile differences—Figure 3 shows the results for the best predictive models at age 16 by GPS deciles. For both intelligence (panel a) and educational

achievement (panel b), the relationship with GPS deciles is linear and the lowest and highest deciles differ substantially. For intelligence, the mean difference (~1 SD) is comparable to 15 IQ points. For educational achievement, the mean difference corresponds to an average 'C' grade for the lowest decile and an average 'A' grade for the highest decile. However, the range of distributions in the lowest and highest deciles overlap considerably, as would be expected from GPS correlations of ~0.32 with intelligence and ~0.40 with educational achievement.

Sex differences—We tested associations for the best prediction model (i.e. MTAG EA3 GPS calculated in Lassosum) separately for males and females in the test set. For intelligence at age 16, the GPS predicted 10.7% of the variance (95% CIs [6.33;16.74]) in males (N= 369, $\beta = 0.334$, se = 0.049) and 10.5% (95% CIs [6.49;15.41]) in females (N = 558, $\beta = 0.329$, se = 0.040). For educational achievement in males (N = 1,105) the GPS predicted 14.2% (95% CIs [10.96;17.86]) of the variance ($\beta = 0.375$, se = 0.027); in females (N = 1,300) estimates were 17.2% (95% CIs [13.51;21.43]; $\beta = 0.420$, se = 0.025). To test the significance of these sex differences, we performed a Fisher's r to z transformation of corresponding correlation coefficients. Sex differences were not significant for intelligence (observed z = -0.066, p = 0.472) nor educational achievement (Observed z = -1.419, p = 0.077).

Multiple regression model—We compared the results from our multi-trait GPS analyses to a simple multiple regression using the five GPS from summary statistics of our multi-trait analyses (IQ, EA, income, age when completed full time education, time spent using computer) to predict intelligence and educational achievement. The multiple regression model predicted similar amounts of variance as the best single multi-trait GPS predictors. For intelligence, the adjusted R² was 8.6% at age 12 and 9.9% at age 16. For educational achievement, the adjusted R² was 9.6% at age 12 and 16.7% at age 16. Results are shown in Supplementary Table S7.

Discussion

Using summary statistics from the latest GWA studies of intelligence (IQ3¹⁴) and educational attainment (EA3¹⁷), we report the strongest polygenic prediction estimates for cognitive-related traits to date. Comparing standard polygenic score approaches, we showed that IQ3 GPS predicts a maximum of 6.73% of the variance in intelligence at age 16, while EA3 GPS predicts 14.78% of the variance in educational achievement at age 16.

In an attempt to boost predictive power, we compared results using state-of-the-art genomics methods that leverage the multivariate nature of traits in order to increase power of GWA summary statistics. We then tested boosted summary statistics across a number of polygenic score approaches, showing that we can predict 10.6% of the variance in intelligence and 15.9% of the variance in educational achievement, both at age 16. These results compare favourably with polygenic prediction estimates from the recent EA3 GWA analysis, whereby a polygenic score constructed from multi-trait summary statistics of educational attainment and three cognitive-related phenotypes predicted up to 13% of the variance in educational attainment and up to 10% in cognitive performance¹⁷, this is especially notable given the

larger discovery sample size employed in that study ($N \sim 1.1$ million including 23andMe). We note that differences between these studies may be attributable to systematic differences at the level of trait measurement (e.g. accuracy of measurement) and sample characteristics (e.g. differences in ancestry; differences in heritability). Nevertheless, this is a good indication that a multi-trait approach to polygenic prediction replicates well across independent samples yielding robust prediction estimates.

We found that trait prediction increased from age 12 to age 16. Polygenic scores become more predictive with age, probably because as the sample approaches adulthood it is closer in age to the samples in which beta weights were estimated in the original GWA studies for IQ3 and EA3. Another possible reason for this finding is that given that heritability of intelligence increases with age³³, the variance that can be predicted by cognitive-related polygenic scores also increases. Lastly, we did not find significant differences in the predictive power of IQ3 and EA3 for males and females.

These results indicate the usefulness of taking into account the multivariate nature of complex traits in polygenic prediction, and add to the possibility of practical use of polygenic scores at the level of individuals³⁴. It is important to note that we randomly split our sample (~50%) to validate our models and assessed performance of prediction models in the test sample in order to avoid overfitting. Because TEDS is a representative sample of the UK population, these prediction estimates are expected to be a close representation of how these models would perform in similar samples. Overall, multi-trait methods were successful in increasing variance predicted; compared to our 'baseline' predictions, estimates increased from 1% to 3%. Multi-trait methods were especially useful in increasing predictive power of the IQ3 GPS, which was constructed using less powerful summary statistics than the EA3 GPS. However, differences in prediction accuracy across the tested combinations of genomic methods seemed to reflect differences in polygenic score approaches rather than in multi-traits approaches. An indication of this intuition was also provided by a formal comparison of R^2 estimates, which showed no consistent differences across multi-trait methods. Yet, reassuringly, there were no dramatic differences in prediction accuracy across polygenic score approaches either, especially when considering approaches that do not perform clumping (thereby losing information across the genome).

One limitation that could affect the interpretability of our findings is that by jointly analysing traits with differing levels of power and genetic overlap, the multi-trait methods considered here might confound the genetic architecture of boosted traits with that of other traits. In this regard, genetic correlations between traits before and after multi-trait analyses and with a control trait, as those reported in Supplementary Methods S4, may indicate the degree to which the genetic architecture of one trait has 'shifted' towards that of others in the multi-trait analysis. This is an important post-hoc test to be considered by future studies employing multi-trait approaches in the context of polygenic prediction.

An ongoing debate concerns the causal mechanisms by which polygenic scores predict phenotypes such as educational achievement and intelligence. Passive gene-environment correlation may be a mechanism underlying the association between polygenic scores and educational attainment. Given parent-child shared genetics (~50%), if EA trait-increasing

variants are correlated with rearing environments which in turn are contributing to attainment, GWAS estimates obtained for EA would be partly picking up genetic effects mediated via the environment. That is, GWAS effect estimates may be due to indirect genetic effects via rearing environments that could reflect both inherited and non-inherited parental DNA. Therefore, the association between an individual's EA polygenic score and cognitive traits could partly reflect an environmentally transmitted parental genetic effect^{17, 35, 36}. Analyses relying on family-based designs have put forward evidence in this regard^{17, 37, 38}. These studies confirmed what have long been acknowledged by twin and adoption studies on the nature of nurture^{39,40}. Separating the different mechanisms of gene-environment interplay by which polygenic scores influence complex traits is an important area of research. However, prediction of individual differences in behavioural phenotypes from polygenic scores can be achieved without an underlying explanatory model.

Finally, a general limitation of all genomic analyses is that they only assess additive effects of common SNPs used on currently SNP arrays. SNP heritability is the ceiling for polygenic score prediction, which is about 20%¹⁴ of the total variance for intelligence and 30%⁴¹ for educational achievement. Viewed in this light, our best polygenic scores predict about half of the SNP heritability. With bigger and better GWA studies and other methodological advances like multivariate approaches, the missing SNP heritability gap will be narrowed. Polygenic scores will only reach their full potential when we are able to close the gap between SNP heritability (about 25%) and family study estimates of heritability (about 50%).

Nonetheless, these polygenic scores predictions are already among the strongest predictors in the behavioural sciences. Because inherited DNA variants do not change during development, polygenic scores are unique predictors in two ways. First, unlike other characteristics of the individual, DNA variants can predict individual differences in adult behaviour from birth. Second, unlike other correlations, associations between DNA variants and behaviour are causal from DNA to behaviour in the sense that there can be no backward causation from behaviour to DNA. These unique features will put genomic prediction of cognitive traits in the front line of the DNA revolution.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We gratefully acknowledge the ongoing contribution of the participants in the Twins Early Development Study (TEDS) and their families. TEDS is supported by a program grant to RP from the UK Medical Research Council (MR/M021475/1 and previously G0901245), with additional support from the US National Institutes of Health (AG046938). The research leading to these results has also received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ grant agreement n° 602768 and ERC grant agreement n° 295366. RP is supported by a Medical Research Council Professorship award (G19/2). This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 721567.

References

1. Plomin R, von Stumm S. The new genetics of intelligence. *Nature reviews Genetics* 2018; 19(3): 148–159.
2. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA et al. 10 Years of GWAS Discovery: Biology, Function, and Translation. *American journal of human genetics* 2017; 101(1): 5–22. [PubMed: 28686856]
3. Pasaniuc B, Price AL. Dissecting the genetics of complex traits using summary association statistics. *Nature reviews Genetics* 2017; 18(2): 117–127.
4. Turley P, Walters RK, Maghzian O, Okbay A, Lee JJ, Fontana MA et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nature genetics* 2018; 50(2): 229–237. [PubMed: 29292387]
5. Maier RM, Zhu Z, Lee SH, Trzaskowski M, Ruderfer DM, Stahl EA et al. Improving genetic prediction by leveraging genetic correlations among human diseases and traits. *Nature Communications* 2018; 9(1): 989.
6. Benyamin B, Pourcain B, Davis OS, Davies G, Hansell NK, Brion MJ et al. Childhood intelligence is heritable, highly polygenic and associated with FBNP1L. *Molecular psychiatry* 2014; 19(2): 253–258. [PubMed: 23358156]
7. Butcher LM, Davis OS, Craig IW, Plomin R. Genome-wide quantitative trait locus association scan of general cognitive ability using pooled DNA and 500K single nucleotide polymorphism microarrays. *Genes, brain, and behavior* 2008; 7(4): 435–446.
8. Davies G, Armstrong N, Bis JC, Bressler J, Chouraki V, Giddaluru S et al. Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N=53949). *Molecular psychiatry* 2015; 20(2) : 183–192. [PubMed: 25644384]
9. Davies G, Marioni RE, Liewald DC, Hill WD, Hagenaars SP, Harris SE et al. Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). *Molecular psychiatry* 2016; 21(6): 758–767. [PubMed: 27046643]
10. Davies G, Tenesa A, Payton A, Yang J, Harris SE, Liewald D et al. Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Molecular psychiatry* 2011; 16(10): 996–1005. [PubMed: 21826061]
11. Plomin R, Hill L, Craig IW, McGuffin P, Purcell S, Sham P et al. A genome-wide scan of 1842 DNA markers for allelic associations with general cognitive ability: a five-stage design using DNA pooling and extreme selected groups. *Behavior genetics* 2001; 31(6): 497–509. [PubMed: 11838529]
12. Trampush JW, Yang ML, Yu J, Knowles E, Davies G, Liewald DC et al. GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium. *Molecular psychiatry* 2017; 22(3): 336–345. [PubMed: 28093568]
13. Sniekers S, Stringer S, Watanabe K, Jansen PR, Coleman JRI, Krapohl E et al. Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nature genetics* 2017; 49 (7): 1107–1112. [PubMed: 28530673]
14. Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, de Leeuw CA et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature genetics* 2018; 50(7): 912–919. [PubMed: 29942086]
15. Rietveld CA, Medland SE, Derringer J, Yang J, Esko T, Martin NW et al. GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science (New York, NY)* 2013; 340(6139): 1467–1471.
16. Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA et al. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 2016; 533(7604): 539–542. [PubMed: 27225129]
17. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature genetics* 2018.

18. Krapohl E, Plomin R. Genetic link between family socioeconomic status and children's educational achievement estimated from genome-wide SNPs. *Molecular psychiatry* 2016; 21(3): 437–443. [PubMed: 25754083]
19. Selzam S, Krapohl E, von Stumm S, O'Reilly PF, Rimfeld K, Kovas Y et al. Predicting educational achievement from DNA. *Molecular psychiatry* 2018; 23(1): 161. [PubMed: 28948970]
20. Rietveld CA, Esko T, Davies G, Pers TH, Turley P, Benyamin B et al. Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proceedings of the National Academy of Sciences of the United States of America* 2014; 111(38): 13790–13794. [PubMed: 25201988]
21. Krapohl E, Patel H, Newhouse S, Curtis CJ, von Stumm S, Dale PS et al. Multi-polygenic score approach to trait prediction. *Molecular psychiatry* 2018; 23(5): 1368–1374. [PubMed: 28785111]
22. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. *Bioinformatics* 2015; 31(9): 1466–1468. [PubMed: 25550326]
23. Vilhjalmsón BJ, Yang J, Finucane HK, Gusev A, Lindstrom S, Ripke S et al. Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *American journal of human genetics* 2015; 97(4): 576–592. [PubMed: 26430803]
24. Mak TSH, Porsch RM, Choi SW, Zhou X, Sham PC. Polygenic scores via penalized regression on summary statistics. *Genetic epidemiology* 2017; 41(6): 469–480. [PubMed: 28480976]
25. Hill WD, Hagenaars SP, Marioni RE, Harris SE, Liewald DCM, Davies G et al. Molecular Genetic Contributions to Social Deprivation and Household Income in UK Biobank. *Current Biology* 2016; 26(22): 3083–3089. [PubMed: 27818178]
26. Seed C. Hail: An Open-Source Framework for Scalable Genetic Data. 2017.
27. Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD et al. Genomic SEM Provides Insights into the Multivariate Genetic Architecture of Complex Traits. *bioRxiv* 2018.
28. Haworth CM, Davis OS, Plomin R. Twins Early Development Study (TEDS): a genetically sensitive investigation of cognitive and behavioral development from childhood to young adulthood. *Twin research and human genetics : the official journal of the International Society for Twin Studies* 2013; 16(1): 117–125. [PubMed: 23110994]
29. Oliver BR, Plomin R. Twins' Early Development Study (TEDS): a multivariate, longitudinal genetic investigation of language, cognition and behavior problems from childhood through adolescence. *Twin research and human genetics : the official journal of the International Society for Twin Studies* 2007; 10(1): 96–105. [PubMed: 17539369]
30. Yang J, Ferreira T, Morris AP, Medland SE, Madden PA, Heath AC et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature genetics* 2012; 44(4): 369–375, s361–363. [PubMed: 22426310]
31. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics* 2007; 81(3): 559–575. [PubMed: 17701901]
32. Krapohl E, Euesden J, Zabaneh D, Pingault JB, Rimfeld K, von Stumm S et al. Phenome-wide analysis of genome-wide polygenic scores. *Molecular psychiatry* 2015; 21: 1188. [PubMed: 26303664]
33. Haworth CMA, Wright MJ, Luciano M, Martin NG, de Geus EJC, van Beijsterveldt CEM et al. The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular psychiatry* 2010; 15(11): 1112–1120. [PubMed: 19488046]
34. Plomin R. *Blueprint: How DNA Makes Us Who We Are*. Allen Lane/Penguin Press: London, 2018.
35. Fletcher JM, Lehrer SF. Genetic lotteries within families. *Journal of health economics* 2011; 30(4): 647–659. [PubMed: 21664708]
36. Pingault J-B, O'Reilly PF, Schoeler T, Ploubidis GB, Rijdsdijk F, Dudbridge F. Using genetic data to strengthen causal inference in observational research. *Nature Reviews Genetics* 2018; 19(9): 566–580.

37. Belsky DW, Domingue BW, Wedow R, Arseneault L, Boardman JD, Caspi A et al. Genetic analysis of social-class mobility in five longitudinal studies. *Proceedings of the National Academy of Sciences* 2018; 115(31): E7275–E7284.
38. Kong A, Thorleifsson G, Frigge ML, Vilhjalmsón BJ, Young AI, Thorgeirsson TE et al. The nature of nurture: Effects of parental genotypes. *Science (New York, NY)* 2018; 359(6374): 424–428.
39. Plomin R, Bergeman CS. The nature of nurture: Genetic influence on “environmental” measures. *Behavioral and Brain Sciences* 2011; 14(3): 373–386.
40. Plomin R. *Genetics and experience: The interplay between nature and nurture*. Sage Publications: Thousand Oaks, CA, 1994.
41. Krapohl E, Plomin R. Genetic link between family socioeconomic status and children’s educational achievement estimated from genome-wide SNPs. *Molecular psychiatry* 2015; 21: 437. [PubMed: 25754083]

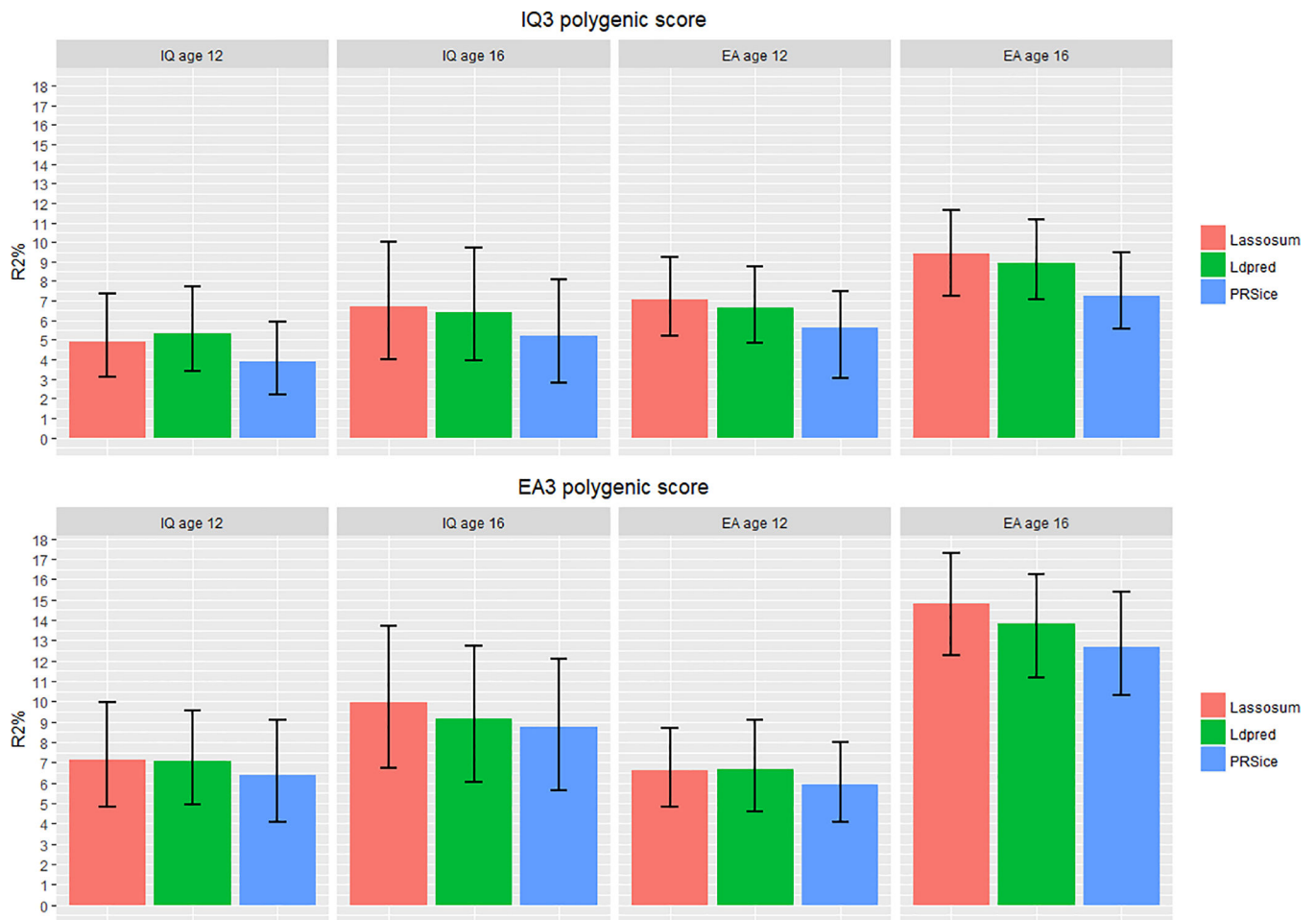


Figure 1. Polygenic score prediction of intelligence (IQ) and educational achievement (EA) at age 12 and 16. Figure shows polygenic prediction accuracy across polygenic score methods. Error bars are bootstrapped 95% confidence intervals based on 1,000 replications.

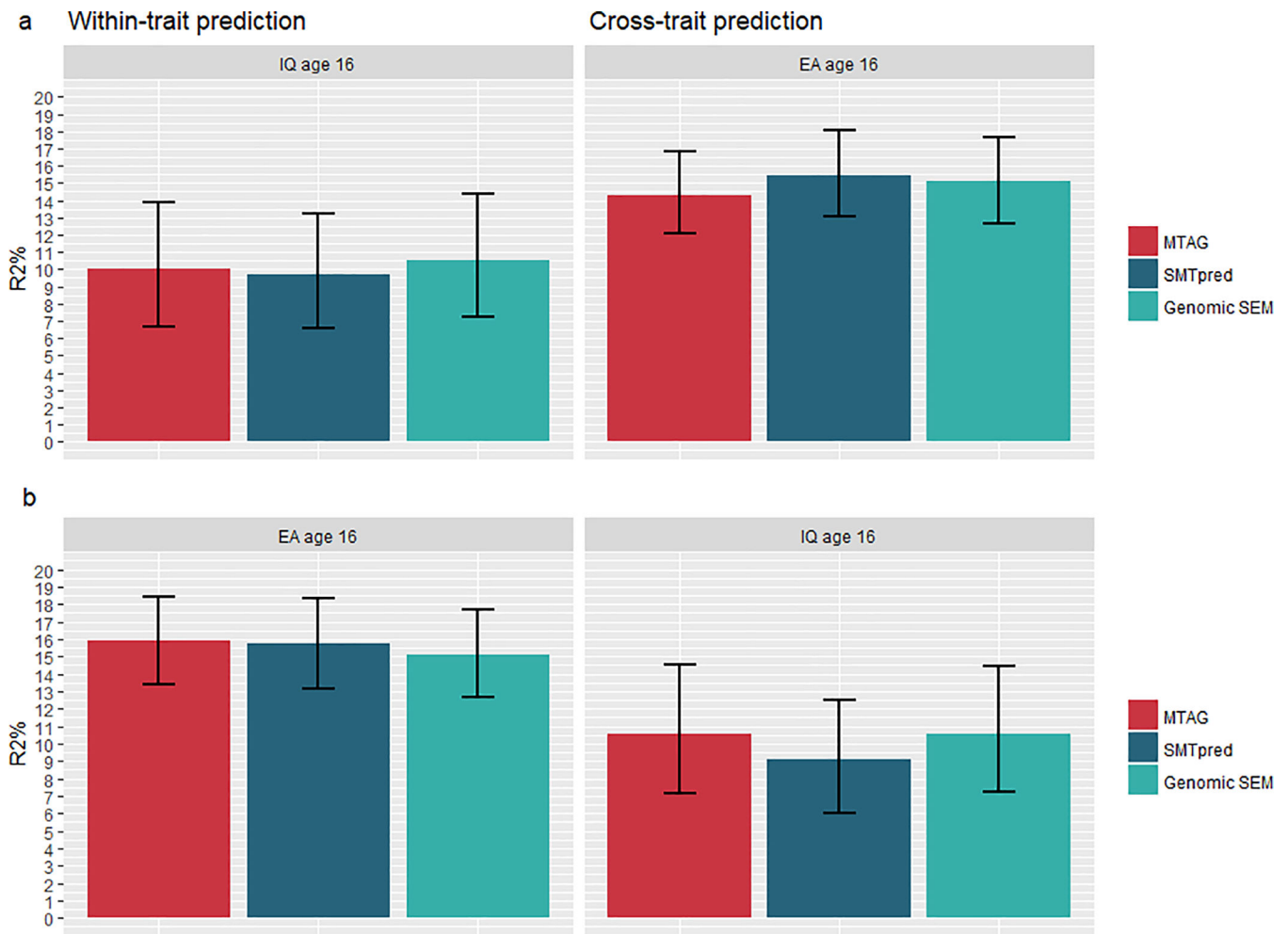


Figure 2. Within-trait and cross-trait polygenic score prediction of intelligence and educational achievement at age 16 across multi-trait methods.
Note. MTAG = MTAG IQ3 (panel a)/ MTAG EA3 (panel b) polygenic scores constructed in Lassosum; SMTpred = IQ3 (panel a)/EA3 (panel b) wMT-SBLUP predictors; Genomic SEM = Common Factor polygenic score constructed in Lassosum (panel a and b). Error bars are bootstrapped 95% confidence intervals based on 1,000 replications.

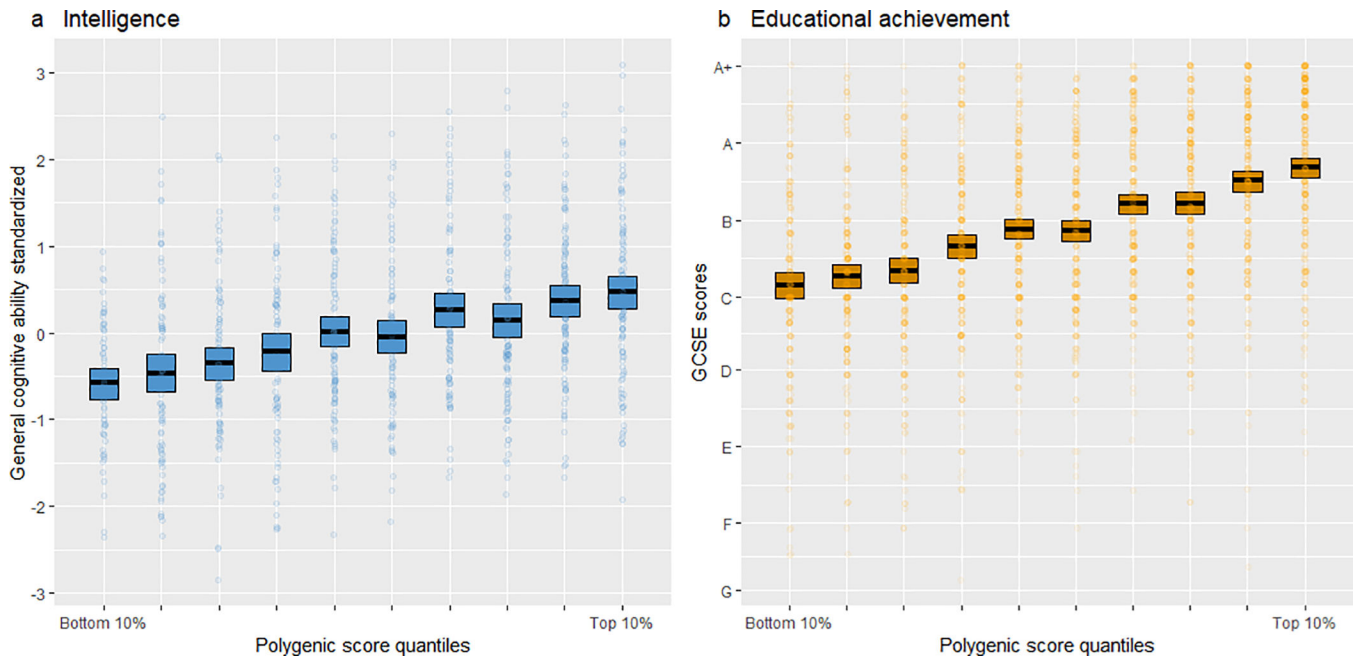


Figure 3. Mean intelligence scores (panel **a**) and mean educational achievement (panel **b**; GCSE grades) at age 16 by GPS deciles for the best polygenic predictors in the test set. Bars represent bootstrapped 95% confidence intervals. Coloured dots represent individual data points.