**The emerging pattern of shared polygenic architecture of psychiatric disorders,
conceptual and methodological challenges**

Olav B Smeland (M.D. Ph.D.)¹, Oleksandr Frei (Ph.D.)¹, Chun-Chieh Fan (M.D. Ph.D.)²,
Alexey Shadrin (Ph.D.)¹, Anders M Dale (Ph.D.)^{3,4,5}, Ole A Andreassen (M.D. Ph.D.)¹

¹NORMENT Centre, Institute of Clinical Medicine, University of Oslo and Division of Mental
Health and Addiction, Oslo University Hospital, 0407 Oslo, Norway;

²Center for Human Development, University of California, San Diego, La Jolla, CA 92093,
United States of America;

³Department of Radiology, University of California, San Diego, La Jolla, CA 92093, United
States of America;

⁴Department of Neuroscience, University of California San Diego, La Jolla, CA 92093, United
States of America;

⁵Center for Multimodal Imaging and Genetics, University of California San Diego, La Jolla,
CA 92093, United States of America;

Corresponding author: Ole A Andreassen

Ole A. Andreassen M.D. Ph.D.
Professor of Biological Psychiatry,
Division of Mental Health and Addiction
University of Oslo and Oslo University Hospital
Kirkeveien 166, 0424 Oslo, Norway
Email: o.a.andreassen@medisin.uio.no
Phone: +47 23027350

Manuscript word count: 3477

Abstract

Genome-wide association studies (GWAS) have transformed psychiatric genetics, and provided novel insights into the genetic etiology of psychiatric disorders. Two major discoveries have emerged: i) the disorders are polygenic, with a large number of common variants each with a small effect, and ii) many genetic variants influence more than one phenotype, suggesting shared genetic etiology across many pairs of traits and disorders. These concepts have the potential to revolutionize the current classification system with distinct diagnostic psychiatric categories, and facilitate development of better treatments. However, to reach clinical impact, we need larger samples and better analytical tools, as most polygenic factors and pleiotropic loci remain undetected. We here present novel statistical approaches that are specifically designed to improve the yield of existing GWAS for polygenic phenotypes. We review how these tools have informed the current knowledge on the genetic architecture of psychiatric disorders, focusing on schizophrenia, bipolar disorder and major depression, and overlap with psychological and cognitive traits. We further discuss application of novel statistical tools for stratification and prediction, that better consider the polygenic architecture of psychiatric phenotypes.

1   **Introduction**

2   Psychiatric disorders are recognized as leading causes of morbidity and are among the most
3   costly disorders to affect humans (DALYs and Collaborators, 2016). At the individual level
4   suffering is large, and the disorders are associated with impaired quality of life and low
5   socioeconomic status. Identifying the underlying pathophysiology for these disorders, as well
6   as resilience factors, is imperative and can lead to major health benefits through better treatment
7   regimens. Further, development of risk prediction in mental disorders could inform prevention
8   strategies and enrich clinical trials. While there has been a remarkable improvement in life
9   expectancy for the general population the last decades, there is a marked social inequality in
10   the field of mental disorders (Laursen et al., 2011). Patients and their families display
11   significantly higher mortality than the general population (Eaton et al., 2008, Ringen et al.,
12   2014), both from natural causes (somatic conditions where cardiovascular disease is most
13   important) and unnatural causes (suicide, homicide or accidents). Register-based studies
14   demonstrate that patients with mental illness have 15-20 years shorter life expectancy than the
15   general population (Wahlbeck et al., 2011). To reduce this gap, knowledge of underlying
16   disease causes and effective prevention strategies are urgently required (Insel, 2010).

17
18   Psychiatric disorders are regarded as complex disorders with heritability estimates between 40-
19   80% (Lichtenstein et al., 2009). While there are clear evidence for rare sequence variants and
20   copy-number variants with large effects associated with schizophrenia (Marshall et al., 2017)
21   and ADHD (Williams et al., 2010), large-scale genome-wide association studies (GWAS)
22   conducted during the last decade have shown that a moderate fraction of the heritability of most
23   psychiatric disorders is accounted for by numerous common genetic variants with small
24   “polygenic” effects (Schizophrenia Working Group of the Psychiatric Genomics Consortium,
25   2014, Demontis et al., 2019, Grove et al., 2019, Howard et al., 2019, Stahl et al., 2019). Due to

1 the revolution in genetics technology and the assembly of large genotyped samples, many
2 genetic variants have successfully been associated with severe psychiatric disorders in recent
3 years. Today, updates from the Psychiatric Genomics Consortium (PGC) include discoveries
4 of 30 genomic loci for bipolar disorder (Stahl et al., 2019), 102 for major depression (Howard
5 et al., 2019), five for autism spectrum disorder (Grove et al., 2019), 12 for ADHD (Demontis
6 et al., 2019), and approximately 250 for schizophrenia (Ripke et al. WCPG 2018). One
7 characteristic finding is the large degree of genetic overlap between mental disorders (Cross-
8 Disorder Group of the Psychiatric Genomics et al., 2013, Brainstorm Consortium, 2018), and
9 between mental disorders and related psychosocial traits (Lo et al., 2017, Day et al., 2018,
10 Savage et al., 2018, Jansen et al., 2019), which may indicate shared molecular genetic
11 mechanisms and possibly overlapping etiology. Yet, despite the assembly of very large GWAS
12 samples, often involving more than hundred thousand participants, most of the polygenic
13 architectures underlying psychiatric disorders still remain undetected (Schizophrenia Working
14 Group of the Psychiatric Genomics Consortium, 2014, Demontis et al., 2019, Grove et al., 2019,
15 Howard et al., 2019, Stahl et al., 2019). This can be attributed to the polygenic nature of
16 psychiatric disorders that poses considerable challenges on analytical methods and GWAS
17 sample size (Sullivan et al., 2018). In short, a GWAS allows for genome-wide analysis of
18 millions of common genetic variants (tag single-nucleotide polymorphisms [SNPs]), estimating
19 their effects on a given phenotype. Given the large numbers of SNPs tested, a GWAS must
20 correct for multiple testing using a stringent threshold of genome-wide significance (typically,
21 $p < 5 \times 10^{-8}$) to avoid false positives. Thus, only a subset of all involved genetic variants is
22 revealed, with a large fraction of the polygenic architecture remaining to be uncovered (i.e.,
23 “the missing heritability”) (Manolio et al., 2009). This has motivated efforts to develop “Big
24 Data” analytical approaches that improve the yield of existing GWAS. In particular,
25 mathematical models building on empirical Bayesian statistical approaches have emerged,

1 which are specifically designed to handle polygenic scenarios, resulting in substantially
2 improved power for genetic discovery (Andreassen et al., 2013b, Schork et al., 2016). Here we
3 review some of the recent discoveries of polygenic architecture in major psychiatric disorders
4 (schizophrenia, bipolar disorder, major depression) enabled by novel statistical tools, which has
5 revealed genetic overlap across psychiatric disorders, psychosocial traits and several somatic
6 traits and diseases. Moreover, we discuss how these tools may improve genetic prediction and
7 estimate discovery trajectories of future GWAS for psychiatric disorders. For example, while
8 the PGC now aims for 1 million genotyped participants for each mental disorder (Sullivan et
9 al., 2018), recent causal mixture modelling analysis (Frei et al., 2019) estimated that this will
10 explain ~60% of the SNP-heritability in schizophrenia and bipolar disorder, but only ~10% in
11 major depression (Figure 1).

13 **Genetic overlap between psychiatric disorders and traits**

14 The increasing wealth of GWAS data now available on human traits and disorders have shown
15 that a large number of genetic variants influence more than one phenotype (Visscher et al.,
16 2017), i.e. they exhibit allelic pleiotropy. This has profound implications for understanding the
17 underlying biology of complex phenotypes. The standard approach to evaluate the polygenic
18 relationship between two phenotypes today is to measure genetic correlation using tools such
19 as polygenic risk scores (Purcell et al., 2009) and linkage disequilibrium (LD) score regression
20 (Bulik-Sullivan et al., 2015). These tools have provided important insights into the shared
21 genetic etiology between human phenotypes, including mental disorders (Visscher et al., 2017,
22 Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium,
23 2018, Brainstorm Consortium, 2018). However, the methods do not provide a complete picture
24 of the complex genetic relationship between polygenic phenotypes. Similar to twin studies,
25 genetic correlations are unable to reveal the individual genetic variants shared between the

phenotypes, which is needed to identify the molecular genetic mechanisms involved. Further, the tools estimating genetic correlation can only detect genetic overlap when the effect directions are consistent (Purcell et al., 2009, Bulik-Sullivan et al., 2015). This is a clear limitation, as increasing evidence shows that overlapping genetic variants between several human phenotypes involve a mixed pattern of allelic effect directions (Baurecht et al., 2015b, Lee et al., 2016, Schmitt et al., 2016, Bansal et al., 2018, Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2018, Smeland et al., 2018, Frei et al., 2019, Smeland et al., 2019).

Cross-trait analytical approaches such as the conditional False Discovery Rate (condFDR) approach complements the standard measures of genetic correlation by allowing identification of individual overlapping variants that did not reach genome-wide significance, and by allowing identification of variants regardless of their allelic effect directions. The condFDR is a model-free strategy designed for polygenic phenotypes inspired by Empirical Bayes approaches (Efron, 2010). It leverages overlapping SNP associations (cross-trait enrichment) between two separate GWAS to improve statistical power for genetic discovery (Andreassen et al., 2013a, Schork et al., 2016). The conjunctive FDR (conjFDR) is a natural extension of the condFDR, which allows discovery of overlapping loci by providing a conservative estimate of the FDR for a SNP association with both phenotypes simultaneously (Andreassen et al., 2013a, Schork et al., 2016). Application of the condFDR and conjFDR approaches has increased genetic discovery and uncovered genetic overlap in a wide specter of complex human traits, including the psychiatric disorders schizophrenia (Andreassen et al., 2013b, Andreassen et al., 2014, Le Hellard et al., 2017, McLaughlin et al., 2017, Smeland et al., 2017b, Smeland et al., 2017a, Shadrin et al., 2018, Smeland et al., 2018, van der Meer et al., 2018, Zuber et al., 2018, Smeland

et al., 2019), bipolar disorder (Andreassen et al., 2013b, Andreassen et al., 2014, Drange et al., 2019, Smeland et al., 2019) and ADHD (Shadrin et al., 2018).

Notably, the conjFDR approach has demonstrated genetic overlap between several phenotypes that are not genetically correlated, such as schizophrenia and brain structure volumes (Smeland et al., 2018), schizophrenia and personality traits (Smeland et al., 2017a), and bipolar disorder and intelligence (Smeland et al., 2019). Moreover, it has helped elucidate the complexity of the genetic relationship between many complex phenotypes, for example that between schizophrenia and cognitive function. It is well established that schizophrenia is associated with cognitive impairment (Kahn and Keefe, 2013), and many genetic studies have demonstrated a negative genetic correlation between schizophrenia and various cognitive measures using tools such as polygenic risk scores (Lencz et al., 2014, Hubbard et al., 2016) and LD score regression, with genetic correlations ranging between -0.2 to -0.4 (Hagenaars et al., 2016, Hill et al., 2016, Liebers et al., 2016, Sniekers et al., 2017, Trampush et al., 2017, Brainstorm Consortium, 2018, Davies et al., 2018, Savage et al., 2018). Complementing these studies, a recent condFDR investigation analyzed large GWAS on schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014) and intelligence (Savage et al., 2018), and identified 75 shared loci at conjFDR<0.01 (Smeland et al., 2019). A gene-set enrichment analysis of the shared loci implicated biological processes related to neurodevelopment, synaptic integrity, and neurotransmission, among others. Among the shared loci, schizophrenia risk was linked to lower intelligence at 61 (81%) of the loci (Smeland et al., 2019). These findings corroborate a prior condFDR study on smaller GWAS samples on cognitive traits which found that schizophrenia risk was associated with poorer cognitive performance at 18 of 21 shared loci, where the implicated genes were globally expressed across the developing and adult human brain (Smeland et al., 2017b). Thus, in addition to identifying more loci shared between

1 schizophrenia and cognitive traits compared to the standard GWAS analysis, these conjFDR
2 studies indicate that the shared genetic etiology between schizophrenia and cognitive function
3 involves a mixture of agonistic and antagonistic effect directions, and is more complex than
4 what is suggested by their moderate negative genetic correlation (Hagenaars et al., 2016, Hill
5 et al., 2016, Liebers et al., 2016, Sniekers et al., 2017, Trampush et al., 2017, Brainstorm
6 Consortium, 2018, Davies et al., 2018, Savage et al., 2018). This is clinically important, and in
7 compliance with some reports that not all patients with schizophrenia perform poorly on
8 cognitive tests (Palmer et al., 1997).

9
10 Several methods for cross-trait GWAS analysis have been developed during the last decade,
11 which have been extensively reviewed elsewhere (Gratten and Visscher, 2016, Schork et al.,
12 2016, Hackinger and Zeggini, 2017, Pasaniuc and Price, 2017). Building on the meta-analysis
13 approach (Willer et al., 2010), many techniques aim to identify shared or unique genomic loci
14 across separate GWAS, including the COMBINE approach (Ellinghaus et al., 2012), restricted
15 and weighted subset search (ASSET) (Bhattacharjee et al., 2012), and compare-and-contrast
16 meta-analysis (CCMA) (Baurecht et al., 2015a). In contrast to such meta-analytical approaches,
17 condFDR analysis intrinsically incorporates multiple testing via the FDR framework by directly
18 working with the entire original set of p-values from two investigated GWAS (Efron, 2010).
19 Newer methods such as MTAG (Turley et al., 2018) or Genomic SEM (Grotzinger et al., 2019)
20 leverage genetic correlation between phenotypes to improve discovery of shared loci. This is a
21 powerful feature for highly correlated phenotypes, but not optimal for phenotypes with a low
22 or non-significant genetic correlation. Conversely, the condFDR method improves genetic
23 discovery by leveraging overlapping SNP associations regardless of the direction of their allelic
24 effects, and may boost discovery of loci jointly influencing phenotypes even in the absence of
25 genome-wide correlation, such as done for bipolar disorder and intelligence (Smeland et al.,

2019). Loci prioritized by standard GWAS analysis or other cross-trait analytical methods can be further interrogated with tools that aim to disentangle LD structure and uncover causal genetic mechanisms. For example, several available Bayesian approaches can explore whether two association signals in the same genomic region obtained from two different GWAS share a single causal variant or multiple causal variants (Giambartolomei et al., 2014, Pickrell et al., 2016).

Variations in polygenicity and heritability define ‘discoverability’

To provide further insights into the genetic relationship between complex human phenotypes, we have developed a statistical model that estimates the number of causal genetic variants influencing a given phenotype (which is termed “polygenicity”) (Holland et al., 2019) and the number of variants unique and shared between phenotypes (Frei et al., 2019). The mathematical models build on a mixture modeling framework (Thompson et al., 2015, Holland et al., 2016), in which only a fraction of causal variants in the genome are assumed to influence a given phenotype, while a null-component is assumed to have no effect on the phenotype. The mixture modelling framework is increasingly applied by novel statistical tools for analysis of complex polygenic phenotypes (Zeng et al., 2018, Zhang et al., 2018). Our model works with GWAS summary statistics, and incorporate detailed LD structures, disentangling their effects on the GWAS signals. Building on this approach, we have introduced the term *discoverability* (Fan et al., 2018). This is defined as the power to detect genetic variants for a given phenotype depending on its unique genomic architecture and GWAS sample size. Given a fixed GWAS sample size, the power to detect novel loci is determined by the effect size distribution of the causal loci. Correspondingly, a larger number of true causal loci (i.e., higher polygenicity) at a fixed heritability, will make SNP effects harder to detect, since they will be increasingly difficult to separate from the background signal (Fan et al., 2018). In addition to estimating

1 polygenicity, the models also estimate the narrow-sense heritability, and the proportion of
2 heritability captured by genome-wide significant SNPs (Frei et al., 2019, Holland et al., 2019).
3 The latter is a function of GWAS sample size, and enables power analysis of existing and future
4 GWAS (Holland et al., 2019). The univariate model thus explains why certain traits have lower
5 yield of genome-wide significant hits despite having larger GWAS sample size and higher
6 heritability (Holland et al., 2019) (Figure 1). For example, even though current GWAS sample
7 sizes are substantially larger for major depression (246,363 cases and 561,190 controls)
8 (Howard et al., 2019) than for schizophrenia (34,241 cases, 45,604 controls, and 1,235 parent-
9 affected offspring trios) (Schizophrenia Working Group of the Psychiatric Genomics
10 Consortium, 2014) and bipolar disorder (20,352 cases and 31,358 controls) (Stahl et al., 2019),
11 recent univariate analysis shows that while the proportion of identified variance is around 3%
12 in schizophrenia, and close to 1% in bipolar disorder, it is even lower in major depression
13 (Figure 1). This both reflects the larger number of variants estimated to influence major
14 depression (14.9k variants) compared to schizophrenia (8.3k variants) and bipolar disorder (6.4
15 variants), as well the lower SNP-heritability of major depression (0.08) compared to the other
16 two disorders (0.45 for schizophrenia, and 0.34 for bipolar disorder) (Holland et al., 2019).
17 Altogether, these parameters yield a lower *discoverability* for major depression variants, and
18 the model estimates that with 1 million GWAS participants, the expected genome-wide
19 significant loci will explain ~60% of SNP-heritability in schizophrenia and bipolar disorder,
20 but less than 10% for major depression (Figure 1). Hence, although the PGC now aims for 1
21 million genotyped participants for each mental disorder (Sullivan et al., 2018), this will
22 seemingly not be sufficient to completely uncover the common variant architecture for these
23 psychiatric disorders using standard statistical tools, in particular not for major depression. This
24 warrants phenotypic refinement to reduce disease heterogeneity or applying more cost-effective

1 statistical approaches to increase the yield of existing and future GWAS, for example by
2 leveraging overlapping genetic signal across traits and disorders to improve discovery.

3
4 The bivariate extension of the causal mixture model can estimate the extent of polygenic
5 overlap between complex phenotypes, allowing shared GWAS participants (Frei et al., 2019).
6 For example, it estimated that there is substantial polygenic overlap between schizophrenia and
7 educational attainment, which involves almost all causal variants for schizophrenia. However,
8 there is a mixture of agonistic and antagonistic effect directions among the shared variants,
9 yielding a low effect size correlation of 0.06 within the shared genomic fraction (Frei et al.,
10 2019). This is in line with the genome-wide correlation of 0.08 estimated between these
11 phenotypes using LD score regression (Okbay et al., 2016, Lee et al., 2018), and prior genetic
12 studies reporting mixed allelic effects among their overlapping genomic loci (Schizophrenia
13 Working Group of the Psychiatric Genomics Consortium, 2014, Okbay et al., 2016, Le Hellard
14 et al., 2017). Moreover, the bivariate model estimated a substantial polygenic overlap between
15 schizophrenia and bipolar disorder, which seems to involve almost all causal variants conferring
16 risk to bipolar disorder (Frei et al., 2019) (Figure 2). Interestingly, the model also estimated that
17 there are smaller fractions of causal variants that are specific to either schizophrenia and bipolar
18 disorder, which may offer important insights into the genetic differences between these
19 disorders. We also find extensive overlap between bipolar disorder and major depression, but
20 less overlap between schizophrenia and major depression (Figure 2). Overall, these data suggest
21 that in order to more completely understand the distinct genetic architecture underlying these
22 disorders, it is important to characterize both their disorder-specific effect size distributions
23 within the shared genomic fractions, as well as the disorder-specific non-overlapping fractions.
24 To this end, three-way causal mixture model analysis may help mapping out unique and

overlapping genetic mechanisms between groups of traits and disorders. This is a subject of our future research.

Improved prediction and clinical utility of polygenic statistical tools

Despite the significant advances in psychiatric genetics the last decade, there is still no utility for individual genetic prediction in clinical psychiatry to aid prevention, diagnostic accuracy and predict therapeutic response and disease course. In comparison, polygenic risk scores have reached promising predictive power for various somatic conditions, but the evidence for clinical use is still sparse (Torkamani et al., 2018, Abraham et al., 2019, Khera et al., 2019). Nevertheless, the discovery of genetic influences underlying mental traits and disorders may already inform psychiatric nosology, epidemiological associations, and provide insights into pathobiological underpinnings (Smoller et al., 2018). For example, the converging evidence that psychiatric disorders share a considerable proportion of genetic risk variants with each other (Cross-Disorder Group of the Psychiatric Genomics et al., 2013, Brainstorm Consortium, 2018, Frei et al., 2019), poses a challenge to the current diagnostic classification systems, in which psychiatric disorders are considered categorically distinct (Smoller et al., 2018). Additional data indicate that psychiatric disorders overlap genetically with a range of normal psychosocial traits such as cognition (Savage et al., 2018), personality (Lo et al., 2017), sleep patterns (Jansen et al., 2019) and social traits (Day et al., 2018). This indicates that most psychiatric disorders and psychosocial traits may exist on continua in genomic space, and are influenced by many overlapping genetic variants. Importantly, these results may support ongoing efforts to develop novel classification systems in which psychiatric disorders are considered continuous with normal variation in neurobiological and behavior dimensions (Cuthbert and Insel, 2010). Such a refinement of the psychiatric diagnostic system may help

1 establishing diagnostic categories that are more closely linked to distinct pathobiological
2 processes.

3
4 The frequently used liability threshold models in genetic testing algorithms are designed to be
5 insensitive of age (Falconer, 1965, Martin et al., 2018). Yet, most psychiatric disorders have
6 strong age-dependent clinical manifestations. To capture time-dependent pathological changes
7 and predict onset of brain diseases, we have developed the Polygenic Hazard Score (PHS)
8 (Desikan et al., 2017), which provides a framework for exploitation of polygenic information
9 towards clinical utility. In short, PHS models the time-dependent disease process by estimating
10 the risk of onset as a hazard function, incorporating genetic variants that influence the age-of-
11 disease-onset (Desikan et al., 2017). By profiling disease risk in the temporal domain, PHS can
12 quantify age-specific genetic risk for Alzheimer's disease and other complex diseases (Desikan
13 et al., 2017, Seibert et al., 2018), providing grounds for clinical prediction and disease risk
14 stratifications. We are currently working to revise and extend the PHS method by integrating
15 other approaches to improve prediction of psychiatric disorders, where an important feature
16 will be to include non-genetic data (Seibert et al., 2018). Although the genetic impact on
17 temporal pathophysiological processes may not be monotonically increased over time for
18 psychiatric disorders, it is of high importance to investigate whether there are polygenic effects
19 that may accelerate or delay disease mechanisms. The PHS algorithms may also aid clinical
20 trials as improved genetic risk stratification can help selecting groups of high-risk individuals
21 for study inclusion that are more likely to develop disease further on or respond to novel
22 therapeutic agents.

24 **Conclusion**

1 Increasing evidence has shown that psychiatric disorders are highly polygenic and that genetic
2 pleiotropy is pervasive among psychiatric disorders and related traits, providing important
3 biological insights into underlying mechanisms. While larger GWAS samples will increase the
4 number of disease-associated variants, recent analyses suggest that not even GWAS sample
5 sizes reaching 1 million participants will uncover most of the SNP-heritability for
6 schizophrenia, bipolar disorder and major depression. Hence, more efficient statistical tools,
7 that better take into account the distinct polygenic architecture underlying each disorder, may
8 help move the field forward. As more disease-associated variants for psychiatric disorders will
9 be uncovered, this may have a profound impact for understanding their underlying etiology and
10 provide novel biomarkers to increase diagnostic accuracy and prediction algorithms.

11

1

2 **Conflict of Interest Disclosures:** O.A.A. has received speaker's honorarium from Lundbeck

3 and is a consultant for Healthlytix. A.M.D is a co-founder of NeuroQuant and HealthLytix.

4 Remaining authors have no conflicts of interest to declare.

5

6 **Funding/Support:** NIH (NS057198; EB00790); NIH NIDA/NCI: U24DA041123; the

7 Research Council of Norway (229129; 213837; 248778; 223273; 249711); the South-East

8 Norway Regional Health Authority (2017-112); KG Jebsen Stiftelsen (SKGJ-2011-36).

References

- Abraham G, Malik R, Yonova-Doing E, Salim A, Wang T, Danesh J et al. (2019). Genomic risk score offers predictive performance comparable to clinical risk factors for ischaemic stroke. *bioRxiv*:689935. doi: 10.1101/689935.
- Andreassen OA, Djurovic S, Thompson WK, Schork AJ, Kendler KS, O'Donovan MC et al. (2013a). Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. *Am J Hum Genet* 92(2):197-209. doi: 10.1016/j.ajhg.2013.01.001.
- Andreassen OA, Harbo HF, Wang Y, Thompson WK, Schork AJ, Mattingdal M et al. (2014). Genetic pleiotropy between multiple sclerosis and schizophrenia but not bipolar disorder: differential involvement of immune-related gene loci. *Molecular psychiatry* 20(2):207-214. doi: 10.1038/mp.2013.195.
- Andreassen OA, Thompson WK, Schork AJ, Ripke S, Mattingdal M, Kelsoe JR et al. (2013b). Improved detection of common variants associated with schizophrenia and bipolar disorder using pleiotropy-informed conditional false discovery rate. *PLoS genetics* 9(4):e1003455. doi: 10.1371/journal.pgen.1003455.
- Bansal V, Mitjans M, Burik CAP, Linner RK, Okbay A, Rietveld CA et al. (2018). Genome-wide association study results for educational attainment aid in identifying genetic heterogeneity of schizophrenia. *Nat Commun* 9(1):3078. doi: 10.1038/s41467-018-05510-z.
- Baurecht H, Hotze M, Brand S, Buning C, Cormican P, Corvin A et al. (2015a). Genome-wide comparative analysis of atopic dermatitis and psoriasis gives insight into opposing genetic mechanisms. *Am J Hum Genet* 96(1):104-120. doi: 10.1016/j.ajhg.2014.12.004.
- Baurecht H, Hotze M, Brand S, Buning C, Cormican P, Corvin A et al. (2015b). Genome-wide comparative analysis of atopic dermatitis and psoriasis gives insight into opposing genetic mechanisms. *American journal of human genetics* 96(1):104-120. doi: 10.1016/j.ajhg.2014.12.004.
- Bhattacharjee S, Rajaraman P, Jacobs Kevin B, Wheeler William A, Melin Beatrice S, Hartge P et al. (2012). A Subset-Based Approach Improves Power and Interpretation for the Combined Analysis of Genetic Association Studies of Heterogeneous Traits. *The American Journal of Human Genetics* 90(5):821-835. doi: <https://doi.org/10.1016/j.ajhg.2012.03.015>.
- Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium (2018). Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell* 173(7):1705-1715 e1716. doi: 10.1016/j.cell.2018.05.046.
- Brainstorm Consortium (2018). Analysis of shared heritability in common disorders of the brain. *Science* 360(6395). doi: 10.1126/science.aap8757.
- Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR et al. (2015). An atlas of genetic correlations across human diseases and traits. *Nature genetics* 47(11):1236-1241. doi: 10.1038/ng.3406.
- Cross-Disorder Group of the Psychiatric Genomics C, Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM et al. (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature genetics* 45(9):984-994. doi: 10.1038/ng.2711.
- Cuthbert BN, Insel TR (2010). Toward New Approaches to Psychotic Disorders: The NIMH Research Domain Criteria Project. *Schizophrenia bulletin*. doi: sbq108 [pii] 10.1093/schbul/sbq108.
- DALYs GBD, Collaborators H (2016). Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 388(10053):1603-1658. doi: 10.1016/S0140-6736(16)31460-X.

Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD et al. (2018). Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat Commun* 9(1):2098. doi: 10.1038/s41467-018-04362-x.

Day FR, Ong KK, Perry JRB (2018). Elucidating the genetic basis of social interaction and isolation. *Nature communications* 9(1):2457. doi: 10.1038/s41467-018-04930-1.

Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E et al. (2019). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature genetics* 51(1):63-75. doi: 10.1038/s41588-018-0269-7.

Desikan RS, Fan CC, Wang Y, Schork AJ, Cabral HJ, Cupples LA et al. (2017). Genetic assessment of age-associated Alzheimer disease risk: Development and validation of a polygenic hazard score. *PLoS medicine* 14(3):e1002258. doi: 10.1371/journal.pmed.1002258.

Drange OK, Smeland OB, Shadrin AA, Finseth PI, Witoelar A, Frei O et al. (2019). Genetic Overlap Between Alzheimer's Disease and Bipolar Disorder Implicates the MARK2 and VAC14 Genes. *Front Neurosci* 13:220. doi: 10.3389/fnins.2019.00220.

Eaton WW, Martins SS, Nestadt G, Bienvenu OJ, Clarke D, Alexandre P (2008). The burden of mental disorders. *Epidemiologic reviews* (Research Support, N.I.H., Extramural Review) 30:1-14. doi: 10.1093/epirev/mxn011.

Efron B (2010). *Large-scale inference : empirical Bayes methods for estimation, testing, and prediction*. Cambridge ; New York: Cambridge University Press.

Ellinghaus D, Ellinghaus E, Nair RP, Stuart PE, Esko T, Metspalu A et al. (2012). Combined analysis of genome-wide association studies for Crohn disease and psoriasis identifies seven shared susceptibility loci. *American journal of human genetics* 90(4):636-647. doi: 10.1016/j.ajhg.2012.02.020.

Falconer DS (1965). The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Annals of Human Genetics* 29(51).

Fan CC, Smeland OB, Schork AJ, Chen CH, Holland D, Lo MT et al. (2018). Beyond heritability: improving discoverability in imaging genetics. *Human molecular genetics* 27(R1):R22-R28. doi: 10.1093/hmg/ddy082.

Frei O, Holland D, Smeland OB, Shadrin AA, Fan CC, Maeland S et al. (2019). Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. *Nat Commun* 10(1):2417. doi: 10.1038/s41467-019-10310-0.

Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C et al. (2014). Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS genetics* 10(5):e1004383. doi: 10.1371/journal.pgen.1004383.

Gratten J, Visscher PM (2016). Genetic pleiotropy in complex traits and diseases: implications for genomic medicine. *Genome medicine* 8(1):78. doi: 10.1186/s13073-016-0332-x.

Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD et al. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav* 3(5):513-525. doi: 10.1038/s41562-019-0566-x.

Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H et al. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature genetics* 51(3):431-444. doi: 10.1038/s41588-019-0344-8.

Hackinger S, Zeggini E (2017). Statistical methods to detect pleiotropy in human complex traits. *Open Biol* 7(11). doi: 10.1098/rsob.170125.

Hagenaars SP, Harris SE, Davies G, Hill WD, Liewald DC, Ritchie SJ et al. (2016). Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N=112 151) and 24 GWAS consortia. *Molecular psychiatry* 21(11):1624-1632. doi: 10.1038/mp.2015.225.

Hill WD, Davies G, Group CCW, Liewald DC, McIntosh AM, Deary IJ (2016). Age-Dependent Pleiotropy Between General Cognitive Function and Major Psychiatric Disorders. *Biological psychiatry* 80(4):266-273. doi: 10.1016/j.biopsych.2015.08.033.

Holland D, Frei O, Desikan R, Fan C-C, Shadrin AA, Smeland OB et al. (2019). Beyond SNP Heritability: Polygenicity and Discoverability of Phenotypes Estimated with a Univariate Gaussian Mixture Model. *bioRxiv*:133132. doi: 10.1101/133132.

Holland D, Wang Y, Thompson WK, Schork A, Chen CH, Lo MT et al. (2016). Estimating Effect Sizes and Expected Replication Probabilities from GWAS Summary Statistics. *Front Genet* 7:15. doi: 10.3389/fgene.2016.00015.

Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirihi M et al. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci* 22(3):343-352. doi: 10.1038/s41593-018-0326-7.

Hubbard L, Tansey KE, Rai D, Jones P, Ripke S, Chambert KD et al. (2016). Evidence of Common Genetic Overlap Between Schizophrenia and Cognition. *Schizophrenia bulletin* 42(3):832-842. doi: 10.1093/schbul/sbv168.

Insel TR (2010). Rethinking schizophrenia. *Nature (Historical Article)* 468(7321):187-193. doi: 10.1038/nature09552.

Jansen PR, Watanabe K, Stringer S, Skene N, Bryois J, Hammerschlag AR et al. (2019). Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat Genet* 51(3):394-403. doi: 10.1038/s41588-018-0333-3.

Kahn RS, Keefe RS (2013). Schizophrenia is a cognitive illness: time for a change in focus. *JAMA psychiatry* 70(10):1107-1112. doi: 10.1001/jamapsychiatry.2013.155.

Khera AV, Chaffin M, Wade KH, Zahid S, Brancale J, Xia R et al. (2019). Polygenic Prediction of Weight and Obesity Trajectories from Birth to Adulthood. *Cell* 177(3):587-596 e589. doi: 10.1016/j.cell.2019.03.028.

Laursen TM, Munk-Olsen T, Gasse C (2011). Chronic somatic comorbidity and excess mortality due to natural causes in persons with schizophrenia or bipolar affective disorder. *PloS one* 6(9):e24597. doi: 10.1371/journal.pone.0024597.

Le Hellard S, Wang Y, Witoelar A, Zuber V, Bettella F, Hugdahl K et al. (2017). Identification of Gene Loci That Overlap Between Schizophrenia and Educational Attainment. *Schizophrenia bulletin* 43(3):654-664. doi: 10.1093/schbul/sbw085.

Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M et al. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature genetics* 50(8):1112-1121. doi: 10.1038/s41588-018-0147-3.

Lee PH, Baker JT, Holmes AJ, Jahanshad N, Ge T, Jung JY et al. (2016). Partitioning heritability analysis reveals a shared genetic basis of brain anatomy and schizophrenia. *Molecular psychiatry* 21(12):1680-1689. doi: 10.1038/mp.2016.164.

Lencz T, Knowles E, Davies G, Guha S, Liewald DC, Starr JM et al. (2014). Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: a report from the Cognitive Genomics consortium (COGENT). *Molecular psychiatry* 19(2):168-174. doi: 10.1038/mp.2013.166.

Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, Sullivan PF et al. (2009). Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* 373(9659):234-239. doi: 10.1016/S0140-6736(09)60072-6.

Liebers DT, Pirooznia M, Seiffudin F, Musliner KL, Zandi PP, Goes FS (2016). Polygenic Risk of Schizophrenia and Cognition in a Population-Based Survey of Older Adults. *Schizophrenia bulletin* 42(4):984-991. doi: 10.1093/schbul/sbw001.

Lo MT, Hinds DA, Tung JY, Franz C, Fan CC, Wang Y et al. (2017). Genome-wide analyses for personality traits identify six genomic loci and show correlations with psychiatric disorders. *Nature genetics* 49(1):152-156. doi: 10.1038/ng.3736.

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ et al. (2009). Finding the missing heritability of complex diseases. *Nature* (Review) 461(7265):747-753. doi: 10.1038/nature08494.

Marshall CR, Howrigan DP, Merico D, Thiruvahindrapuram B, Wu W, Greer DS et al. (2017). Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nature genetics* 49(1):27-35. doi: 10.1038/ng.3725.

Martin AR, Daly MJ, Robinson EB, Hyman SE, Neale BM (2018). Predicting Polygenic Risk of Psychiatric Disorders. *Biological psychiatry*. doi: 10.1016/j.biopsych.2018.12.015.

McLaughlin RL, Schijven D, van Rheenen W, van Eijk KR, O'Brien M, Kahn RS et al. (2017). Genetic correlation between amyotrophic lateral sclerosis and schizophrenia. *Nat Commun* 8:14774. doi: 10.1038/ncomms14774.

Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA et al. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 533(7604):539-542. doi: 10.1038/nature17671.

Palmer BW, Heaton RK, Paulsen JS, Kuck J, Braff D, Harris MJ et al. (1997). Is it possible to be schizophrenic yet neuropsychologically normal? *Neuropsychology* 11(3):437-446.

Pasaniuc B, Price AL (2017). Dissecting the genetics of complex traits using summary association statistics. *Nature reviews Genetics* 18(2):117-127. doi: 10.1038/nrg.2016.142.

Pickrell JK, Berisa T, Liu JZ, Segurel L, Tung JY, Hinds DA (2016). Detection and interpretation of shared genetic influences on 42 human traits. *Nature genetics* 48(7):709-717. doi: 10.1038/ng.3570.

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(7256):748-752. doi: 10.1038/nature08185.

Ringen PA, Engh JA, Birkenaes AB, Dieset I, Andreassen OA (2014). Increased mortality in schizophrenia due to cardiovascular disease - a non-systematic review of epidemiology, possible causes, and interventions. *Frontiers in psychiatry / Frontiers Research Foundation* 5:137. doi: 10.3389/fpsy.2014.00137.

Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, de Leeuw CA et al. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature genetics* 50(7):912-919. doi: 10.1038/s41588-018-0152-6.

Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511(7510):421-427. doi: 10.1038/nature13595.

Schmitt J, Schwarz K, Baurecht H, Hotze M, Folster-Holst R, Rodriguez E et al. (2016). Atopic dermatitis is associated with an increased risk for rheumatoid arthritis and inflammatory bowel disease, and a decreased risk for type 1 diabetes. *J Allergy Clin Immunol* 137(1):130-136. doi: 10.1016/j.jaci.2015.06.029.

Schork AJ, Wang Y, Thompson WK, Dale AM, Andreassen OA (2016). New statistical approaches exploit the polygenic architecture of schizophrenia--implications for the underlying neurobiology. *Curr Opin Neurobiol* 36:89-98. doi: 10.1016/j.conb.2015.10.008.

Seibert TM, Fan CC, Wang Y, Zuber V, Karunamuni R, Parsons JK et al. (2018). Polygenic hazard score to guide screening for aggressive prostate cancer: development and validation in large scale cohorts. *BMJ* 360:j5757. doi: 10.1136/bmj.j5757.

Shadrin AA, Smeland OB, Zayats T, Schork AJ, Frei O, Bettella F et al. (2018). Novel Loci Associated With Attention-Deficit/Hyperactivity Disorder Are Revealed by Leveraging

Polygenic Overlap With Educational Attainment. *J Am Acad Child Adolesc Psychiatry* 57(2):86-95. doi: 10.1016/j.jaac.2017.11.013.

Smeland OB, Wang Y, Lo MT, Li W, Frei O, Witoelar A et al. (2017a). Identification of genetic loci shared between schizophrenia and the Big Five personality traits. *Sci Rep* 7(1):2222. doi: 10.1038/s41598-017-02346-3.

Smeland OB, Wang Y, Frei O, Li W, Hibar DP, Franke B et al. (2018). Genetic Overlap Between Schizophrenia and Volumes of Hippocampus, Putamen, and Intracranial Volume Indicates Shared Molecular Genetic Mechanisms. *Schizophrenia bulletin* 44(4):854-864. doi: 10.1093/schbul/sbx148.

Smeland OB, Bahrami S, Frei O, Shadrin A, O'Connell K, Savage J et al. (2019). Genome-wide analysis reveals extensive genetic overlap between schizophrenia, bipolar disorder, and intelligence. *Molecular psychiatry*. doi: 10.1038/s41380-018-0332-x.

Smeland OB, Frei O, Kauppi K, Hill WD, Li W, Wang Y et al. (2017b). Identification of Genetic Loci Jointly Influencing Schizophrenia Risk and the Cognitive Traits of Verbal-Numerical Reasoning, Reaction Time, and General Cognitive Function. *JAMA psychiatry* (Original Investigation) 74(10):1065-1075. doi: 10.1001/jamapsychiatry.2017.1986.

Smoller JW, Andreassen OA, Edenberg HJ, Faraone SV, Glatt SJ, Kendler KS (2018). Psychiatric genetics and the structure of psychopathology. *Molecular psychiatry*. doi: 10.1038/s41380-017-0010-4.

Sniekers S, Stringer S, Watanabe K, Jansen PR, Coleman JRI, Krapohl E et al. (2017). Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nature genetics* 49(7):1107-1112. doi: 10.1038/ng.3869.

Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V et al. (2019). Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nature genetics* 51(5):793-803. doi: 10.1038/s41588-019-0397-8.

Sullivan PF, Agrawal A, Bulik CM, Andreassen OA, Borglum AD, Breen G et al. (2018). Psychiatric Genomics: An Update and an Agenda. *Am J Psychiatry* 175(1):15-27. doi: 10.1176/appi.ajp.2017.17030283.

Thompson WK, Wang Y, Schork AJ, Witoelar A, Zuber V, Xu S et al. (2015). An Empirical Bayes Mixture Model for Effect Size Distributions in Genome-Wide Association Studies. *PLoS genetics* 11(12):e1005717. doi: 10.1371/journal.pgen.1005717.

Torkamani A, Wineinger NE, Topol EJ (2018). The personal and clinical utility of polygenic risk scores. *Nature reviews Genetics* 19(9):581-590. doi: 10.1038/s41576-018-0018-x.

Trampush JW, Yang ML, Yu J, Knowles E, Davies G, Liewald DC et al. (2017). GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium. *Molecular psychiatry* 22(3):336-345. doi: 10.1038/mp.2016.244.

Turley P, Walters RK, Maghzian O, Okbay A, Lee JJ, Fontana MA et al. (2018). Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat Genet* 50(2):229-237. doi: 10.1038/s41588-017-0009-4.

van der Meer D, Rokicki J, Kaufmann T, Cordova-Palomera A, Moberget T, Alnaes D et al. (2018). Brain scans from 21,297 individuals reveal the genetic architecture of hippocampal subfield volumes. *Molecular psychiatry*. doi: 10.1038/s41380-018-0262-7.

Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA et al. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *American journal of human genetics* 101(1):5-22. doi: 10.1016/j.ajhg.2017.06.005.

Wahlbeck K, Westman J, Nordentoft M, Gissler M, Laursen TM (2011). Outcomes of Nordic mental health systems: life expectancy of patients with mental disorders. *Br J Psychiatry* 199(6):453-458. doi: 10.1192/bjp.bp.110.085100.

1 Willer CJ, Li Y, Abecasis GR (2010). METAL: fast and efficient meta-analysis of
2 genomewide association scans. *Bioinformatics* 26(17):2190-2191. doi:
3 10.1093/bioinformatics/btq340.
4 Williams NM, Zaharieva I, Martin A, Langley K, Mantripragada K, Fossdal R et al. (2010).
5 Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: a
6 genome-wide analysis. *Lancet* 376(9750):1401-1408. doi: 10.1016/S0140-6736(10)61109-9.
7 Zeng J, de Vlaming R, Wu Y, Robinson MR, Lloyd-Jones LR, Yengo L et al. (2018).
8 Signatures of negative selection in the genetic architecture of human complex traits. *Nature*
9 *genetics* 50(5):746-753. doi: 10.1038/s41588-018-0101-4.
10 Zhang Y, Qi G, Park JH, Chatterjee N (2018). Estimation of complex effect-size distributions
11 using summary-level statistics from genome-wide association studies across 32 complex
12 traits. *Nature genetics* 50(9):1318-1326. doi: 10.1038/s41588-018-0193-x.
13 Zuber V, Jonsson EG, Frei O, Witoelar A, Thompson WK, Schork AJ et al. (2018).
14 Identification of shared genetic variants between schizophrenia and lung cancer. *Sci Rep*
15 8(1):674. doi: 10.1038/s41598-017-16481-4.

Figures legends

Figure 1. Power plots for schizophrenia (SCZ, blue), bipolar disorder (BIP, orange), major depression (MD, green), educational attainment (EDU, red) and height (purple) estimated using the causal mixture model (Holland et al., 2019). The plots were originally presented in the paper by Holland et al. (2019). Proportion of SNP-heritability, captured by genome-wide significant SNPs, projected to current and future GWAS sample sizes, N. Values for current GWAS sample sizes are shown in parentheses.

Figure 2. Venn diagram of unique and shared polygenic components at the causal level, showing polygenic overlap (grey) between schizophrenia (SCZ, blue), bipolar disorder (BIP, orange), and major depression (MD); the numbers indicate the estimated quantity of causal variants (in thousands) per component, explaining 90% of SNP heritability in each phenotype, followed by the standard error. The size of the circle reflects the degree of polygenicity. The diagrams were generated using the bivariate causal mixture model (Frei et al., 2019), and were originally presented in the paper by Frei et al. (2019).

1 **Figure 2.**

2