



# HHS Public Access

Author manuscript

*Biol Psychiatry*. Author manuscript; available in PMC 2019 August 16.

Published in final edited form as:

*Biol Psychiatry*. 2018 June 01; 83(11): e51–e53. doi:10.1016/j.biopsych.2018.04.007.

## Polygenic Risk Scores: What Are They Good For?

**Amanda B. Zheutlin,**

Center for Genomic Medicine, Massachusetts General Hospital, Boston, Massachusetts

**David A. Ross**

Department of Psychiatry, Yale University, New Haven, Connecticut.

Matt was in many ways the stereotypical “all-American” teenager: a multisport athlete, class president, and valedictorian of his high school. It seemed preordained that he would attend his dream college and continue this extraordinary trajectory. But something went wrong. In his first year of college, he struggled academically and became increasingly withdrawn. He sought treatment through the university’s health services. He was diagnosed with depression but engaged ambivalently with the clinic, with intermittent adherence with his medications and therapy appointments. His therapist worked hard to understand Matt’s resistance to treatment and to establish a therapeutic alliance. Matt ultimately disclosed—with a deep sense of shame—what he considered both his darkest secret and fear. When he was 6 years old, his father began behaving erratically, went missing for long periods of time; his father was committed against his will to a state psychiatric hospital, where he eventually died. Tearfully, Matt turns to you and asks, “What if the same thing is happening to me?”

“Madness” has been described since antiquity. However, the search for its biological basis has proven more elusive. Psychosis has always been described as genetic—at least since the concept of a “gene” was first put forth by Gregor Mendel. But it was not until 1959 that scientists were first able to measure genetic variation (Figure 1). These original tests could only identify large chromosomal abnormalities—but that was enough to discover the cause of conditions like Down syndrome. Patricia Jacobs, one of the pioneers of this method, went on to describe one of the first mutations ever to be linked to a psychiatric disorder: a translocation that appeared in all seven patients with schizophrenia in a Scottish pedigree in chromosome 1 (1). (The gene that holds the mutation was later named for it: disrupted in schizophrenia [*DISC1*].) But further investigation came up short; the mutation has never been found outside this pedigree, and many features of the initial discovery have been called into question (2).

Nonetheless, the promise of genomics for transforming medicine was powerful, and the search for schizophrenia risk genes continued. Collections of pedigrees amounting to hundreds of individuals were probed further, identifying several new mutations using linkage analysis. Investigations of more than 1000 candidate genes—all hypothesized to play some role in disease pathophysiology (e.g., *COMT*, which codes for an enzyme that breaks down

---

Address correspondence to Amanda B. Zheutlin, Ph.D., Massachusetts General Hospital, Center for Genomic Medicine, 185 Cambridge Street, CPZN-6, Boston, MA 02114; azheutlin@mgh.harvard.edu.

The authors report no biomedical financial interests or potential conflicts of interest.

prefrontal dopamine)—were carefully executed (3). It was decades of hard work. And yet by the early 2000s, only a handful of variants with weak associations had been described—and, even more discouragingly, most did not replicate across studies. Schizophrenia was clearly genetic—so where were all the genes?

The completion of the Human Genome Project in 2003 revolutionized the field. With 99.99% of the human genome sequenced, it became possible to easily scan the entire genome and to measure genetic variation never before assayed. As technologies became cheaper, faster, and higher resolution, genome-wide association studies (GWASs) that surveyed millions of commonly varying polymorphisms to compare cases versus controls quickly became standard practice. Science declared the study of individual human genetic variation the “breakthrough of the year” in 2007. New discoveries in breast cancer, myocardial infarction, diabetes, and many other diseases were reported hand over fist. Psychiatry was sure to be next.

In 2009, Purcell *et al.* (4) published the first GWAS of schizophrenia. The data were gathered from a consortium of researchers who pooled cohorts, collecting 6909 cases and controls—more individuals than any study in psychiatric genetics to date. It was clear from the onset that this would be a landmark study for the field. But the results were unexpected. Instead of uncovering a complete list of risk loci, they found that there were perhaps thousands of risk-conferring variants, most of which had a very small effect. And because the effects were so small, the sample size that would be needed to definitively establish which were real would need to be orders of magnitude larger—something completely unprecedented at the time.

The authors were not content to simply wait. The risky variants were there, even if it was not clear which were most important. Surely, the cumulative data from across thousands of possible variants could be used to predict an individual’s risk for an illness. And so Purcell *et al.* (4) offered the idea of a polygenic risk score (PRS)—a score reflecting the sum of all known risk alleles, weighted by how risky each variant was known to be. It was an elegant solution: a single value representing an individual’s overall genetic risk for schizophrenia.

As clever as the approach was, Purcell *et al.* (4) were still only able to capture 2.4% of the variation in case status with their initial sample. The next step for the field was to begin aggregating massively larger cohorts. Such an undertaking was rife with challenges: studies were expensive; funding was hard to secure (especially for such a nascent enterprise); the logistics of storing and processing the data demanded new technology; and researchers were initially resistant to a kind of collaboration that bucked long-standing institutional structures (e.g., giving up primary authorship). Ultimately, though, through hard-won academic collaborations, plummeting costs, and involvement from a booming direct-to-consumer genotyping industry, massive samples became a reality. A more recent GWAS of schizophrenia with nearly 150,000 participants more than doubled variance explained to 7%—the highest now in psychiatry (5).

The end game for PRSs, in some sense, has always been personalized medicine: a single blood test from which a clinician can estimate risk across an unlimited number of diseases

and traits. Especially when combined with other data, it might be possible to anticipate—and mitigate—risk, years or decades in advance. And while this may sound a bit far off (for almost everything, it is far off), it could soon be a reality for a few medical conditions.

For coronary heart disease, individuals within the top 2.5% of a PRS have the same increased risk (fourfold) as other commonly considered factors, such as familial hypercholesterolemia (a disease caused by a single genetic mutation) (6). If clinicians had access to this information, it would enable them to offer the same counseling and interventions that they do for patients meeting other high-risk criteria (6). Similarly, a breast cancer PRS can stratify women based on 10-year and lifetime risk (7). Based on United Kingdom screening guidelines, if implemented in clinical care ubiquitously, the PRS alone could identify women likely to account for 17% of total breast cancer cases in the population. This could allow the system to optimize screening strategies (e.g., via mammogram timing and frequency) to better align with individuals' risk. Translating these findings from large cohort studies into clinical care will be challenging, but the stage is now set for piloting clinical integration (6). In psychiatry, we are still at least several years from a similarly robust PRS, but such a score could facilitate mapping early disease course and the optimization of early intervention strategies.

Beyond risk stratification, PRSs have also pushed us to think more deeply about what psychiatric diseases are and how they are related to each other. An age-old question, historically we have defined psychiatric illnesses using lists of symptoms, most recently as codified in the DSM-5. It is possible that the addition of genetic information could help organize patients by the cause of their symptoms. This could allow for personalized treatment assignment and even new drug discovery based on key shared genetic substrates. As one example, based solely on symptomatology, depression and bipolar disorder seem highly similar (both involve a constellation of mood- and energy-related items) whereas schizophrenia seems categorically different (the primary symptoms are psychotic). Genetically, though, the picture is quite different: bipolar disorder and schizophrenia have a much higher degree of overlap in risk variants than do bipolar disorder and depression (68% vs. 47%). Fascinatingly, this same pattern is also reflected in gene expression across the brain (8). [And, of course, consistent with these findings, emerging evidence is demonstrating a broader role for the use of antipsychotics in bipolar disorder than there is for antidepressants (9).] This is one of the most exciting aspects of modern genetics— as ongoing research further elucidates the genetic landscape, we are discovering both shared biological substrates [such as calcium channel signaling (10)] and disease-specific genetic variation; together, these offer great potential to improve both nosology and clinical care in psychiatry.

So where does this leave us with Matt? The art of the clinical encounter is—and will always remain—in the ability to engage with patients through therapeutic communication. Much of that work today involves helping individuals cope with the intrinsic uncertainty of their situations. As our field continues to evolve, we hope that new research will chip away at this uncertainty and, critically, that improved understanding of the biological basis of disease will empower clinicians to identify and deliver the most effective interventions at the most appropriate times.

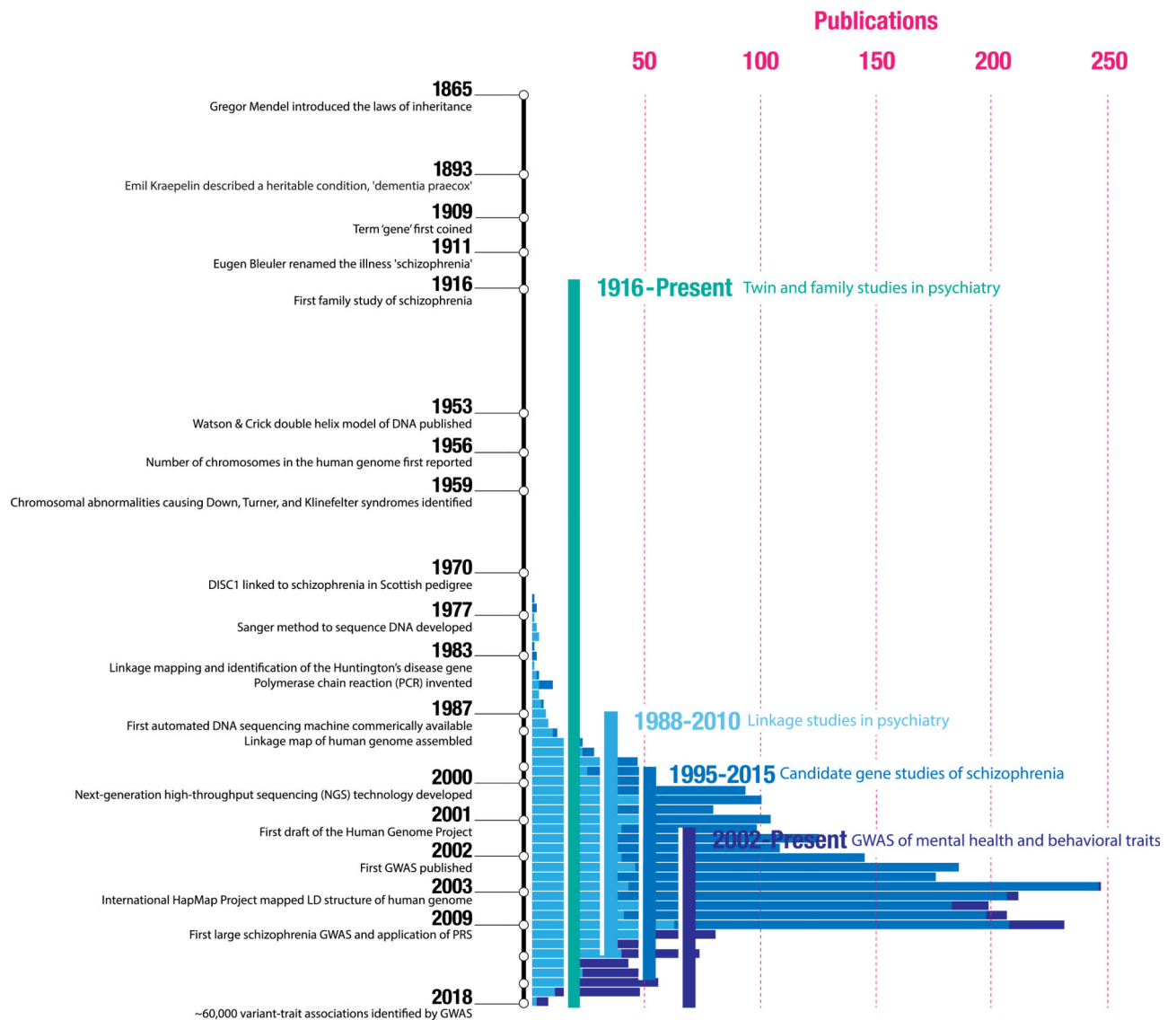
## Acknowledgments and Disclosures

Clinical Commentaries are produced in collaboration with the National Neuroscience Curriculum Initiative (NNCI). David Ross, in his dual roles as co-chair of the NNCI and as Education Editor of Biological Psychiatry, manages the development of these commentaries but plays no role in the decision to publish each commentary. The NNCI is supported by the National Institutes of Health Grant Nos. R25 MH10107602S1 and R25 MH08646607S1.

We thank Amanda Wang for her role in developing the figure.

## References

1. Jacobs P, Brunton M, Frackiewicz A, Newton M, Cook P, Robson E (1970): Studies on a family with three cytogenetic markers. *Ann Hum Genet* 33:325–336.
2. Sullivan PF (2013): Questions about DISC1 as a genetic risk factor for schizophrenia. *Mol Psychiatry* 18:1050–1052. [PubMed: 24056909]
3. Henriksen MG, Nordgaard J, Jansson LB (2017): Genetics of schizophrenia: Overview of methods, findings and limitations. *Front Hum Neurosci* 11:322. [PubMed: 28690503]
4. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, et al. (2009): Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 460:748–752. [PubMed: 19571811]
5. Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014): Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511:421–427. [PubMed: 25056061]
6. Khera AV, Chaffin M, Aragam K, Emdin CA, Klarin D, Haas M, et al. (2017): Genome-wide polygenic score to identify a monogenic riskequivalent for coronary disease [published online ahead of print Nov 15]. *bioRxiv*.
7. Mavaddat N, Pharoah PDP, Michailidou K, Tyrer J, Brook MN, Bolla MK, et al. (2015): Prediction of breast cancer risk based on profiling with common genetic variants. *J Natl Cancer Inst* 107:1–15.
8. Gandal MJ, Haney JR, Parikshak NN, Leppa V, Ramaswami G, Hartl C, et al. (2018): Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* 359:693–697. [PubMed: 29439242]
9. Goodwin GM, Haddad PM, Ferrier IN, Aronson JK, Barnes TRH, Cipriani A, et al. (2016): Evidence-based guidelines for treating bipolar disorder: Revised third edition recommendations from the British Association for Psychopharmacology. *J Psychopharmacol* 30: 495–553. [PubMed: 26979387]
10. Smoller JW, Craddock N, Kendler K, Lee PH, Neale BM, Nurnberger JI, et al. (2013): Identification of risk loci with shared effects on five major psychiatric disorders: A genome-wide analysis. *Lancet* 381: 1371–1379. [PubMed: 23453885]



**Figure 1.**

Selected events from the methodological evolution of genetic analysis and, specifically, for psychiatric genetics, are shown on the vertical timeline (not to scale). The x-axis also illustrates a broad estimate of the number of articles per year for linkage studies in psychiatry (derived from PubMed), candidate gene studies in schizophrenia (SZGene; updated through 2012), and mental health and behavioral genome-wide association studies (GWASs) (from the National Human Genome Research Institute–European Bioinformatics Institute GWAS catalog). While the notion that schizophrenia was heritable was confirmed in the first half of the 20th century by twin and family studies, it only became possible to link genetic variation to psychiatric outcomes in the last 50 years. Cytogenetic studies, which detected large chromosomal abnormalities, were the first to investigate the genetic underpinnings of schizophrenia. Linkage studies scanned for patterns of disease cosegregation in large pedigrees (i.e., regions where a genetic divergence within a family consistently tracked with the individuals that developed the illness) to nominate broad loci

that might harbor risk variants. Candidate gene studies moved genetic association outside of pedigrees to a case-control design. Here, the effect of a variant hypothesized to play a role in disease pathophysiology (e.g., variants within genes involved in dopaminergic signaling) on risk for an illness was assessed by comparing cases and controls. Finally, GWASs—a hypothesis-free approach where millions of variants were compared between cases and control subjects—became possible after the human genome was sequenced and catalogs of its structure were assembled (e.g., HapMap). Today, this is still the most popular approach to understanding how common genetic variation relates to disease. LD, linkage disequilibrium; PRS, polygenic risk score.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript