



Polygenic Scores for ADHD: A Meta-Analysis

James J. Li^{1,2,3} · Quanfa He^{1,2}

Accepted: 18 January 2021 / Published online: 25 January 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC part of Springer Nature 2021

Abstract

Attention-deficit/hyperactivity disorder (ADHD) is a highly heritable neurodevelopmental disorder that is known to have a polygenic (i.e., many genes of individually small effects) architecture. Polygenic scores (PGS), which characterize this polygenicity as a single score for a given individual, are considered the state-of-the-art in psychiatric genetics research. Despite the proliferation of ADHD studies adopting this approach and its clinical implications, remarkably little is known about the predictive utility of PGS in ADHD research to date, given that there have not yet been any systematic or meta-analytic reviews of this rapidly developing literature. We meta-analyzed 12 unique effect sizes from ADHD PGS studies, yielding an $N=40,088$. These studies, which included a mixture of large population-based cohorts and case–control samples of predominantly European ancestry, yielded a pooled ADHD PGS effect size of $r_{\text{random}}=0.201$ (95% CI = [0.144, 0.288]) and an $r_{\text{fixed}}=0.190$ (95% CI = [0.180, 0.199]) in predicting ADHD. In other words, ADHD PGS reliably account for between 3.6% (in the fixed effects model) to 4.0% (in the random effects model) of the variance in broadly defined phenotypic ADHD. Findings provide important insights into the genetics of psychiatric outcomes and raise several key questions about the impact of PGS on psychiatric research moving forward. Our review concludes by providing recommendations for future research directions in the use of PGS, including new methods to account for comorbidities, integrating bioinformatics to elucidate biological pathways, and leveraging PGS to test mechanistic models of ADHD.

Keywords Polygenic scores · Genome-wide association study · ADHD · Psychiatric genetics · Meta-analysis

Attention-deficit/hyperactivity disorder (ADHD) is a neurodevelopmental disorder with an estimated worldwide prevalence of 7.2% (Thomas et al., 2015). ADHD is reliably associated with a range of adverse social, behavioral, and emotional outcomes in later life, including low educational attainment (Kuriyan et al., 2013), depression and substance misuse (Agnew-Blais et al., 2018), criminality (Fletcher & Wolfe, 2009), and poor physical health (Kuriyan et al., 2013). The negative impact of childhood ADHD is pervasive in ways beyond its effects on the individual; it also affects society through the substantial economic burden it places on

mental health systems and education services (Doshi et al., 2012).

Although decades of behavioral and molecular genetics research have indicated the critical role that genes play in the etiology of ADHD (Chang et al., 2013; Chen et al., 2017; Faraone & Larsson, 2018; Nikolas & Burt, 2010), recent studies leveraging genome-wide methods have begun to challenge some of our previous knowledge about the genetics of ADHD while ushering in a new set of tools to our pursuit of uncovering its genetic architecture. One method that has gained considerable traction in psychiatric genetics research are polygenic scores (PGS) (Wray et al., 2007). Study adoption of PGS methods has been rapid; from 2007 to 2013, there were a total of 138 published studies cataloged on PubMed involving the use of PGS (three of which were related to ADHD). In just 2018 *alone*, there were 279 studies published on PGS, which is more than double the total number of published PGS studies prior to 2013. Given the high degree of heritability of many complex traits and disorders, there is optimism that PGS can eventually be used in the prediction of human health and behavioral outcomes

✉ James J. Li
james.li@wisc.edu

¹ Department of Psychology, University of Wisconsin, Madison, WI, USA

² Waisman Center, University of Wisconsin, WI, Madison, USA

³ Center for Demography of Health and Aging, University of Wisconsin, WI, Madison, USA

(Anderson et al., 2019; Martin et al., 2019; Torkamani et al., 2018). Despite its tremendous promise, little is known about whether psychiatric PGS are particularly useful for clinical prediction. Furthermore, as more studies involving PGS are published, it is paramount to consider the limitations and future directions of PGS research. This study examines the predictive performance of ADHD PGS by conducting the first meta-analytic review of the ADHD PGS literature.

Genome-Wide Association Studies of ADHD

Before discussing the PGS approach, we first summarize genome-wide association studies (GWAS) of ADHD, a critical aspect of PGS computations. GWAS examine the association between single nucleotide polymorphism (SNP) variation in the genome and a quantitative trait of interest (Hirschhorn & Daly, 2005). Most GWAS to-date have employed microarrays to genotype hundreds of thousands to several million SNPs in individuals. Because of the volume of SNP-phenotype associations tested and the expected small effect of even a significant association, GWAS have far larger sample size requirements than behavioral genetic or candidate gene studies (Visscher et al., 2017). Thus, nearly all major GWAS of psychiatric outcomes are led by consortia, consisting of international research teams that combine their cohorts to form a large pooled sample (e.g., Psychiatric Genomics Consortium; PGC) (Sullivan, 2010).

PGC-led efforts have produced two meta-analyses specific to ADHD GWAS, the first of which was performed by B. M. Neale and colleagues (2010). This GWAS was comprised of four independent child and adolescent cohorts of predominantly Western European descent, resulting in a pooled study sample of 2,064 parent-ADHD proband trios, 896 ADHD probands and 2,455 controls. Participants in all cohorts were assessed for ADHD using a semi-structured clinical interview (keyed to either the *DSM-IV* or *ICD-10*, depending on the population) conducted with a caregiver. The majority of ADHD probands were diagnosed with combined-type ADHD ($n = 868$). Genetic data were imputed using the HapMap Phase III reference panel, resulting in 1,206,463 SNPs analyzed in the GWAS. Test statistics from each cohort were transformed into Z scores, which was the effect size statistic in the GWAS. The GWAS yielded no genome-wide significant findings (i.e., $p < 5 \times 10^{-8}$). Regions on chromosome 7 (e.g., *SHFM1*), 8 (e.g., *CHMP7*), and 11 (e.g., *DHCR7* and *NADSYN1*) were implicated based on having the most SNPs that were among the top 50 associations, but the authors pointed to being underpowered in their sample to detect small genetic effects. Other noted limitations of the GWAS included: measurement variability and differences in ADHD referral patterns across cohorts, unmeasured effects from rare and

copy number variants, and unaccounted environmental differences (e.g., diet, culture) due to having drawn from different populations (Neale et al., 2010).

The most recent meta-analysis of ADHD GWAS was conducted by Demontis and colleagues (2019). The research team analyzed genetic data from 12 child and adult cohorts, resulting in 20,183 ADHD probands and 35,191 controls. However, the vast majority of the probands and controls ($n = 14,583$ and $22,494$ respectively) came from a single child and adult cohort from Denmark (iPSYCH). The 11 other cohorts were aggregated by the PGC and represented smaller child and adult samples from Europe, Canada, United States and China. To address the possibility of population stratification, genetic principal components were included in their analysis. Imputation was performed using the 1000 Genomes Project Phase 3 reference panel, resulting in 8,047,421 variants analyzed in the GWAS. In iPSYCH, ADHD status was determined by a psychiatrist according to the *ICD-10*. ADHD status in the PGC samples were assessed using semi-structured clinical interviews (e.g., Schedule for Affective Disorders and Schizophrenia for School-Age Children, K-SADS; Child and Adolescent Psychiatric Assessment, CAPA). Twelve unique SNPs were identified in the full GWAS sample, including a locus in the *FOXP2* gene, which is believed to play a role in learning and speech (Schreiweis et al., 2014) and *SEMA6D*, which is believed to play a role in embryonic brain development and educational attainment (Okbay et al., 2016). Notably, none of the regions that were implicated in the earlier B.M. Neale et al. (2010) GWAS were genome-wide significant in the more recent GWAS. Furthermore, one of the loci (in *SPAG16* on chromosome 2) failed to pass the significance threshold when only the European ancestry subsample was meta-analyzed. And the 12 loci that were genome-wide significant only captured a small fraction of variance in ADHD, with odds ratios for each of the loci ranging from 1.077 to 1.198 (the SNP heritability of ADHD was an estimated 0.22). A number of important limitations should be noted, including the high degree of age heterogeneity in the pooled sample (combining children and adults), diagnostic heterogeneity from the use of different measures of ADHD, and an overrepresentation of cases and controls from a single Western European cohort (i.e., iPSYCH). Results from these GWAS have made it abundantly clear that many genes with individually small effects, rather than a few genes with large effects, are likely involved in the etiology of ADHD. As sample sizes for these GWAS have increased, so too has the rate of genetic discovery (Visscher et al., 2017). Given the prominence of consortia-led efforts towards gene identification, GWAS sample sizes are projected to be large enough to capture a substantial amount of the common genetic variation underlying the psychiatric disorders in time (Sullivan, 2010).

Polygenic Scores (PGS)

PGS leverage of our knowledge from GWAS into a single score that characterizes an individual's polygenic (via common SNPs) liability for a trait of interest (Wray et al., 2007). Notably, PGS have been widely referred to as a polygenic “risk” scores (PRS) in the broader literature (Anderson et al., 2019; Bogdan et al., 2018). The inclusion of the term “risk” is perhaps a misnomer because the traditional scoring approach does not differentiate true risk alleles (i.e., those that confer an increased liability to a trait in question) from unassociated variants (The International Schizophrenia Consortium, 2009). Furthermore, there is emerging evidence that PGS may also have promotive associations, particularly for those on the “low” end of the distribution (Krapohl et al., 2016; Li, 2019b; Plomin et al., 2009; Torkamani et al., 2018). We therefore prefer the more general term “polygenic score,” which is the nomenclature used throughout the review.

The traditional PGS approach takes an ensemble of genetic variants from a GWAS sample, referred to as the “discovery” population, and computes a weighted linear composite of these variants for individuals sampled from an independent “target” population (Anderson et al., 2019; The International Schizophrenia Consortium, 2009):

$$\text{PGS} = \sum_i \text{SNP}_{ij} \beta_j$$

where the PGS for individual i in a target population is the summation of the total number of alleles, j , for a SNP in individual i multiplied by the SNP's effect size, β , taken from GWAS conducted on a separate discovery population. In PGS, the number of SNPs included in the computation is based on the p -value threshold (P_T) set in the discovery GWAS. A more liberal threshold, e.g. $P_T = 1$, captures much more of the genetic signal into the PGS computation, but also more unassociated variants as well. Conversely, setting a more conservative threshold, e.g. $P_T < 0.01$, captures less of the genetic signal but also fewer unassociated variants that may contribute to a noisy signal (Dudbridge, 2013). Although most studies have used $P_T < 0.5$ as a matter of convention (see example set by The International Schizophrenia Consortium, 2009), a “best fit” P_T can also be empirically derived by selecting the P_T that explains the largest amount of variance in the phenotype as measured in the target population (Euesden, Lewis, & O'Reilly, 2016). There are, however, issues with model overfit (Benjamini et al., 2001) and a lack of generalizability when using a P_T that is optimally-predictive only for a particular phenotype in a particular target population.

Purcell and colleagues (2009) published the first major application of PGS in psychiatric genetics (The International Schizophrenia Consortium, 2009). Using a discovery GWAS on schizophrenia, which at the time consisted of 3,322 European proband individuals and 3,587 controls, schizophrenia PGS (at GWAS $P_T < 0.5$) explained approximately 3% of the variance in schizophrenia in a completely independent target sample, with larger effects at increasingly more liberal GWAS P_T . Moreover, they replicated the schizophrenia PGS association in two European ancestry target samples, once again showing that schizophrenia PGS (at GWAS $P_T < 0.5$) explained between 2.3 – 3.2% of the variance in schizophrenia. A systematic review identified 31 studies involving the use of schizophrenia PGS in relation to a broad set of outcomes (Mistry et al., 2018b). The authors noted that a meta-analysis was not conducted due to a lack of enough information provided by most PGS studies as well excessive variability in both the types of outcomes examined and discovery populations across each study. The same authors also conducted a review of studies on bipolar and major depressive disorder PGS, making similarly limited conclusions with respect to the state-of-the-field (Mistry et al., 2018a). Excessive outcome heterogeneity and the current lack of reporting standards in PGS studies of psychiatric disorders obfuscates our knowledge about the predictive utility of PGS for these outcomes, which has implications on future research in this area and its clinical applications.

The Current Study

Recent years have seen psychiatric genetics research developed rapidly, given the combination of publicly-available GWAS summary statistics for psychiatric disorders, the relatively low cost of microarrays and genotyping (Visscher et al., 2017), and the availability of open-source software for GWAS and PGS computations (Euesden et al., 2016; Purcell et al., 2007). Despite the rapid emergence ADHD PGS studies, there has yet to be a systematic or meta-analytic review of this literature. This study provides the first meta-analysis of the ADHD PGS literature and provides recommendations for future PGS studies.

Due to the possibility of heterogeneity in the effect sizes of ADHD PGS across studies, we explored a number of potential moderators, including the GWAS discovery sample, publication year, sampling method of the target sample (i.e., population-based vs. case-control), informant type of the ADHD measurement (i.e., parent only vs. multi-informant), and whether the ADHD outcome was measured

categorically or continuously. Because of the greater power afforded by larger GWAS discovery samples (Chatterjee et al., 2013), we hypothesized that more recently published and larger discovery samples (i.e., Demontis et al., 2019) would be associated with larger PGS effect sizes than older studies using smaller discovery samples (i.e., Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; Neale et al., 2010). We also hypothesized that studies that focused on clinical populations (i.e., case–control) would have larger effect sizes than studies that sampled from general populations, especially given prior research showing that psychiatric PGS are more predictive in clinical populations than non-clinical ones (Savage et al., 2018). With respect to the various measurement methods used to assess ADHD across studies (e.g., categorical vs. continuous, parent-only vs. multi-informant), we make no specific hypotheses regarding genetic prediction estimates given that we are not aware of any prior studies that have considered analyzing the effects of measurement method on PGS effect sizes specifically.

Method

Eligibility Criteria

This review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009). Eligibility criteria for the meta-analysis were as follows: 1) published in a peer reviewed journal, 2) written in English, 3) used an ADHD GWAS discovery sample, 4) used an independent target

population (i.e., target participants cannot overlap with those in the ADHD GWAS discovery sample), 5) derived ADHD PGS on said target population, and 6) examined associations between ADHD PGS and ADHD (broadly defined). In the final meta-analyzed sample, we also removed studies that featured the same target sample, but used an older (i.e., lesser powered) discovery GWAS sample. The decision to focus on phenotypic ADHD in the ADHD PGS meta-analysis was driven by 1) the larger number of studies having measured this phenotype relative to a non-ADHD phenotype, and 2) the importance of establishing the construct validity of the ADHD PGS, which remains a crucial endeavor in the psychological sciences (Cronbach & Meehl, 1955). In order to facilitate comparisons between ADHD PGS studies as well as to reduce some of the heterogeneity that typifies this literature, we only included ADHD PGS studies that employed ADHD GWAS disseminated by the PGC: B. M. Neale et al. (2010), the Cross-Disorder Group of the Psychiatric Genomics Consortium (2013), and Demontis et al. (2019).

Search Procedure and Data Extraction

A flow chart of the meta-analysis procedure is presented in Supplemental Fig. 1. Data collection was conducted by Q.H. The review focused on all studies published from January 2010 to January 2020. Potential studies were identified in PubMed and Google Scholar databases. Searches targeted studies with all possible variants of the term “ADHD” (e.g., “attention deficit disorder,” “attention,” “hyperactivity”) and variants of the term “polygenic score” (e.g., “polygen*,” “polygenic risk scores,” “profile scores,” “risk profile scores”). A full list of the search criteria employed in the current review are listed in

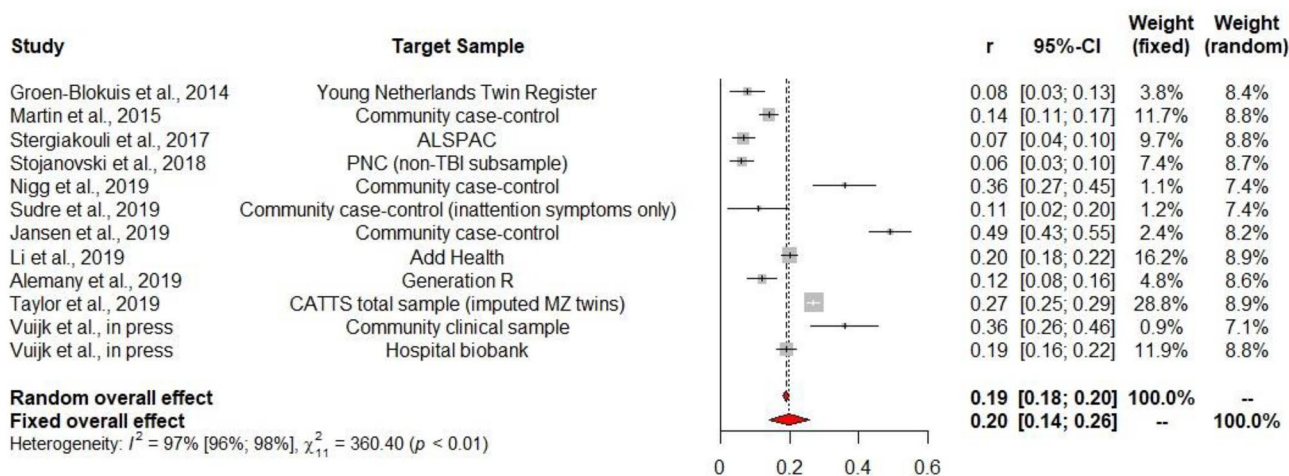


Fig. 1 Forest plot of ADHDPGS effect sizes

Supplemental Table 1. Study authors were contacted in cases where relevant information about a study was missing (e.g., sample demographics, association statistics, *p*-values, etc.).

Effect Sizes

To determine the overall effect size of the association between ADHD PGS on phenotypic ADHD, Pearson's *r* effect sizes were computed for each study (using the “*esc*” package in *R*), thus allowing us to combine studies that used a case–control design (i.e., logistic regression) with population-based cohorts (i.e., linear regression or correlations). We used analyzed (rather than total) sample sizes in the effect size computation. Some studies reported odds ratios (OR's) or standardized betas (β 's) for multiple ADHD phenotypes (e.g., parent versus teacher report, questionnaire versus interview) and at multiple GWAS P_T . Rather than including each of these effect sizes into the meta-analysis, we instead included only the strongest reported association of the ADHD PGS on ADHD for each study. We assumed that these estimates would be more consistent with the effects reported in studies that did not test multiple ADHD PGS associations at multiple thresholds and phenotypes.

Analyses

All analyses were performed in *R*, Version 3.6.2. The meta-analysis was performed in the *R* packages “*meta*”, “*dmatar*”, and “*metafor*”. We reported both fixed- and random-effects models to estimate the pooled effect size across ADHD PGS studies (Borenstein et al., 2010). Between-study heterogeneity was estimated using Cochran's *Q* test, which measures the weighted and summed difference between the observed effect sizes and the pooled effect, and Higgins' and Thompson's I^2 , which assesses the percentage of variability in effect sizes not due to sampling error (Higgins & Thompson, 2002). High heterogeneity may be suggestive of excessive study differences, which can complicate interpretations of the meta-analytic pooled effect size (Higgins & Thompson, 2002). For potential moderators, we also explored whether 1) the effects of the GWAS discovery sample that was used, 2) the publication year, 3) the sampling strategy of the target sample (i.e., case–control or population-based), 4) the primary method of measurement (i.e., parent report only vs. multi-informant), and 5) ADHD was measured categorically or continuously, moderated the association of ADHD PGS and phenotypic ADHD effect sizes via meta-regression. Finally, we evaluated the possibility of publication bias in two ways: the Egger's test of the intercept (Egger et al., 1997), which tests asymmetry in the effect size estimates relative to sample sizes, and *p*-curve analysis (Simonsohn et al., 2014), which plots the distribution of the *p*-values across studies and estimates the

true meta-analytic effect size after accounting for possible “*p*-hacking.”

Results

Summary of ADHD PGS Studies

There were 18 studies that met our criteria for the meta-analysis, although only 12 effect sizes were ultimately included in the meta-analysis, yielding a pooled $N=40,088$ (see Fig. 1). Three of the 18 studies (Albaugh et al., 2019; Benca et al., 2017; Hawi et al., 2018) did not report the information needed to compute effect sizes and could not be included in the meta-analysis. We also identified several studies that published on the same target samples: Child and Adolescent Twin Study in Sweden (CATSS) (Brikell et al., 2018; Taylor et al., 2019), Generation R (Alemany et al., 2019; P. R. Jansen et al., 2018), the Avon Longitudinal Study of Parents and Children (ALSPAC) (Riglin et al., 2016; Stergiakouli et al., 2017), and a community case–control (Nigg et al., 2018, 2019). To avoid sample overlap in our effect sizes, we included the study among each pair that either featured the most recent GWAS discovery sample (Generation R: Alemany et al., 2019; ALSPAC: Stergiakouli et al., 2017), the study that used that utilized a larger portion of the target sample (CATSS: Taylor et al., 2019; ALSPAC: Stergiakouli et al., 2017) or the study most recently published (Nigg et al., 2019). Table 1 provides an abbreviated summary of the 12 effect sizes included in the meta-analysis; more detailed characteristics of each study can be found in Supplemental Table 1.

We highlight several characteristics of the studies included in our meta-analysis (Table 1). The most used GWAS P_T was $P_T < 0.50$. ADHD was measured across studies in a variety of ways, including parent-rated questionnaires (e.g., Child Behavior Checklist, CBCL; Strengths and Difficulty Questionnaire, SDQ), semi or fully-structured clinical interview (e.g., Kiddie Schedule for Affective Disorders and Schizophrenia; K-SADS, Diagnostic Interview Schedule for Children; DISC) or a combination of both methods. Seven studies included in the meta-analysis used population-based target samples where ADHD was measured continuously (e.g., symptom counts or latent factors), except for Li (2019a, b). All studies included in the meta-analysis accounted for age, child sex, and population stratification effects via genetic principal components (PCs) as covariates in their association analyses. However, not all studies included the same number of genetic PCs (see Supplemental Table 3 for additional details about each study). All studies consisted of target populations that were predominantly or entirely of Western European descent.

Meta-analysis, Heterogeneity, and Moderation Analysis

In the random-effects model, the pooled effect size of the ADHD PGS was $r_{\text{random}} = 0.201$, 95% CI = [0.144, 0.288], $p < 0.001$ ($r^2_{\text{random}} = 0.040$) (see Fig. 1). In the fixed-effects model, the pooled effect size of the ADHD PGS was $r_{\text{fixed}} = 0.190$, 95% CI = [0.180, 0.199], $p < 0.001$ ($r^2_{\text{fixed}} = 0.036$). However, there was evidence of excessive heterogeneity, Cochran's $Q = 360.40$, $df = 11$, $p < 0.001$; $I^2 = 96.9\%$, 95% CI = [95.9–97.8%]. A multiple meta-regression analysis showed significant associations of sampling [i.e., population-based (0) vs case-control (1)] of the target sample, $B = 0.128$, $se = 0.050$, $p = 0.015$, 95% CI = [0.030, 0.225] and the use of continuous (1) vs. categorical (2) measures of ADHD, $B = 0.115$, $se = 0.051$, $p = 0.025$, 95% CI = [0.014, 0.216] on the effect sizes of the PGS across studies. That is, studies with target samples that featured cases and controls ($n = 5$ studies) and used categorical measures of ADHD ($n = 5$ studies) reported greater effect sizes than studies that used population-based ($n = 7$ studies) sampling methods and continuous measures of ADHD ($n = 7$ studies). There was no significant association of the discovery sample, $B = 0.057$, $se = 0.075$, $p = 0.446$, 95% CI = [-0.090, 0.205], publication year, $B = 0.001$, $se = 0.027$, $p = 0.960$, 95% CI = [-0.052, 0.054], and informant type, $B = -0.048$, $se = 0.060$, $p = 0.426$, 95% CI = [-0.166, 0.070] on the effect sizes. Notably, only 3 of the 12 studies in our meta-analysis used a non-Demontis et al. (2019) GWAS discovery sample, and only 2 studies of the 12 studies used a multi-informant measurement method as opposed to a parent-only method. Thus, one reason why we may not have detected any moderating effects of the discovery sample or informant method is due to the relative lack of variability in these variables. Overall, 73.7% of the heterogeneity of the effect sizes was accounted for by the five moderators we tested.

Publication Bias

The Egger's test for effect size asymmetry showed a non-significant deviation from the $Y = 0$ intercept ($B = -0.580$, 95% CI = [-8.028, 6.868], $p = 0.883$) and no clear evidence of publication bias. We then performed a p -curve analysis, a tool that corrects for inflated effect sizes that publication bias produces. For any given sample size, the bigger the "true" effect, the more right-skewed the expected p -curve (Simonsohn et al., 2014). If a "true" effect exists, one expects lower significant p -values than higher significant p -values. Figure 2 shows the results of the p -curve analysis of the 12 meta-analyzed effect sizes. The p -curve analysis indicated that at 99% power, 92% of the p -values are < 0.01 .

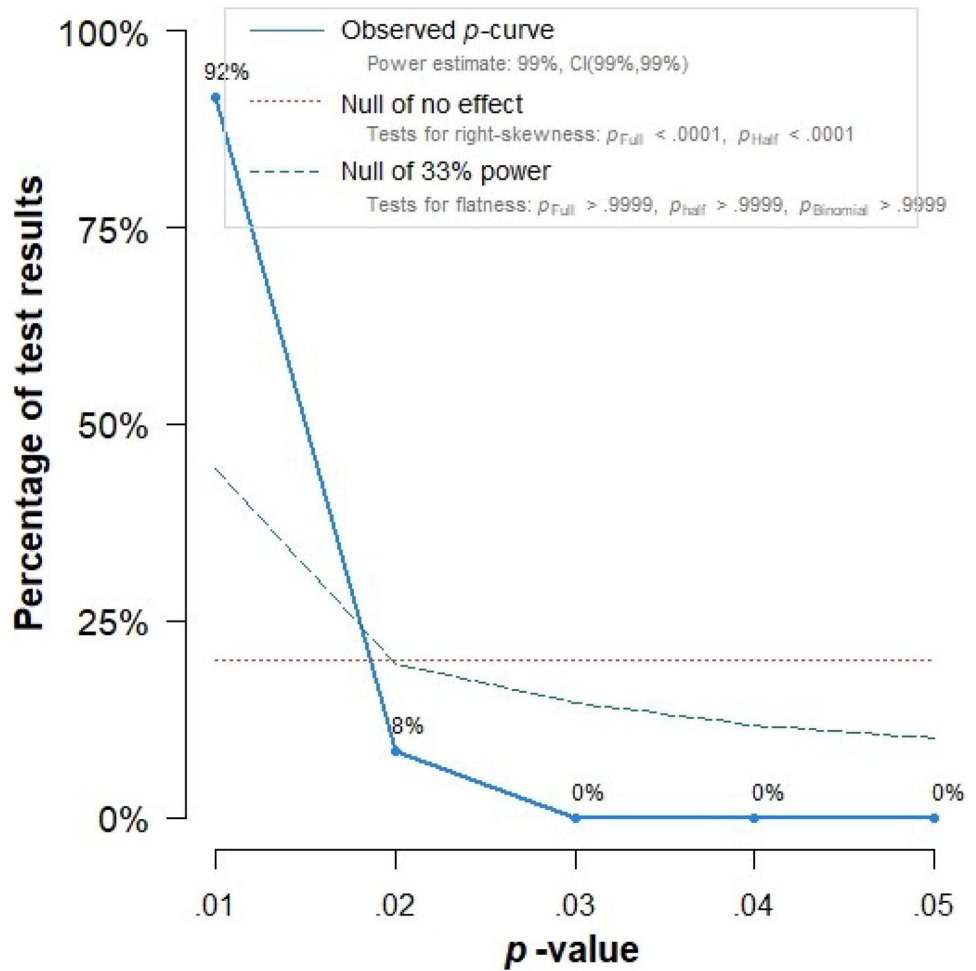
In fact, all 12 of the studies had a p -values lower than 0.025, reflecting a significant right-skewedness of the p -curve ($p < 0.0001$). Collectively, results from Egger's test and p -curve analysis suggest no strong evidence of publication bias or "p-hacking" in the meta-analysis.

Discussion

To our knowledge, this is the first meta-analysis of ADHD PGS studies. The meta-analysis included 12 unique effect sizes spanning population-based and case-control target samples (albeit, predominantly of European ancestry) with a pooled $N = 40,008$. Overall, ADHD PGS were consistently and significantly associated with phenotypic ADHD across studies. ADHD PGS accounted for between 3.6% (in the fixed effects model) to 4.0% (in the random effects model) of the variance in broadly defined phenotypic ADHD. This prediction estimate is in line with PGS estimates observed for other psychiatric disorders with similarly large GWAS samples, including schizophrenia (Mistry et al., 2018b), bipolar disorder (Mistry et al., 2018a; Stahl et al., 2019), and autism spectrum disorder (Grove et al., 2019). This estimate reflects the tremendous progress being made in psychiatric genetics, especially when considering that just over a decade ago the dominant methodology for directly quantifying genetic risk for complex traits like ADHD was via candidate genes, an approach that not only failed to capture much of the variation in ADHD, but also required us to suspend our long-held belief that most complex traits are driven by polygenic, rather than monogenic, influences (Lander & Schork, 1994). As we enter the post-GWAS era, our findings provide important insights into the state-of-the-science in the genetics of ADHD and lay a foundation for future directions in this field.

Results indicate that ADHD PGS reliably predict ADHD symptoms and diagnosis across different samples, with every study reporting a statistically significant effect size. However, there was excessive heterogeneity as we observed a wide range of reported effect sizes. For instance, ADHD PGS effect sizes were $r = 0.062$ ($r^2 = 0.004$) on the lowest end (Stojanovski et al., 2019) and $r = 0.490$ ($r^2 = 0.240$) on the highest end (Jansen et al., 2019). Stojanovski et al. (2019) may have been relatively underpowered given that they conducted their association analyses on subsamples of individuals with and without traumatic brain injuries in the Philadelphia Neurodevelopmental Cohort (PNC), a population-based dataset children, adolescents and young adults from the United States. In contrast, Jansen et al.'s (2019) target sample was ascertained from a psychiatric outpatient hospital in the Netherlands that oversampled for children with ADHD (controls were recruited from

Fig. 2 *p*-curve analysis of publication bias and *p*-hacking *Note.* The *p*-curve reflects the distribution of *p*-values as a function of the “true” underlying effect. For any given sample size, the bigger the effect, the more right-skewed the expected *p*-curve becomes (Simonsohn et al., 2014). If a “true” effect exists, one expects lower significant *p*-values than higher significant *p*-values. The observed *p*-curve indicates that at 99% power, all 12 statistically significant results (100%) had $p < 0.025$



the general population). In fact, studies with the largest effect sizes in our meta-analysis were typically outpatient or case–control populations that oversampled or focused on individuals with ADHD and related disorders (Jansen et al., 2019; Nigg et al., 2019; Vuijk et al., 2019). The GWAS cohorts in the discovery sample were largely recruited from clinical populations themselves, potentially making them a better ‘match’ to other case–control studies (Savage et al., 2018).

Qualitative differences between the discovery sample and target populations might have contributed to stronger predictive performance of ADHD PGS in clinical or case–control samples relative to population-based samples. The classic liability threshold model (Gottesman & Shields, 1967) provides a plausible theory for why matching the composition of the discovery and target samples in PGS studies may be crucial for optimizing the PGS signal. Clinical populations may thus reflect a different genetic liability distribution than those in the general population, as individuals from clinical populations also typically present with more severe psychopathology and additional comorbidities (Savage et al., 2018). Thus, sampling

from a clinical population in the ADHD GWAS may have attenuated the ADHD PGS predictive effect in the general population (Supplemental Fig. 2), where a higher PGS would be required to meet the liability threshold in a general population than a clinical population. Existing GWAS discovery samples for ADHD seem optimized for studies where the target population is also a case–control sample and/or a clinical population, and possibly less so for samples featuring a general population. We note, however, that there may be other differences besides sampling type that could contribute to differences in the PGS distribution (e.g., age of the sample, informants that were used to phenotype, ancestry, etc.). Future studies should examine these effects by comparing the performance of PGS based on various discovery samples as they pertain to different target populations. Furthermore, the excessive study heterogeneity could also be partly due to the different GWAS P_T used across studies, although most studies used $P_T < 0.50$. We already discussed the problem with this convention, including issues with model overfit (Benjamini et al., 2001) and a lack of generalizability when using a standard P_T across studies and populations. As it stands,

Table 1 Abbreviated summary of ADHD PGS studies included in meta-analysis

Author and year	Discovery GWAS	Target sample	Analytic N	Target ages	Strongest reported effect size	S.E. or C.I	<i>p</i> *
Groen-Blokuis et al., 2014	Cross Disorder PGC (2013)	Young Netherlands Twin Register	1612	7–13	$\beta=0.08$	<i>S.E.</i> =0.03	0.0067
Martin et al., 2015	Neale (2010)	"UK-based" cases and controls	4670	6–17	OR = 1.15	95% CI=1.04–1.27	0.009
Stergiakouli et al., 2017	Cross Disorder (2013)	Avon Longitudinal Study of Parents and Children (ALSPAC)	4164	7–17	$\beta=0.07$	<i>S.E.</i> = .03	0.0043
Stojanovski et al., 2019	Demontis (2019)—Caucasian Only	Philadelphia Neurodevelopmental Cohort (PNC), No-TBI subgroup	1233	8–21	$T=3.5, df=1224$	Not reported	0.004
Nigg et al., 2019	Demontis (2019)—Caucasian Only	Community case–control	514	7–11	OR = 1.43; $r^2=0.033$	CI = 1.17–1.75	0.0004
Sudre et al., 2020	Demontis (2019)	Community case–control	489	> 3	$\beta=0.11$	<i>s.e.</i> = .046	0.02
Jansen et al., 2019	Demontis (2019)	Community case–control	280	3–18	$\beta=0.49$; OR = 1.625; $r^2=0.045$	Not reported	4.7E-08
Aleman et al., 2019	Demontis (2019)	Generation R	1053	8–11	$\beta=0.120$	<i>s.e.</i> = .000	<0.0001
Li et al., 2019	Demontis (2019)	National Longitudinal Study of Adolescent to Adult Health (Add Health)	7088	< 12	OR = 1.22	CI = 1.10–1.36	<0.001
Taylor et al., 2019	Demontis (2019)	Child and Adolescent Twin Study in Sweden (CATSS)	13,412	9–12	$\beta=0.268$	<i>s.e.</i> = .029	<0.0001
Vuijk et al., in press	Demontis (2019)	Community clinical sample	433	7–18	OR = 1.44, $r^2=0.020$	CI = 1.14–1.81	0.0011
Vuijk et al., in press	Demontis (2019)	Hospital biobank	5140	19–60	OR = 1.21, $r^2=0.004$	CI = 1.07–1.37	0.0028

Additional information about the ADHD PGS studies are available in Supplemental Table 1

*Decimal values shown here reflect the number of the decimals that were reported in the original article

there is currently no established, universal, or agreed upon GWAS P_T with which to specify in deriving a PGS. Few, if any psychiatric PGS studies have employed a *separate* discovery population, independent from the GWAS, to select an optimal P_T for the target population. At the same time, selecting an optimal GWAS P_T this way limits the validity and reliability of PGS as predictors since it means that a different GWAS P_T will have to be used each time a different sample is analyzed. We recommend future studies consider reporting PGS effects sizes using a GWAS P_T of 1.0, which incorporates all available genotypic information without the need for model selection. While it is highly unlikely that all genes contribute to the various outcomes, the computation of the PGS accordingly downweights trivial variants based its small (or null) GWAS effect size.

Limitations of the Meta-analysis

Some limitations of the meta-analysis should be noted. First, we focused our review on studies that exclusively used ADHD GWAS summary statistics that were publicly available via the PGC, rather than studies that used GWAS summary statistics from smaller/local samples. This decision was motivated by our desire to enhance the comparability across ADHD PGS studies and to subsequently perform a meta-analysis of these studies. However, we have already noted that the small overall effect size of the ADHD PGS across studies may have been due to the discrepancy between the ADHD discovery sample (i.e., case–control samples) and the target samples (predominantly population-based). Meta-analyzed GWAS of population-based samples for ADHD

(e.g., UK Biobank, 23andMe) thus offer a compelling direction for a future GWAS in this area. Second, we noted that some studies reported multiple ADHD outcomes at multiple P_T , thus providing multiple effect sizes with which to include in our meta-analysis. We mentioned that we selected the most significant effect size that was reported in each study, in part because we assumed that studies that did not report multiple effect sizes likely also just reported what was most robust. Thus, the pooled effect size reported in this meta-analysis is likely more liberal of an estimate than if we had only included the least significant effect sizes across studies. Third, nearly every study in our meta-analysis (as well as the GWAS discovery cohorts) consisted of individuals of predominantly (Western) European ancestries. PGS are known to be imprecise in non-European populations because these populations are underrepresented in GWAS discovery samples (Martin et al., 2019). Although well-powered transancestral GWAS studies of psychiatric outcomes have just recently emerged (Bigdeli et al., 2017; Walters et al., 2018), the lack of racial diversity is concerning and has the potential for exacerbating health disparities if and when genomic information becomes widely used for clinical applications (Martin et al., 2019). Fourth, all studies in our meta-analysis assumed an underlying linearity with respect to the PGS. This ignores the possibility of non-linear or non-additive allelic effects in the PGS computation, which is problematic when considering that not all genes have the same mode of inheritance. Additionally, emerging studies have shown that polygenic risks may not be linearly associated with psychiatric risks, especially at the upper end of the PGS distribution (Khera et al., 2018; Li, 2019b).

Future Directions

Accounting for comorbidity There is shared genetic variation between *most* of the major disorders in the *DSM* (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013). Psychiatric comorbidity complicates the interpretation of PGS for a single disorder because these scores likely contain a mixture of pleiotropic variants and variants that are unique to the disorder. It is important to disambiguate the genetic signals that are shared from those that are specific to discrete disorders in order to enhance the prediction accuracy of trait-specific PGS. In the absence of accounting for this phenomenon, it is unsurprising that psychiatric GWAS, and by extension, PGS studies informed by these GWAS, have yet to yield any high impact discoveries and large effect sizes, respectively.

Most well-powered psychiatric GWAS do not allow for the direct identification of higher-order or pleiotropic SNP effects because they typically employ case–control samples, where cases are recruited and measured on the presence of

a single (or a few related) disorder of interest. *Genomic Structural Equation Modeling (GSEM)* was developed as a way to identify the variants that affect cross-trait liability (i.e., influencing a general factor) and trait-specific liability leveraging only the GWAS summary statistics for these phenotypes (Grotzinger et al., 2019). One of the more promising applications of *GSEM* is in PGS. The proof of this concept was demonstrated by modeling a general factor PGS from GWAS summary statistics of schizophrenia and major depressive disorder in *GSEM* (Grotzinger et al., 2019). The general factor PGS was more predictive of the phenotypic general factor, psychotic experiences, depression, mania, anxiety and post-traumatic stress disorder in the completely independent UK Biobank dataset than any univariate version of the PGS (Grotzinger et al., 2019). Their findings not only show what a substantial effect that the general factor has on any single psychiatric outcome, but also potential gains in prediction accuracy and specificity when we control for the general factor disorder specific PGS studies.

Integrating bioinformatics into PGS computations A common criticism of PGS is that the presence of a statistical association does not necessarily reveal information about the biological mechanisms underlying the trait (Schaub et al., 2012; Subramanian et al., 2005). While gene expression panels such as the Encyclopedia of DNA Elements (The ENCODE Project Consortium, 2011) and the Epigenomics Roadmap Project (Roadmap Epigenomics Consortium et al., 2015) have generated a wealth of data on the linkages between genetic variation and gene function, their applications for PGS have yet to fully develop and take hold in the field. Software are available to functionally annotate GWAS summary statistics from gene expression panels like ENCODE and the Roadmap Project (e.g., *DAVID*; Dennis et al., 2003; *GenoSkyline*; Lu et al., 2016). These approaches integrate bioinformatics with GWAS to partition the SNP heritability by how strongly enriched the genetic signals are in human cells and tissues. Using *GenoSkyline*, which corrects for multiple testing and linkage disequilibrium (LD), genes for ADHD have been shown to be significantly overrepresented in the anterior caudate of the brain, although enrichment in other brain regions were also implicated to a lesser degree (Supplemental Fig. 3). Another gene expression assay for ADHD GWAS (which did not control for multiple testing or LD) detected gene enrichment across a broader range of brain regions, including the anterior cingulate gyrus (ACC), the anterior caudate, and the dorsolateral prefrontal cortex (DLPFC) (Demontis et al., 2019). These structures collectively play a crucial role in human reward processing and decision-making (Volkow et al., 2011). Thus, in the case of ADHD, functional annotations of the GWAS confirmed the regions of the brain that we had long suspected in its etiology (Li,

2018; Luman et al., 2005), but they also provide us with clues as to which brain regions are more relevant to ADHD than others.

Once GWAS summaries have been mapped onto gene expression panels, the resultant information can then be directly leveraged into PGS computations, where SNPs that are overrepresented in functionally annotated sites can be prioritized and weighted in the PGS computation. This contrasts with the traditional approach to computing PGS that prioritize SNPs based solely on statistical significance. PGS using highly enriched GWAS signals via *AnnoPred* significantly outperformed prediction estimates from a traditional PGS approach in predicting several complex phenotypes, including Crohn's disease, breast cancer, rheumatoid arthritis, Type-II diabetes, and celiac disease (Hu et al., 2017). We are not aware of any applications of this approach in psychiatric genetics at this time but integrating bioinformatics into PGS computations can promisingly lead to new hypotheses or provide evidence in support of existing hypotheses with respect to the biological processes underlying many complex traits.

Testing mechanistic theories of psychiatric outcomes PGS can also be leveraged to validate psychological constructs or to test new or existing models of heritable psychiatric disorders. One empirical framework is the Polygenic-Phenotypic Mediation Model (PPMM) (Li, 2019a; Li et al., 2019), which involves comparing two mediational models of a given trait B: 1) a phenotypic model, in which the independent variable for trait A is measured using traditional psychological methods, such as self-report or a behavioral paradigm, and 2) a polygenic model, in which the independent variable is characterized by a PGS of trait A. Both models are tested in relation to another trait B, where concordance between the phenotypic and polygenic mediation models provides robust validation of a mechanism involved in trait B and robust evidence of a genetic relationship between the two traits (Supplemental Fig. 4). The strength of PPMM is that the two models complement each other by addressing their respective limitations. Polygenic models, which are derived mechanically from completely independent GWAS datasets, are impervious to many of the methodological confounds that typify a traditional phenotypic model (i.e., common method variance, self-reporter bias). Phenotypic models on the other hand, address limitations of the polygenic model by virtue of providing construct validity of the PGS, whereby the genetic associations between the two traits also reflect observed associations as well. Furthermore, temporally separating measures of traits A and B allow researchers to test even more robust models of causality, particularly in cases where specific mechanisms are theorized to intervene in these pathways.

PPMM was first illustrated in the context of ADHD in examining that causal pathways of psychosocial risk from early childhood ADHD to later antisocial behaviors in adulthood (Li, 2019a). Using the PPMM framework, results of the study revealed that children with ADHD exhibited greater difficulties in school, which in turn predicted their antisocial development during adulthood even after accounting for the simultaneous effects of supportive parenting and peer effects during adolescence. Importantly, this indirect effect was completely replicated in the polygenic model of ADHD. Another study using the same PPMM framework showed that phenotypic and polygenic models of neuroticism indirectly predict late life depression via mid-life stressful life events and low social support in the Wisconsin Longitudinal Study (Li et al., 2019). Collectively, these studies show how PPMM could be used to validate theories about the psychosocial mechanisms involved in complex psychiatric development.

The two study examples focused on psychosocial factors as *mechanisms* rather than as *moderators* of genetic liability. This represents a significant departure from convention in the gene-environment literature. In fact, gene-environment correlation effects are quite pervasive in the psychopathology literature and are likely to explain why gene-environment interactions are so difficult to detect and replicate (Knafo & Jaffee, 2013). Furthermore, PPMM can be used to test more proximal mechanisms of genetic liability, including neurobiological or cognitive *endophenotypes* (Gottesman & Gould, 2003). One of the studies included in the meta-analysis (Nigg et al., 2018) showed that ADHD PGS was indirectly associated with phenotypic ADHD via working memory and arousal/alertness. New paradigms such as PPMM can be used to incorporate measures at multiple levels of analysis, including those rigorously assayed in the laboratory, to shed new light on biological as well as psychosocial mechanisms of risk for ADHD and other psychiatric outcomes (Hinshaw, 2018).

Conclusion

The rapid rise in PGS studies published in just the last half decade reflects the multidisciplinary enthusiasm the approach has garnered, from scientists who view PGS as a new model for explaining the genetic variation underlying complex traits (Wray et al., 2014), to those who view PGS as a tool for clinical prediction and diagnosis (Anderson et al., 2019; Martin et al., 2019; Torkamani et al., 2018). Prediction estimates from ADHD PGS should become even more robust as our GWAS sample sizes increase (Chatterjee et al., 2016). For instance, the increase from just a few thousand (B.

M. Neale et al., 2010) to several tens of thousands of ADHD cases and controls (Demontis et al., 2019) in GWAS have only just enabled the detection of genome-wide significant findings for ADHD. We note that simulations of genetic data indicate that increases in the proportion of phenotypic variance explained by PGS in several complex traits is non-linearly associated with the GWAS sample size (Chatterjee et al., 2013). Greater prediction accuracy depends on the heritability of the trait as well, where more heritable traits (e.g., height) require fewer samples to achieve the same gains in prediction accuracy than less heritable and presumably more complex traits. PGS are projected to be a major part of precision medicine moving forward as our technology improves and the science becomes more refined.

At the same time, it is also important to mitigate any unbridled enthusiasm around this conceivably not-too-distant future where our genes are used to predict our health and behavior. The current lack of diversity in GWAS studies of ADHD crucially limits the utility of ADHD PGS to those with Western European-ancestries. Unless substantially greater efforts are made to collect genetic samples from a broader range (i.e., non-Western European descent) of populations, PGS studies will only continue to be disproportionately focused on Western European-descent populations (Martin et al., 2019). Furthermore, history has been fraught with infamous examples of the social consequences that stemmed from strong beliefs about genetic determinism (Epstein, 2003; Galton, 1904). Prominent experts have already warned about the potential invasions of privacy and cavalier presentations of genetic information that direct-to-consumer genetic testing companies have introduced to mass audiences (Zettler et al., 2014). As psychiatric genetics continues to rise, we must be careful not to also perpetuate the old and dangerous stereotype that our destiny is in our genes. Despite the valuable information they provide about us, genes alone do not determine who we become.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10802-021-00774-4>.

Acknowledgements The authors acknowledge contributions from Alexis Hosch and Hannah Glasrud, who provided valuable assistance during the initial phases of the study. J.J.L. was supported in part by a core grant to the Waisman Center from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (U54 HD090256) and the Wisconsin Alumni Research Foundation. J.J.L. and Q.H. were also supported by the UW-Madison Center for Human Genomics and Precision Medicine.

Compliance with Ethical Standards

Conflicts of interest The authors have no relevant financial or non-financial interests to disclose.

Ethical approval This is a meta-analytic review of existing literature. No human subjects were involved in this study. Per UW-Madison's Exempt Research Policy, no ethical approval was required.

References

- Agnew-Blais, J. C., Polanczyk, G. V., Danese, A., Wertz, J., Moffitt, T. E., & Arseneault, L. (2018). Young adult mental health and functional outcomes among individuals with remitted, persistent and late-onset ADHD. *The British Journal of Psychiatry*, *213*(3), 526–534. <https://doi.org/10.1192/bjp.2018.97>.
- Albaugh, M. D., Hudziak, J. J., Ing, A., Chaarani, B., Barker, E., Jia, T., & Potter, A. (2019). White matter microstructure is associated with hyperactive/inattentive symptomatology and polygenic risk for attention-deficit/hyperactivity disorder in a population-based sample of adolescents. *Neuropsychopharmacology*, *44*(9), 1597–1603. <https://doi.org/10.1038/s41386-019-0383-y>.
- Aleman, S., Jansen, P. R., Muetzel, R. L., Marques, N., El Marroun, H., Jaddoe, V. W. V., & White, T. (2019). Common Polygenic Variations for Psychiatric Disorders and Cognition in Relation to Brain Morphology in the General Pediatric Population. *Journal of the American Academy of Child & Adolescent Psychiatry*, *58*(6), 600–607. <https://doi.org/10.1016/j.jaac.2018.09.443>.
- Anderson, J. S., Shade, J., DiBlasi, E., Shabalin, A. A., & Docherty, A. R. (2019). Polygenic risk scoring and prediction of mental health outcomes. *Current Opinion in Psychology*, *27*, 77–81. <https://doi.org/10.1016/j.copsyc.2018.09.002>.
- Benca, C. E., Derringer, J. L., Corley, R. P., Young, S. E., Keller, M. C., Hewitt, J. K., & Friedman, N. P. (2017). Predicting Cognitive Executive Functioning with Polygenic Risk Scores for Psychiatric Disorders. *Behavior Genetics*, *47*(1), 11–24. <https://doi.org/10.1007/s10519-016-9814-2>.
- Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., & Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behavioural Brain Research*, *125*(1), 279–284. [https://doi.org/10.1016/S0166-4328\(01\)00297-2](https://doi.org/10.1016/S0166-4328(01)00297-2).
- Bigdeli, T. B., Ripke, S., Peterson, R. E., Trzaskowski, M., Bacanu, S.-A., Abdellaoui, A., & Kendler, K. S. (2017). Genetic effects influencing risk for major depressive disorder in China and Europe. *Translational Psychiatry*, *7*(3), e1074. <https://doi.org/10.1038/tp.2016.292>.
- Bogdan, R., Baranger, D. A. A., & Agrawal, A. (2018). Polygenic Risk Scores in Clinical Psychology: Bridging Genomic Risk to Individual Differences. *Annual Review of Clinical Psychology*, *14*(1), 119–157. <https://doi.org/10.1146/annurev-clinpsy-050817-084847>.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97–111. <https://doi.org/10.1002/jrsm.12>.
- Brikell, I., Larsson, H., Lu, Y., Pettersson, E., Chen, Q., Kujala, H., & Martin, J. (2018). The contribution of common genetic risk variants for ADHD to a general factor of childhood psychopathology. *Molecular Psychiatry*. <https://doi.org/10.1038/s41380-018-0109-2>.
- Chang, Z., Lichtenstein, P., Asherson, P. J., & Larsson, H. (2013). Developmental Twin Study of Attention Problems: High Heritabilities Throughout Development. *JAMA Psychiatry*, *70*(3), 311–318. <https://doi.org/10.1001/jamapsychiatry.2013.287>.
- Chatterjee, N., Shi, J., & García-Closas, M. (2016). Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nature Reviews Genetics*, *17*(7), 392–406. <https://doi.org/10.1038/nrg.2016.27>.

- Chatterjee, N., Wheeler, B., Sampson, J., Hartge, P., Chanock, S. J., & Park, J.-H. (2013). Projecting the performance of risk prediction based on polygenic analyses of genome-wide association studies. *Nature Genetics*, *45*(4), 400–405. <https://doi.org/10.1038/ng.2579>.
- Chen, Q., Brikell, I., Lichtenstein, P., Serlachius, E., Kuja-Halkola, R., Sandin, S., & Larsson, H. (2017). Familial aggregation of attention-deficit/hyperactivity disorder. *Journal of Child Psychology and Psychiatry*, *58*(3), 231–239. <https://doi.org/10.1111/jcpp.12616>.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. <https://doi.org/10.1037/h0040957>.
- Cross-Disorder Group of the Psychiatric Genomics Consortium. (2013). Identification of risk loci with shared effects on five major psychiatric disorders: A genome-wide analysis. *Lancet (London, England)*, *381*(9875), 1371–1379. [https://doi.org/10.1016/S0140-6736\(12\)62129-1](https://doi.org/10.1016/S0140-6736(12)62129-1).
- Demontis, D., Walters, R. K., Martin, J., Mattheisen, M., Als, T. D., Agerbo, E., & Neale, B. M. (2019). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics*, *51*(1), 63. <https://doi.org/10.1038/s41588-018-0269-7>.
- Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., & Lempicki, R. A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology*, *4*(9), R60. <https://doi.org/10.1186/gb-2003-4-9-r60>.
- Doshi, J. A., Hodgkins, P., Kahle, J., Sikirica, V., Cangelosi, M. J., Setyawan, J., & Neumann, P. J. (2012). Economic Impact of Childhood and Adult Attention-Deficit/Hyperactivity Disorder in the United States. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(10), 990–1002.e2. <https://doi.org/10.1016/j.jaac.2012.07.008>.
- Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLOS Genetics*, *9*(3), e1003348. <https://doi.org/10.1371/journal.pgen.1003348>.
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. <https://doi.org/10.1136/bmj.315.7109.629>.
- Epstein, C. J. (2003). Is modern genetics the new eugenics? *Genetics in Medicine*, *5*(6), 469–475. <https://doi.org/10.1097/01.GIM.0000093978.77435.17>.
- Euesden, J., Lewis, C. M., & O'Reilly, P. F. (2016). *PRSice: Polygenic Risk Score software v1.25*. 24.
- Faraone, S. V., Larsson, H. (2018) Genetics of attention deficit hyperactivity disorder. *Molecular Psychiatry* *1*. <https://doi.org/10.1038/s41380-018-0070-0>.
- Fletcher, J., & Wolfe, B. (2009). Long-term Consequences of Childhood ADHD on Criminal Activities. *The Journal of Mental Health Policy and Economics*, *12*(3), 119–138.
- Galton, F. (1904). Eugenics: Its Definition, Scope, and Aims. *American Journal of Sociology*, *10*(1), 1–25.
- Gottesman, I. I., & Shields, J. (1967). A polygenic theory of schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*, *58*(1), 199–205.
- Gottesman, I. I., & Gould, T. D. (2003). The Endophenotype Concept in Psychiatry: Etymology and Strategic Intentions. *American Journal of Psychiatry*, *160*(4), 636–645. <https://doi.org/10.1176/appi.ajp.160.4.636>.
- Groen-Blokhuis, M. M., Middeldorp, C. M., Kan, K. J., Abdellaoui, A., Van Beijsterveldt, C. E., Ehli, E. A., ... & Psychiatric Genomics Consortium ADHD Working Group. (2014). Attention-deficit/hyperactivity disorder polygenic risk scores predict attention problems in a population-based sample of children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*, 1123–1129. <https://doi.org/10.1016/j.jaac.2014.06.014>.
- Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., & Tucker-Drob, E. M. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature Human Behaviour*, *1*. <https://doi.org/10.1038/s41562-019-0566-x>.
- Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., & Børglum, A. D. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics*, *51*(3), 431. <https://doi.org/10.1038/s41588-019-0344-8>.
- Hawi, Z., Yates, H., Pinar, A., Arnatkeviciute, A., Johnson, B., Tong, J., & Bellgrove, M. A. (2018). A case-control genome-wide association study of ADHD discovers a novel association with the tenascin R (TNR) gene. *Translational Psychiatry*, *8*(1), 1–8. <https://doi.org/10.1038/s41398-018-0329-x>.
- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. <https://doi.org/10.1002/sim.1186>.
- Hinshaw, S. P. (2018). Attention Deficit Hyperactivity Disorder (ADHD): Controversy, Developmental Mechanisms, and Multiple Levels of Analysis. *Annual Review of Clinical Psychology*, *14*(1), 291–316. <https://doi.org/10.1146/annurev-clinpsy-050817-084917>.
- Hirschhorn, J. N., & Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics*, *6*(2), 95–108. <https://doi.org/10.1038/nrg1521>.
- Hu, Y., Lu, Q., Powles, R., Yao, X., Yang, C., Fang, F., & Zhao, H. (2017). Leveraging functional annotations in genetic risk prediction for human complex diseases. *PLOS Computational Biology*, *13*(6), e1005589. <https://doi.org/10.1371/journal.pcbi.1005589>.
- Jansen, A. G., Dieleman, G. C., Jansen, P. R., Verhulst, F. C., Posthuma, D., & Polderman, T. J. C. (2019). Psychiatric Polygenic Risk Scores as Predictor for Attention Deficit/Hyperactivity Disorder and Autism Spectrum Disorder in a Clinical Child and Adolescent Sample. *Behavior Genetics*. <https://doi.org/10.1007/s10519-019-09965-8>.
- Jansen, P. R., Polderman, T. J. C., Bolhuis, K., van der Ende, J., Jaddoe, V. W. V., Verhulst, F. C., & Tiemeier, H. (2018). Polygenic scores for schizophrenia and educational attainment are associated with behavioural problems in early childhood in the general population. *Journal of Child Psychology and Psychiatry*, *59*(1), 39–47. <https://doi.org/10.1111/jcpp.12759>.
- Khera, A. V., Chaffin, M., Aragam, K. G., Haas, M. E., Roselli, C., Choi, S. H., & Kathiresan, S. (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics*, *50*(9), 1219. <https://doi.org/10.1038/s41588-018-0183-z>.
- Knafo, A., & Jaffee, S. R. (2013). Gene-environment correlation in developmental psychopathology. *Development and Psychopathology*, *25*(1), 1–6. <https://doi.org/10.1017/S0954579412000855>.
- Krapohl, E., Euesden, J., Zabaneh, D., Pingault, J.-B., Rimfeld, K., von Stumm, S., & Plomin, R. (2016). Phenome-wide analysis of genome-wide polygenic scores. *Molecular Psychiatry*, *21*(9), 1188–1193. <https://doi.org/10.1038/mp.2015.126>.
- Kuriyan, A. B., Pelham, W. E., Molina, B. S. G., Waschbusch, D. A., Gnagy, E. M., Sibley, M. H., & Kent, K. M. (2013). Young Adult Educational and Vocational Outcomes of Children Diagnosed with ADHD. *Journal of Abnormal Child Psychology*, *41*(1), 27–41. <https://doi.org/10.1007/s10802-012-9658-z>.
- Lander, E. S., & Schork, N. J. (1994). Genetic dissection of complex traits. *Science*, *265*(5181), 2037–2048. <https://doi.org/10.1126/science.8091226>.
- Li, J. J., (2018) Children's Reward and Punishment Sensitivity Moderates the Association of Negative and Positive Parenting Behaviors in Child ADHD Symptoms. *Journal of Abnormal Child Psychology*, 1–14. <https://doi.org/10.1007/s10802-018-0421-y>.
- Li, J. J. (2019a). Assessing phenotypic and polygenic models of ADHD to identify mechanisms of risk for longitudinal trajectories of

- externalizing behaviors. *Journal of Child Psychology and Psychiatry*, 0(0). doi: <https://doi.org/10.1111/jcpp.13071>
- JJ Li 2019 The positive end of the polygenic score distribution for ADHD: A low risk or a protective factor? *Psychological Medicine* 1–10 <https://doi.org/10.1017/S0033291719003039>
- Li, J. J., Hilton, E. C., Lu, Q., Hong, J., Greenberg, J. S., & Mailick, M. R. (2019). Validating psychosocial pathways of risk between neuroticism and late life depression using a polygenic score approach. *Journal of Abnormal Psychology*, 128(3), 200–211. <https://doi.org/10.1037/abn0000419>.
- Lu, Q., Powles, R. L., Wang, Q., He, B. J., & Zhao, H. (2016). Integrative Tissue-Specific Functional Annotations in the Human Genome Provide Novel Insights on Many Complex Traits and Improve Signal Prioritization in Genome Wide Association Studies. *PLoS Genetics*, 12(4), e1005947. <https://doi.org/10.1371/journal.pgen.1005947>.
- Luman, M., Oosterlaan, J., & Sergeant, J. A. (2005). The impact of reinforcement contingencies on AD/HD: A review and theoretical appraisal. *Clinical Psychology Review*, 25(2), 183–213. <https://doi.org/10.1016/j.cpr.2004.11.001>.
- Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., & Daly, M. J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics*, 51(4), 584. <https://doi.org/10.1038/s41588-019-0379-x>.
- Martin, J., Hamshere, M. L., Stergiakouli, E., O'Donovan, M. C., & Thapar, A. (2015). Neurocognitive abilities in the general population and composite genetic risk scores for attention-deficit hyperactivity disorder. *Journal of Child Psychology and Psychiatry*, 56, 648–656. <https://doi.org.ezproxy.library.wisc.edu/10.1111/jcpp.12336>.
- Mistry, S., Harrison, J. R., Smith, D. J., Escott-Price, V., & Zammit, S. (2018a). The use of polygenic risk scores to identify phenotypes associated with genetic risk of bipolar disorder and depression: A systematic review. *Journal of Affective Disorders*, 234, 148–155. <https://doi.org/10.1016/j.jad.2018.02.005>.
- Mistry, S., Harrison, J. R., Smith, D. J., Escott-Price, V., & Zammit, S. (2018b). The use of polygenic risk scores to identify phenotypes associated with genetic risk of schizophrenia: Systematic review. *Schizophrenia Research*, 197, 2–8. <https://doi.org/10.1016/j.schres.2017.10.037>.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097–e1000097. <https://doi.org/10.1371/journal.pmed.1000097>.
- Neale, B. M., Medland, S. E., Ripke, S., Asherson, P., Franke, B., Lesch, K.-P., & Nelson, S. (2010). Meta-Analysis of Genome-Wide Association Studies of Attention-Deficit/Hyperactivity Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(9), 884–897. <https://doi.org/10.1016/j.jaac.2010.06.008>.
- Nigg, J. T., Gustafsson, H. C., Karalunas, S. L., Ryabinin, P., McWeeney, S. K., Faraone, S. V., & Wilmot, B. (2018). Working Memory and Vigilance as Multivariate Endophenotypes Related to Common Genetic Risk for Attention-Deficit/Hyperactivity Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 57(3), 175–182. <https://doi.org/10.1016/j.jaac.2017.12.013>.
- Nigg, J. T., Karalunas, S. L., Gustafsson, H. C., Bhatt, P., Ryabinin, P., Mooney, M. A., ... Wilmot, B. (2019). Evaluating chronic emotional dysregulation and irritability in relation to ADHD and depression genetic risk in children with ADHD. *Journal of Child Psychology and Psychiatry*, n/a(n/a). doi: <https://doi.org/10.1111/jcpp.13132>
- Nikolas, M. A., & Burt, S. A. (2010). Genetic and environmental influences on ADHD symptom dimensions of inattention and hyperactivity: A meta-analysis. *Journal of Abnormal Psychology*, 119(1), 1–17. <https://doi.org/10.1037/a0018010>.
- Okbay, A., Beauchamp, J. P., Fontana, M. A., Lee, J. J., Pers, T. H., Rietveld, C. A., & Benjamin, D. J. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature*, 533(7604), 539–542. <https://doi.org/10.1038/nature17671>.
- Plomin, R., Haworth, C. M. A., & Davis, O. S. P. (2009). Common disorders are quantitative traits. *Nature Reviews Genetics*, 10(12), 872–878. <https://doi.org/10.1038/nrg2670>.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, 81(3), 559–575. <https://doi.org/10.1086/519795>.
- Riglin, L., Collishaw, S., Thapar, A. K., Dalsgaard, S., Langley, K., Smith, G. D., & Thapar, A. (2016). Association of Genetic Risk Variants With Attention-Deficit/Hyperactivity Disorder Trajectories in the General Population. *JAMA Psychiatry*, 73(12), 1285–1292. <https://doi.org/10.1001/jamapsychiatry.2016.2817>.
- Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., ... Kellis, M. (2015). Integrative analysis of 111 reference human epigenomes. *Nature*, 518(7539), 317–330. <https://doi.org/10.1038/nature14248>.
- Savage, J. E., Salvatore, J. E., Aliev, F., Edwards, A. C., Hickman, M., Kendler, K. S., & Dick, D. M. (2018). Polygenic Risk Score Prediction of Alcohol Dependence Symptoms Across Population-Based and Clinically Ascertained Samples. *Alcoholism, Clinical and Experimental Research*, 42(3), 520–530. <https://doi.org/10.1111/acer.13589>.
- Schaub, M. A., Boyle, A. P., Kundaje, A., Batzoglou, S., & Snyder, M. (2012). Linking disease associations with regulatory information in the human genome. *Genome Research*, 22(9), 1748–1759. <https://doi.org/10.1101/gr.136127.111>.
- Schreiweis, C., Bornschein, U., Burguière, E., Kerimoglu, C., Schreier, S., Dannemann, M., & Graybiel, A. M. (2014). Humanized Foxp2 accelerates learning by enhancing transitions from declarative to procedural performance. *Proceedings of the National Academy of Sciences*, 111(39), 14253–14258. <https://doi.org/10.1073/pnas.1414542111>.
- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). p-Curve and Effect Size: Correcting for Publication Bias Using Only Significant Results. *Perspectives on Psychological Science*, 9(6), 666–681. <https://doi.org/10.1177/1745691614553988>.
- Stahl, E. A., Breen, G., Forstner, A. J., McQuillin, A., Ripke, S., Trubetskoy, V., & Sklar, P. (2019). Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nature Genetics*, 51(5), 793–803. <https://doi.org/10.1038/s41588-019-0397-8>.
- Stergiakouli, E., Davey Smith, G., Martin, J., Skuse, D. H., Viechtbauer, W., Ring, S. M., & St Pourcain, B. (2017). Shared genetic influences between dimensional ASD and ADHD symptoms during child and adolescent development. *Molecular Autism*, 8(1), 18. <https://doi.org/10.1186/s13229-017-0131-2>.
- Stojanovski, S., Felsky, D., Viviano, J. D., Shahab, S., Bangali, R., Burton, C. L., & Wheeler, A. L. (2019). Polygenic Risk and Neural Substrates of Attention-Deficit/Hyperactivity Disorder Symptoms in Youths With a History of Mild Traumatic Brain Injury. *Biological Psychiatry*, 85(5), 408–416. <https://doi.org/10.1016/j.biopsych.2018.06.024>.
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., & Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43), 15545–15550. <https://doi.org/10.1073/pnas.0506580102>.
- Sudre, G., Frederick, J., Sharp, W., Ishii-Takahashi, A., Mangalmurti, A., Choudhury, S., & Shaw, P. (2020). Mapping associations between polygenic risks for childhood neuropsychiatric disorders, symptoms of attention deficit hyperactivity disorder, cognition, and the brain. *Molecular Psychiatry*, 25, 2482–2492. <https://doi.org.ezproxy.library.wisc.edu/10.1038/s41380-019-0350-3>.

- Sullivan, P. F. (2010). The Psychiatric GWAS Consortium: Big Science Comes to Psychiatry. *Neuron*, *68*(2), 182–186. <https://doi.org/10.1016/j.neuron.2010.10.003>.
- Taylor, M. J., Martin, J., Lu, Y., Brikell, I., Lundström, S., Larsson, H., & Lichtenstein, P. (2019). Association of Genetic Risk Factors for Psychiatric Disorders and Traits of These Disorders in a Swedish Population Twin Sample. *JAMA Psychiatry*, *76*(3), 280–289. <https://doi.org/10.1001/jamapsychiatry.2018.3652>.
- The ENCODE Project Consortium. (2011). A User's Guide to the Encyclopedia of DNA Elements (ENCODE). *PLOS Biology*, *9*(4), e1001046. <https://doi.org/10.1371/journal.pbio.1001046>.
- The International Schizophrenia Consortium. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, *460*(7256), 748–752. <https://doi.org/10.1038/nature08185>.
- Thomas, R., Sanders, S., Doust, J., Beller, E., & Glasziou, P. (2015). Prevalence of Attention-Deficit/Hyperactivity Disorder: A Systematic Review and Meta-analysis. *Pediatrics*, *135*(4), e994–e1001. <https://doi.org/10.1542/peds.2014-3482>.
- Torkamani, A., Wineinger, N. E., & Topol, E. J. (2018). The personal and clinical utility of polygenic risk scores. *Nature Reviews Genetics*, *19*(9), 581. <https://doi.org/10.1038/s41576-018-0018-x>.
- Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *The American Journal of Human Genetics*, *101*(1), 5–22. <https://doi.org/10.1016/j.ajhg.2017.06.005>.
- Volkow, N. D., Wang, G.-J., Fowler, J. S., Tomasi, D., & Telang, F. (2011). Addiction: Beyond dopamine reward circuitry. *Proceedings of the National Academy of Sciences*, *108*(37), 15037–15042. <https://doi.org/10.1073/pnas.1010654108>.
- Vuijk, P. J., Martin, J., Braaten, E. B., Genovese, G., Capawana, M. R., O'Keefe, S. M., & Doyle, A. E. (2019). Translating Discoveries in Attention-Deficit/Hyperactivity Disorder Genomics to an Outpatient Child and Adolescent Psychiatric Cohort. *Journal of the American Academy of Child & Adolescent Psychiatry*. <https://doi.org/10.1016/j.jaac.2019.08.004>.
- Walters, R. K., Polimanti, R., Johnson, E. C., McClintick, J. N., Adams, M. J., Adkins, A. E., & Agrawal, A. (2018). Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nature Neuroscience*, *21*(12), 1656–1669. <https://doi.org/10.1038/s41593-018-0275-1>.
- Wray, N. R., Goddard, M. E., & Visscher, P. M. (2007). Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Research*, *17*(10), 1520–1528. <https://doi.org/10.1101/gr.6665407>.
- Wray, N. R., Lee, S. H., Mehta, D., Vinkhuyzen, A. A. E., Dudbridge, F., & Middeldorp, C. M. (2014). Research Review: Polygenic methods and their application to psychiatric traits. *Journal of Child Psychology and Psychiatry*, *55*(10), 1068–1087. <https://doi.org/10.1111/jcpp.12295>.
- Zettler, P. J., Sherkow, J. S., & Greely, H. T. (2014). 23andMe, the Food and Drug Administration, and the Future of Genetic Testing. *JAMA Internal Medicine*, *174*(4), 493–494. <https://doi.org/10.1001/jamainternmed.2013.14706>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.