Rigor and Reproducibility in Genetic Research on Eating Disorders

Running title: Genetic Research on Eating Disorders

Christopher Hübel, M.D.[1,2,3], Virpi Leppä Ph.D.[3], Gerome Breen Ph.D.[1,2], Cynthia M. Bulik Ph.D.[3,4,5]

[1]Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK

[2]UK National Institute for Health Research (NIHR) Biomedical Research Centre for Mental Health, South London and Maudsley Hospital, London, UK

[3]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

[4]Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC

[5]Department of Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, NC

Correspondence to: Dr. Bulik, Department of Psychiatry, University of North Carolina at Chapel Hill, CB #7160, 101 Manning Drive, Chapel Hill, NC 27599-7160, USA; e-mail: cbulik@med.unc.edu

Acknowledgements

Dr. Bulik acknowledges funding from the Swedish Research Council (VR Dnr: 538-2013-8864). CM Bulik reports: Shire (grant recipient, Scientific Advisory Board member) and Pearson and Walker (author, royalty recipient). This study represents independent research part funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. Dr. Breen has received grant funding from and served as a consultant to Eli Lilly and has received honoraria from Illumina. All interests unrelated to this work. Dr. Leppä and Dr. Hübel have nothing to disclose. We gratefully thank Jonathan Coleman for the consulting during the writing of this manuscript.

Abstract

Objective: We explored both within-method and between-method rigor and reproducibility in the field of eating disorders genetics.

Method: We present critical evaluation and commentary on component methods of genetic research (family studies, twin studies, molecular genetic studies) and discuss both successful and unsuccessful efforts in the field.

Results: Eating disorders genetics has had a number of robust results that converge across component methodologies. Familial aggregation of eating disorders, twin-based heritability estimates of eating disorders, and genome-wide association studies (GWAS) all point toward a substantial role for genetics in eating disorders etiology and support the premise that genes do not act alone. Candidate gene and linkage studies have been less informative historically.

Discussion: The eating disorders field has entered the GWAS era with studies of anorexia nervosa. Continued growth of sample sizes is essential for rigorous discovery of actionable variation. Molecular genetic studies of bulimia nervosa, binge-eating disorder, and other eating disorders are virtually nonexistent and lag seriously behind other major psychiatric disorders. Expanded efforts are necessary to reveal the fundamental biology of eating disorders, inform clinical practice, and deliver new therapeutic targets.

Keywords: anorexia nervosa, bulimia nervosa, binge-eating disorder, GWAS, polygenic risk score, genetic correlation, pathway

1. Introduction

The goal of genetic research in eating disorders is to decipher basic mechanisms that contribute to or modulate risk of developing these complex phenotypes. Eating disorders, like all

psychiatric illnesses, are complex traits that that are influenced by hundreds or even thousands of variants of small to moderate effect (i.e., polygenicity) together with environmental factors (Gelernter, 2015; Plomin, Haworth, & Davis, 2009).

Genetic findings inform a wide array of functional investigations from genetically modified cells to animal models, and can potentially guide new drug development or repurposing of existing medications (Breen et al., 2016; Yilmaz, Hardaway, & Bulik, 2015). Health science research for eating disorders is underfunded relative to their contribution to global disability adjusted life-years (Insel, 2015; Murray, Pila, Griffiths, & Le Grange, 2017). Consequently, the use of resources must be optimized to guarantee the generation of reproducible and robust findings. Genetic research offers myriad opportunities to understand eating disorders on molecular, cellular, and systems biology levels and is envisioned to improve risk prediction, therapeutic interventions, and preventive strategies (Sullivan et al., 2018).

DNA is the underlying blueprint of the organism and interindividual changes in this blueprint are partially responsible for the observed differences between individuals. Our genome is assigned at conception and remains essentially stable throughout life. This stability means that it is prone to far fewer confounding factors than are necessary to attend to in other fields, including epidemiology and epigenetics. Specifically, genetics is robust to reverse causation, and, with epidemiologically informed statistical analyses, allows for causal interpretation of findings. For genomics to be optimally informative, it is important to achieve the broadest possible coverage of the genome.

There are tens of millions of common genetic variants (i.e., present in at least 1% of the population), of many different types, throughout the genome. The most common variants are single base pair changes, such as adenine (A) to thymine (T), which are referred to as single nucleotide polymorphisms or SNPs. Current technology based on DNA microarrays allows

direct measurement ("genotyping") of SNPs at between 500,000 to 5 million positions. Thereafter, imputation methods are used in conjunction with reference panels from whole genome sequencing (1000 Genomes Project Consortium et al., 2012; International HapMap Consortium, 2005; McCarthy et al., 2016) to boost the information content in the data by allowing the prediction of genotypes at positions not measured on the arrays, thereby increasing coverage of the genome and improving power to detect associations.

Eating disorders genetics to date has yielded many robust, rigorous, and reproducible results, both within and across component methodologies (see Table 1). In this review, we will focus on family, twin, and molecular genetic approaches highlighting convergent results and briefly review the history of candidate gene studies in eating disorders. First, we introduce the concept of heritability, which is central to understanding genetic studies of eating disorders. We recommend returning to this section after reading the entire paper to consolidate your understanding of the concept. Second, we briefly introduce the methodologies covered in this review: family, twin, linkage, candidate gene, genome-wide association studies (GWAS), their follow-on approaches, whole genome and exome sequencing, and epigenetics. Finally, we review four critical concepts necessary to evaluate the rigor of eating disorders genetics literature in an informed manner: phenotype, relatedness, population stratification, and sample size. Of note, comprehensive reviews of the findings of genetic studies in eating disorders are available, see Brandys et. al (2015), Breithaupt, Hübel, & Bulik (2018), Davis (2015), Hinney and Volckmar (2013), Yilmaz et al. (2015).

2. Heritability (h^2)

In general, heritability (h^2) refers to the proportion of phenotypic variance due to inherited (non-environmental) factors (see Box 1).

Box 1: Narrow-sense vs. broad-sense heritability

The reader should be mindful of the difference between h^2 (narrow-sense heritability) which captures the proportion of variance attributed to variation in additive genetic values and H^2 (broad-sense heritability), which captures the proportion of variance attributed to variation in total genetic values (Visscher, Hill, & Wray, 2008).

Numerically, h^2 estimates range from 0 (i.e., no effect of genes on phenotypic variation) to 1.0 (genes are the only contributor to phenotypic variation). The h^2 estimates for eating disorders can be derived from family studies (Javaras et al., 2008; Strober, Lampert, Morrell, Burroughs, & Jacobs, 1990), twin studies (Baker, Mitchell, Neale, & Kendler, 2010; Bulik et al., 2006, 2010; Bulik, Sullivan, & Kendler, 1998; Dellava, Thornton, Lichtenstein, Pedersen, & Bulik, 2011; Javaras et al., 2008; Klump, Miller, Keel, McGue, & Iacono, 2001; Kortegaard, Hoerder, Joergensen, Gillberg, & Kyvik, 2001; Mazzeo et al., 2009, 2010; Mitchell et al., 2010; Trace et al., 2013; Wade, Bulik, Neale, & Kendler, 2000), adoption studies (Klump, Suisman, Burt, McGue, & Iacono, 2009), and genomic studies (Duncan et al., 2017). The extent to which h^2 estimates across various methods converge is one measure of reproducibility in eating disorder genetics. As described below, differences in h^2 estimates across approaches should be expected and indeed are observed (reasons for this are covered in greater detail in section SNP-based h^2; Table 1). Critical to recognize is that h^2 is a population statistic (not an individual measure) and refers to a given population at a given point in time. The occurrence and frequencies of genetic variants can differ across populations for several reasons, including chance, migration, mutation, and selection (Wray & Visscher, 2008). Therefore, applying results from one population to another and developing predictors, such as polygenic risk scores, require well-matched

population samples. Sampling bias should also be considered: ascertainment approaches and inclusion and exclusion criteria can limit the representativeness of the sample and the applicability of the results to the general population. For example, gender, ethnicity, type of disorder, age, geography, and socioeconomic status should all be considered when designing ascertainment strategies.

2.1 Twin-based heritability (Twin-h^2)

Twin-based h^2 estimates are derived from the phenotypic similarity (i.e., phenotypic correlations) between monozygotic (identical twins) and dizygotic twins (fraternal twins), their known proportion of shared genomes (100% and 50%, respectively), and their expected shared environment (Rijsdijk & Sham, 2002). The standard ACE twin model (Figure 1) captures the relative contribution of additive genetic (A), common environmental (C), and unique environmental (E) factors on trait variation. Importantly, in standard twin models, these are not measured, but rather inferred. The proportion of variance of a trait accounted for by A, C, and E is equal to the square of the connecting paths (a, c, and e, respectively). In this model, a^2 represents the h^2 estimate, c^2 captures the variance of the shared environmental factors (all factors that serve to make twins similar), and e^2 captures the variance of all factors that serve to make twins different and includes error of measurement (Figure 1).

Twin-based h^2 (twin-h^2 or a^2) estimates of most behavioral traits (including eating disorders) fall roughly between 30% and 80% (Polderman et al., 2015). Twin-h^2 estimates are influenced by several factors including: a) how the trait under investigation is measured, b) trait prevalence, c) the population, d) sex, e) developmental stage or age, f) time of measurement, and g) geographical area (Visscher et al., 2008). Therefore, h^2 estimates for the same trait do and are expected to differ across studies, across populations, and across time.

2.2 Single nucleotide polymorphism-based heritability (SNP-h^2)

The h^2 contributed by common genetic variation can be calculated from genotype data from SNP arrays or GWAS summary statistics (SNP-h^2; Yang, Zeng, Goddard, Wray, & Visscher, 2017). Modern methods for estimating SNP-h^2 include genome-wide complex trait analysis (GCTA; Yang, Lee, Goddard, & Visscher, 2011) and linkage disequilibrium (LD) score regression (LDSC; Bulik-Sullivan et al., 2015). GCTA derives estimates by comparing the degree to which individuals share their genomes to their similarity on a particular trait. LD score regression estimates h^2 from regressing the association statistics for SNPs against their summed correlation with nearby SNPs (for further details, see Yang et al., 2017).

As mentioned previously, estimates of SNP-h^2 are typically and expectedly lower than twin-h^2 estimates. This has held true in twin studies in eating disorders and other psychiatric disorders (Table 1; Anttila et al., 2016; Cross-Disorder Group of the Psychiatric Genomics Consortium et al., 2013). The SNP-h^2 for anorexia nervosa (AN) is ~20% (Duncan et al., 2017) in comparison to twin-h^2 estimates ranging from 28% to 74% (Bulik et al., 2006, 2010; Dellava et al., 2011; Klump et al., 2001; Kortegaard et al., 2001; Mazzeo et al., 2009; Wade et al., 2000). This differential is consistent with other complex traits and is likely due to fundamental differences in what each measure captures. Whereas SNP-h^2 captures h^2 due primarily to common variation, twin-h^2 also captures h^2 due to rare genetic variations (mutations) as well as passive, evocative, and active gene-environment correlations (McAdams et al., 2014), and gene-environment interactions (Vinkhuyzen & Wray, 2015). It is important to note, however, that although h^2 estimates are a central outcome of twin studies, the goals of GWAS, in contrast, reach far beyond h^2 estimates and rather seek to identify actual causal variants that affect disease risk.

3. Core Methods in the Genetics of Eating Disorders

3.1 Family history and family studies

The familial aggregation of AN, bulimia nervosa (BN), and binge-eating disorder (BED) have been widely documented and replicated (Hudson et al., 2006; Javaras et al., 2008; Kassett et al., 1989; Lilenfeld et al., 1998; Mangweth et al., 2003; Stein et al., 1999; Strober, Freeman, Lampert, Diamond, & Kaye, 2000, 2001; Strober et al., 1990; Strober, Morrell, Burroughs, Salkin, & Jacobs, 1985). In their most basic form, family studies are designed to determine whether a trait or disorder "runs in families". The "family history" method in which one or a few family members act as informants about the status of other family members is less effortful but also less informative than the family study method (Hudson, Pope, Jonas, & Yurgelun-Todd, 1983; Hudson, Pope, Jonas, Yurgelun-Todd, & Frankenburg, 1987; Keck et al., 1990; Rivinus et al., 1984). In the "family study" method, every family member is queried or interviewed about their own history (and can also serve as an informant on other family members). Although methods do exist to determine the extent to which genes affect the familial aggregation of a trait or illness from a family study (Hudson et al., 2006; Javaras et al., 2008; Kassett et al., 1989; Lilenfeld et al., 1998; Mangweth et al., 2003; Stein et al., 1999; Strober et al., 2000, 2001, 1990, 1985), they are more effective in identifying the existence of familial aggregation rather than the extent to which genetic or environmental factors cause that aggregation (Visscher et al., 2008).

Population level family studies compare the overall population prevalence with the disease risk of family members with an index case in their family (a proband) and yield relative risk estimates taking into account the relationship to the proband (McGuffin & Burke, 2014). For instance, siblings of individuals with AN were four times more likely to suffer from AN than

siblings of unaffected control probands in a Danish nationwide registry study (Steinhausen, Jakobsen, Helenius, Munk-Jørgensen, & Strober, 2015).

More sophisticated family designs can also be conducted in large populations when different degrees of relatedness in multigenerational registers are used to tease apart genetic and environmental factors whilst controlling for genetic confounding of these environmental factors (i.e., gene-environment correlations). For example, a child may grow up in an environment where both parents smoke: the parents will most probably have passed on variants of genes which make the child more likely to smoke but may also create a family environment where the child is exposed to smoking behavior and second hand smoking (i.e., passive gene-environment correlation). The child may actively choose its environment where its peers smoke (i.e., active gene-environment correlation) and then may ask for cigarettes and receive them (i.e., evocative gene-environment correlation; Narusyte et al., 2008). In this design, multigenerational registers that can specify different family relationships between individuals who share known and varying proportions of genetic factors and environmental factors (Figure 2). Comparing the degree of genetic relatedness and similarities of the trait or disorder over several generations—both expressed in correlations—allow the researcher to make cautious causal inferences and estimate the relative contribution of genetic and environmental factors (D'Onofrio, Lahey, Turkheimer, & Lichtenstein, 2013).

3.2 Twin studies

After verifying familial aggregation of a trait or disorder in family studies, the next logical step is to determine the extent to which the observed familiality is due to genetic or environmental factors. Twin studies are an approach to address this question and were the mainstay of genetic epidemiology for decades (Figure 1). Encouragingly, estimates of the genetic

and environmental contributions to eating disorders represent robust findings in our field (Polderman et al., 2015). Twin studies conducted on patients with eating disorders in European-ancestry populations yielded h^2 point estimates ranging from 28% to 74% for AN, from 55% to 65% for BN, and from 33% to 45% for BED (Baker et al., 2010; Bulik et al., 2006, 2010, 1998; Dellava et al., 2011; Javaras et al., 2008; Klump et al., 2001; Kortegaard et al., 2001; Mazzeo et al., 2009, 2010; Mitchell et al., 2010; Trace et al., 2013; Wade et al., 2000). Due to the low prevalence of the disorders and thus relatively few cases in twin registries, most estimates have fairly wide confidence intervals (Yilmaz et al., 2015).

The validity of twin results rest on several important assumptions. First, twin studies assume that identical (monozygotic or MZ) twins are 100% genetically identical, whereas fraternal (dizygotic or DZ) twins share on average 50% of their segregating alleles (i.e., version of genes inherited from their parents; Rijsdijk & Sham, 2002). As our ability to query the genome advances, it is becoming clear that minor genetic and epigenetic differences do exist between MZ twins (Bruder et al., 2008; Charney, 2012; Wong et al., 2010); however, the impact of these differences on results is minimal (Liu, Molenaar, & Neiderhiser, 2017). Second, twin studies assume that environmental factors of relevance to the trait under investigation are shared equally between MZ and DZ twins (i.e., the equal environments assumption or EEA). Two large studies tested the EEA by, first, correcting twin-h^2 estimates by a measure of similarity between twins (Felson, 2014) and, second, by investigating the similarity between twins who had been falsely assigned as DZ when they were actually biologically MZ (Conley, Rauscher, Dawes, Magnusson, & Siegal, 2013). Neither approach revealed a gross violation of the EEA. Third, twin studies assume that the prevalence of trait under investigation is similar in twins as in the general population. For example, one study reported a 1.5-fold greater prevalence of AN in twins than in singletons in Sweden (Goodman, Heshmati, Malki, & Koupil, 2014). Although this

observation requires replication, it highlights the importance of comparing prevalences in twins relative to the general population to ensure that twin-specific risk factors are not operative.

3.3 Candidate gene studies

Candidate gene studies should be viewed in their historical context, which was limited by the fact that large scale genotyping arrays that concurrently measure several million genetic markers only became available and affordable in the last decade. In the candidate gene approach, a researcher chooses one or a few genes (i.e., extremely low coverage of the genome) that they hypothesize could contribute to eating disorders and compares the frequency of the variants of that gene in a sample of cases versus a sample of controls. Due to the many shortcomings of the candidate gene studies, they have fallen out of favor as an approach to the initial identification of genes involved in the etiology of psychiatric disorders (Duncan & Keller, 2011; Vinkhuyzen & Wray, 2015). Candidate gene studies are limited by pre-existing knowledge about the trait under study and about the selected genes, are physically limited to loci within and near genes (e.g., promoter, coding regions, and introns), and have typically been grossly underpowered. In addition, the same subtle biases that can affect genome-wide studies such as cryptic relatedness and population stratification can affect candidate gene studies.

In eating disorders, the selection of candidate genes was based on educated guesses about the biology of eating disorders (e.g., genes related to serotonin which is known to influence both appetite and mood). Knowing now that complex traits such as eating disorders are likely to be influenced by hundreds if not thousands of genes, it is understandable that these complex behavioral syndromes could not be explained via the function of one or a few genes. Many of the samples collected during the candidate gene era have been meaningfully carried forward to genome-wide studies. Current standards dictate that candidate gene studies should only be

employed to replicate findings from existing GWAS and never as an initial step in genetic analyses. No candidate gene study of eating disorders has ever produced robust and reproducible findings (Yilmaz et al., 2015).

3.4 Linkage studies

The fundamental premise of linkage studies is based on the observation that the genetic variants in genes that reside physically close on a chromosome remain linked during meiosis (i.e., specialized type of cell division to create germ cells) and tend to do so over several generations if those variants are close together. Linkage analysis allows the assessment of the strength of linkage between genetic variants characterizing a region and disorder or disorder status in families. In particular, this methodology is favored for the detection of large effect genetic mutations, especially those that have sufficient penetrance to cause observable dominant or recessive inheritance of disorders. We now know that such examples are rare in psychiatric disorders, [although they do exist, such as the chromosome 1:11 translocation in schizophrenia (St Clair et al., 1990)]. However, linkage studies conducted on AN and BN examining affected relative pairs have by and large not yielded robust or reproducible findings (Bacanu et al., 2005; Bergen et al., 2003; Bulik et al., 2003; Devlin et al., 2002; Grice et al., 2002).

3.5 Genome-wide association studies (GWAS)

GWAS is currently the workhorse of psychiatric genetics, and, as the name implies, is a genome-wide approach. Unlike candidate gene studies, a GWAS employs millions of markers across the genome and compares their frequency between cases and controls for binary traits, or their correlation with a trait, such as BMI, in a population. Sample sizes are usually large—in the tens or even hundreds of thousands. Given the sheer number of comparisons, most of the results

that cross the nominal association threshold of 0.05 do so by chance—so rigorous control for multiple testing must be employed. There are approximately 10^{-8} independent tests involved in a GWAS. Consequently, the Bonferroni multiple testing corrected p-value of $5x10^{-8}$ is regarded as genome-wide significant, which is crucial for rigor in GWAS (Dudbridge & Gusnanto, 2008; Risch & Merikangas, 1996). Given this stringent threshold, many true associations remain nonsignificant often due to practical limitations in power, which is dependent on sample size and allele frequencies of variants of interest in the sample.

Very large sample sizes are required to achieve adequate statistical power for the detection of variants within the expected effect size range for complex disorders, usually between odds ratios (OR) of 1.0 and 1.2 in case-control studies (Park et al., 2011; W. Y. S. Wang, Barratt, Clayton, & Todd, 2005). Although researchers used to larger effect sizes may find these to be small, they must be interpreted in the context of the polygenic model of complex traits in which thousands of genes in concert influence a phenotype. Interpreting GWAS effect sizes in the tradition of Cohen inevitably leads to false conclusions, results in misinterpretation, and premature rejection of results (Chen, Cohen, & Chen, 2010). Sample size is the critical ingredient for GWAS success and can be estimated using publicly available power calculators (Purcell, Cherny, & Sham, 2003; Sham & Purcell, 2014). The Psychiatric Genomics Consortium (PGC) has facilitated achieving required sample sizes in the tens to hundreds of thousands by facilitating unprecedented international collaborations (Sullivan et al., 2018). Research in schizophrenia and major depressive disorder has revealed a nearly linear relationship between the number of detected genome-wide significant regions (i.e., associations) and increases in sample size (Hyman, 2014). Current goals of the PGC are to achieve sample sizes of 100,000 for each major psychiatric disorder, including eating disorders.

Such collaboration also enabled the identification of the first genome-wide significant locus for AN by the Eating Disorders Working Group of the PGC. The locus on chromosome 12 has been previously implicated in a range of autoimmune disorders, such as rheumatoid arthritis and type 1 diabetes (Duncan et al., 2017). Until now, no GWAS of BN or BED have been conducted, but sample collection is underway.

Other factors that are important in evaluating GWAS include the screening of controls for the disorder or trait under study, especially if the disorders are common in the population, as well as for strongly correlated phenotypes. The rationale for this is a control group that contains a large proportion of cases decreases the statistical power to detect the association of a genetic variant with the disorder. Ancestry, relatedness, population stratification, and reliable phenotype assessment are crucial in GWAS (see section 4) and should be carefully considered in the study design. Extensive quality control is applied to genotype data: data sets are checked for genotyping quality (i.e., SNP and individual missingness); expected frequencies of the alleles in the population (i.e., deviation from Hardy-Weinberg equilibrium); and, if imputed, for imputation quality (i.e., INFO score; for detailed information see Coleman et al., 2016). Finally, replication of identified associations between genetic variants and a disorder should be pursued in an independent data set, which is often difficult due to the large sample sizes required for GWAS, but will become easier with the emergence of large biobanks.

The interpretation of GWAS results is complicated as common variant association signals can come from loci well outside of any known genes and from loci that are hard to annotate functionally or understand their precise role in the disorder. If a genomic region associated with a trait harbors genes, it may well be more than 30 genes, making the selection of which to carry forward in animal or organoid models challenging (Sekar et al., 2016). Estimates show that it is possible to create a hypothesis from about 80% of the associated and annotated genes in GWAS

underscoring the problem of gene selection (Hawrylycz et al., 2012). Additionally, identified genomic regions can interact with genes that are located further away on the same or other chromosomes. Genetic variants located within the alpha-ketoglutarate dependent dioxygenase (*FTO*) gene, for instance, are robustly associated with type 2 diabetes and obesity; however, these genetic variants interact with iroquois homeobox 3 (*IRX3*) which is several hundred thousand base pairs away from *FTO* due to complicated folding structures of the human chromosomes. *IRX3* appears to be the gene most strongly affecting BMI (Claussnitzer et al., 2015; Smemo et al., 2014). Genetic variants that are very far away from implicated genes—not just neighboring genes—can have surprisingly strong effects on the observed phenotype.

3.6 Follow-on approaches

A common question raised after loci are identified via GWAS is: "What next?" This is a complex question worthy of its own paper, and indeed several authors have addressed the question extensively (de Leeuw, Neale, Heskes, & Posthuma, 2016; Maier, Visscher, Robinson, & Wray, 2017; Yang et al., 2017). GWAS results can be prioritized and carried forward into systematic analyses including polygenic risk scoring (PRS), SNP-based genetic correlations, and pathway analyses in order to estimate genetic liabilities, identify genetic overlap between traits, understand extended biological pathways, and inform studies of functional biology, such as animal, cell, and organoid studies. These follow-on approaches combine many SNPs of small effect to capture the polygenic nature of complex traits. For our purposes in addressing rigor and reproducibility in eating disorders genetics, we will focus on a few widely employed methods. Reproducibility in animal (Peers, South, Ceuppens, Bright, & Pilling, 2014) and cell studies (González, Boué, & Izpisúa Belmonte, 2011; Maherali & Hochedlinger, 2008; Young-Pearse & Morrow, 2016) are reviewed elsewhere.

3.7 Polygenicity: polygenic risk scoring

The results generated by GWAS (i.e., summary statistics) contain a plethora of information. In addition to significant loci, genomic regions below the genome-wide significance threshold of $5x10^{-8}$ also carry valuable information. To capture the full polygenic signal, one can calculate PRS with software such as PLINK (Chang et al., 2015) or PRSice (Euesden, Lewis, & O'Reilly, 2015). A PRS is the sum of the genetic risk alleles carried by an individual, multiplied/weighted by their effect size. The effect sizes are determined from the largest available discovery GWAS. For instance, assume that A represents the risk allele at a given locus. As all individuals carry two alleles at the same locus, they can either carry two risk alleles (AA), one risk allele (A and T), or no risk alleles (TT). The risk allele "count" in the locus is then multiplied by the effect size (i.e., odds ratio or beta) and all risk loci are summed to form the total PRS value. The higher the PRS, the more risk variants an individual carries. For a practical guide see (Maier et al., 2017). PRS can be incorporated into regression analyses as predictors of, for example, case status, treatment outcome, or disorder severity (Figure 3).

PRS prediction studies require independent discovery and target samples. Summary statistics from a GWAS for a trait or disorder in the discovery sample are used to create the PRS, which is then applied to individuals in the target sample. In a regression model, we can then test how well the PRS derived from the discovery sample predicts case status in the target sample.

Rigorous application of PRS should ensure that the discovery and target samples are truly independent, with the same individuals or close relatives not represented in both samples. In addition, genetic variants in close physical proximity are not independent and therefore must be clumped to a single marker (Dudbridge, 2013, 2016; Wray et al., 2013, 2014). Inclusion of correlated markers would lead to overfitting of the model and inflate the association between

PRS and trait. It is also important to use two samples stemming from the same ancestry group and to correct for population stratification within these groups (see below). Furthermore, it is important to determine the genome-wide significance threshold that should be used to decide which genetic markers to include in the PRS. This can be achieved by fitting regression models at different genome-wide significance thresholds. This method, however, is subject to multiple testing because fitting a regression at every p value threshold represents a statistical test. This can be solved by permutation (Phipson & Smyth, 2010).

PRS need to be interpreted on the population level, meaning the absolute value of an individual's PRS needs to be compared with the PRS of the rest of the population. The distribution of PRS, therefore, is often divided into deciles. Furthermore, PRS should not be interpreted deterministically. Even though an individual may carry a substantial number of genetic risk variants (i.e., have a high PRS), they may never develop a disorder due to either protective environmental factors, absence of exposure to environmental triggers, or other genetic factors that could contribute to resilience. The opposite can also occur: an individual with a low PRS may develop a disorder because of adverse environmental exposures or other genetic factors that influence risk.

Other applications of PRS exist. PRS can be incorporated into imaging genetics to investigate the relationship between an individual's polygenic load and brain structure (Dima & Breen, 2015). PRS can also be used to investigate gene x environment interactions. For example, Agerbo et al. (2015) discovered a positive G x E interaction for schizophrenia where a higher schizophrenia PRS interacted with a positive family history increasing liability for the psychiatric disorder. However, this finding requires replication in a larger independent sample. Furthermore, genome-wide by environment interaction studies (GWEIS) can measure gene x environment interactions. For example, an interaction between stressful life events and genetic

variants associated with depressive symptoms has been reported (Dunn et al., 2016) and another larger investigation in the UK Biobank revealed an interaction between self-reported traumatic experiences and genetic variants associated with depression (Coleman, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, UK Biobank Mental Health Consortium, Eley, & Breen, 2018). However, both findings also require replication.

3.8 Pleiotropy: genetic correlations

Pleiotropy (i.e., the extent to which genetic variants influence two or more traits) is common in genetics and may explain the frequent co-occurrence of specific traits and disorders (i.e., comorbidity). The calculation of SNP-based genetic correlations (r_g) from GWAS summary statistics is one way to determine the extent to which disorders or traits share causal genetic variants. Like other correlations, SNP-based r_g have values between 0 and 1 and can be either positive or negative. A positive r_g indicates that the same genetic variants influence both traits in the same direction, and a negative r_g indicates that the same genetic variants influence both traits, but in opposite directions. For example, significant r_g from the latest AN GWAS, revealed a positive r_g between AN and schizophrenia and between AN and a cross disorder phenotype from the PGC, both confirming that AN is genetically related to other psychiatric disorders (Duncan et al., 2017). Intriguingly, a palette of significant positive genetic correlations emerged with favorable metabolic parameters (high density lipoprotein measures), and a parallel palette of negative genetic correlations with what are considered to be unfavorable metabolic parameters (insulin and glucose-related parameters). In addition, AN was negatively genetically correlated with BMI and obesity measures. Together these observations provide insight to the manner in which shared genetic variants influence AN and other psychiatric and metabolic/anthropometric

traits and underscore the importance of addressing metabolic as well as psychiatric aspects of the illness.

3.9 Biology: pathway analyses

Methods developed in systems biology allow us to discern functional pathways that are activated in eating disorders. Genes rarely function entirely independently, but rather are embedded in a highly complex network of pathways that interact with each other leading to downstream effects. Molecular components that are parts of reactions, such as nucleic acids and proteins, build networks of biological interactions and are grouped into pathways within those databases. Examples of these biological pathways are innate immune function, apoptosis, and neurotransmitter signaling.

Pathway analysis reduces the number of tests conducted by grouping genes into pathways; however, correction for multiple comparisons remains necessary. An extensive review on pathway analyses can be found by Kao et al. (2017). Although the current AN GWAS was underpowered to detect significant pathways associated with the disorder, larger sample sizes will lead to more robust signals and will facilitate future pathway analyses (Duncan et al., 2017). Pathway analysis results have the potential to reveal tissues that are relevant to the disorders, can inform downstream cell and animal studies that are designed to further explicate the biology of the illness or aspects of the illness, and can inform pharmacologic targets for interventions.

3.10. Drug repurposing and development

If evidence emerges for a specific drug-related pathway from a genome-wide study, the drug is more likely to show efficacy in the treatment of the disorder under study (Nelson et al., 2015). GWASs have already demonstrated the ability to identify and confirm previously known drug

targets, such as the dopamine receptor D2 for schizophrenia, the common target of antipsychotic medications (Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium, 2011). The genetic signal from GWAS, therefore, can be used as part of a drug discovery or drug repurposing pipeline to prioritize drug targets and their binding partners (Gaspar & Breen, 2017).

4. Factors to Consider When Evaluating Rigor and Reproducibility of the Genetics of Eating Disorders

4.1 Phenotype

A phenotype is the set or a subset of the observable characteristics of an individual. The validity and reliability of the phenotype studied is an important factor to consider in evaluating the genetics literature. In eating disorders genetics, many different phenotypes have been employed. In the absence of biomarkers or laboratory values that reliably index the presence or absence of an eating disorder, the most valued phenotype in research has been a diagnosis conducted via a structured clinical interview, by a trained clinician. Even though the structured clinical interview was the benchmark against which all other phenotypes were measured, it is still subjective, not perfectly reliable, and a potential source of error in genetic studies of eating disorders. This approach is also impractical and expensive for large studies. Moreover, diagnostic flux across eating disorder presentations is well-documented and reduces the reliability of eating disorders diagnoses (Castellini et al., 2011; Eddy et al., 2008; Ekeroth, Clinton, Norring, & Birgegård, 2013; Milos, Spindler, Schnyder, & Fairburn, 2005).

Other methods that capture diagnostic phenotypes include clinical records, national patient registers, self-report of ever having had an eating disorder, parent report, and self-report of symptoms that are combined through an algorithm into a diagnosis. Each of these has

advantages and disadvantages that should be considered carefully. The measurement of a phenotype should be also considered when evaluating reproducibility, as different ways to define an eating disorder can measure slightly different characteristics or phenotypes, which in turn affects the variants we discover.

With statistical power always being a critical consideration, and the prevalence of eating disorders being fairly low, diagnostic categories have often been relaxed in order to increase the number of available cases. This boost in sample size, however, could theoretically hamper gene discovery for the intended phenotype, as we might sample a more heterogeneous sample. For example, the twin-h^2 of AN decreases as diagnostic criteria broaden (Dellava et al., 2011).

Even more common is the use of self-report questionnaires that capture dimensions purported to be relevant to the etiology of eating disorders such as body dissatisfaction, dietary restraint, drive for thinness, etc. In reality, we have no idea whether these constructs index the same underlying biology/genetics that the core diagnoses of AN, BN, and BED do. Yet we often assume that we are speaking about highly related constructs. Some of the methods outlined above such as computing SNP-based r_g between these continuous phenotypes and eating disorders diagnoses will allow us to address this question directly.

Other approaches to consider include incorporating continuous dimensions from the Research Domain Criteria (RDoC) model to measure domains relevant to eating disorders such as reward, inhibition, and social communication (Wildes & Marcus, 2015) and incorporating objective measures including, for example, biomarkers, digital phenotyping, and accelerometer measurements to improve measurement precision and reduce error.

4.2 Relatedness

If the researcher is conducting a genetic case-control association study comparing the genomes of cases to the genomes of controls, and several related individuals are in the case sample, the sample might be biased enough to cause trouble (see Box 2).

Box 2: Potential Impact of Relatedness

Imagine that John hears about a genetic study of AN and given his history of the illness, decides to participate by donating a blood sample for DNA. At a holiday gathering, he tells his three siblings and all 21 of his cousins about the study. Although none of them has talked openly about their eating disorder, seven of them have had AN, and they all go home, login, and sign up for the study.

The fact that related individuals share large proportions of their segregating alleles (i.e., genomes) on average (e.g., identical twins 100%; full siblings on average 50%; first cousins on average 12.5%) increases the likelihood that related individuals carry the same genetic alteration at a certain genomic locus (Figure 2; Astle & Balding, 2009). If a sample includes cryptic relatedness, any observed genetic differences between cases and controls could actually be due to relatedness rather than the eating disorder phenotype under investigation. Control for this can be achieved calculating a relatedness matrix from a genome-wide marker data that estimates the degree of relatedness between individuals in the data set. These estimates can be included in the statistical model, which controls for their genomic similarity and enables the researcher to retain the participants (Hellwege et al., 2017). Another approach is to exclude related individuals (i.e., individuals who share more than 0.03%, such as second cousins or the grandchild of a cousin) from all analyses (Coleman et al., 2016).

4.3 Population stratification and leveraging diverse populations

Population stratification refers to systematic differences in allele frequency at certain positions throughout the genome in groups of individuals that are ancestrally from different geographical locations. When an undetected or unaccounted for systematic bias in allele frequencies between populations coincides with systematic difference in a phenotype under investigation, it can cause false positive or negative results (Box 3; Hellwege et al., 2017).

Box 3: Example of Population Stratification

One classic example of an association caused by population stratification is that of the genetic variant encoding increased lactase persistence into adulthood with height. Lactase persistence refers to the continued activity of the lactase enzyme which is responsible for digesting lactose in milk into adulthood. In most mammals (including humans), the activity of lactase is dramatically reduced after weaning. In humans this results in lactose intolerance. Epidemiologically, both height and lactase persistence increase towards Northern Europe, although there is no causal relationship between the two traits. Therefore, if one analyses people of European ancestry without taking into account their region of origin, the lactase persistence variant and height appear to be associated. If one properly matches cases and controls by their ancestry, e.g., Greeks with Greeks and Scandinavians with Scandinavians, the association disappears.

Most researchers are aware of the threat of population stratification when comparing broader ancestral groups (e.g., Africans, Asians, Europeans, Native Americans, etc.), but population

stratification can even obscure results within the same continent or country, if ancestry is not properly accounted for (Marchini, Cardon, Phillips, & Donnelly, 2004).

Population stratification reflects drift (i.e., random chance of alleles being dropped out or being enriched in a population), bottlenecks (i.e., a small ancestral population at any time point in the population's history), and non-random mating (i.e., you are more likely to mate with someone who lives near you and is ancestrally similar to you; Price, Zaitlen, Reich, & Patterson, 2010). Even in this highly mobile world, the pattern is remarkably visible in our genomes. An all-European investigation by Novembre et. al. (2008) basically recreated the geographical map of Europe by projecting genetic similarity between study participants onto a two-dimensional space (Figure 4).

Most studies now use principal components to correct for ancestry effects by including ancestry principal components as covariates (Hellwege et al., 2017; Pritchard, Stephens, & Donnelly, 2000; Y. Zhang, Shen, & Pan, 2013). Although large ancestrally homogeneous samples are powerful, including individuals from diverse ancestral backgrounds can increase the power of a genetic study (Pulit, Voight, & de Bakker, 2010). Statistical simulations comparing European-only with diverse ancestry GWAS showed that multiethnic samples increase power especially for the detection of variants with frequencies lower than 5% (Pulit et al., 2010). Thus, admixed populations may be another powerful and underused resource for genetic studies (J. Zhang & Stram, 2014).

4.4 Sample size

Large sample sizes are clearly essential to GWAS success, but the question of how big is big enough is highly relevant to eating disorder genetic rigor and reproducibility. For each

psychiatric disorder, genomic discovery has accelerated as sample size has increased. The current goal of the PGC is to achieve sample sizes of 100,000 for each of the ten disorders under study. However, the optimal sample size depends to a large extent on the genetic architecture of the illness under study. Figure 5, presents the discovery trajectory that can be expected in complex traits such as eating disorders (adapted from Sullivan et al., 2018). This curve illustrates the "dead zone" when sample sizes are inadequate for discovery of significant hits, through sample-size-induced acceleration of discovery, to the ultimate asymptote. The PGC recommends striving for a "good enough" point where "most genes are identified at least once and the majority of genes in salient biological processes are highlighted" (Sullivan et al., 2018). With this in mind, it is clear that statistical power in the initial two AN GWAS (Boraska et al., 2014; K. Wang et al., 2011) occupied the dead zone. Combining these efforts and adding new samples has edged eating disorders genetics into the discovery phase; however, many more samples are needed in order to reach the "good enough" point.

A point worth considering is that an individual's genome has one extraordinary advantage, namely stability over time. A given individual has to be sampled and genotyped only once and the resultant data can then be used repeatedly to address a host of research questions. For example, if an individual was a regular heavy smoker and had bulimia nervosa, their genotype data could contribute to both a smoking and a bulimia nervosa GWAS. Given the need for large samples for genomic discovery, mega-analytic and meta-analytic approaches are common and appropriate methods to combine GWAS data from large international consortia. Such approaches are even more essential in the search for rare variants (Evangelou & Ioannidis, 2013; Lin & Zeng, 2010; Sung et al., 2014). Heterogeneity, however, can be introduced from various sources, including phenotype definition, the inclusion of individuals from diverse ancestral groups, and different genotyping platforms or imputation software. Well designed

meta- and mega-analyses take these issues into account and address them appropriately (Evangelou & Ioannidis, 2013).

5. Rare variation and copy number variants

The genetic architecture of a trait comprises all genetic variants, their effect sizes, frequencies, interactions with each other, and with environmental factors through epigenetic factors (Timpson, Greenwood, Soranzo, Lawson, & Richards, 2018). Rare variants (i.e., present in less than 1% of the population) can also contribute to etiology and can be investigated via whole-genome (WGS) or whole-exome sequencing (WES). In schizophrenia, this approach identified genetic variants that were not captured by GWAS and are assumed to be etiologically relevant (Richards et al., 2016). WGS, however, is still a costly (around $2,000 per genome) and therefore not yet affordable for large-scale epidemiological studies (Schwarze, Buchanan, Taylor, & Wordsworth, 2018; Timpson et al., 2018). If we assume that rare variants are present in about 1% of the population, in an optimal scenario, we would need to sample 1,000 cases to identify 10 carriers, which should be compared to at least 1,000 controls. This may require even larger sample sizes than GWAS. Copy number variants (CNVs) are large-scale structural changes of the genome that span more than 1 kb (i.e., 1000 base pairs), including insertions, deletions, and duplications. Most CNVs are inherited, but some arise *de novo* (Thapar & Cooper, 2013). Carriers of large, rare, pathogenic CNVs often show profound phenotypes including intellectual disability and malformations, rendering them unlikely candidates for eating disorders if not accompanied by other severe clinical manifestations. However, it is possible that smaller CNVs that are more difficult to analyze with current methods could still be relevant for eating disorders. Investigations of CNVs in AN have showed no association so far (Yilmaz et al., 2017), but efforts are currently underway on larger sample sizes.

6. Epigenetics

Epigenetics investigate mechanisms of gene regulation that are hypothesized to mediate the interaction of our genome with environmental stimuli. Three types of epigenetic modifications have been described: methylation, histone modifications, and non-coding RNAs. A comprehensive review on epigenetic modification and their mechanisms can be found at (Bartel, 2009; Roadmap Epigenomics Consortium et al., 2015). All three modifications change the accessibility of DNA and enable or disable the translation of certain DNA segments into proteins. Epigenetic research in eating disorders has not yet yielded consistent findings. Methylation is the only epigenetic modification to have been studied and most studies focused on AN. Very small sample sizes have hampered interpretation and replication. For the future, longitudinal epigenome-wide association studies (EWAS) that control for environmental confounders, such as sex, age, smoking, alcohol consumption, medication, and diet, and that include thousands of eating disorder cases, will complement other genetic approaches. A comprehensive review of epigenetics is beyond the scope of this article, but can be found elsewhere (Hübel, Marzi, Breen, & Bulik, 2018; Thaler & Steiger, 2017).

7. Conclusion

Intensifying the search for contributing genetic factors in eating disorders is essential not only to explicate the biological causal mechanisms underlying the illnesses, but also to improve our ability to understand why environmental factors differentially influence individuals. We acknowledge the important role of environment in the etiology of eating disorders. Indeed, the contribution of environment is inherent in the fact that eating disorders are not 100% heritable.

One advantage of lagging behind other disorders in terms of progress in psychiatric genetics, is that teams working on other conditions such as schizophrenia and major depressive disorder, have paved the way by developing and refining methodologies and analytic strategies. This focuses our efforts on amassing adequate sample sizes to enable the identification of underlying biological processes. Rigorous genetic investigations will not only improve our understanding of the molecular biology of eating disorders, but will also allow us to identify resilience and risk factors involved in the complex interplay of genetics, development, and a multitude of environmental factors implicated in eating disorders (Wray et al., 2014). Poorly conducted analyses, inaccurate interpretations, and overly optimistic reporting of spurious associations can lead to a waste of resources and distort our understanding of the biological mechanisms for years to come (Henriksen, Nordgaard, & Jansson, 2017). Reliable and replicated results, in contrast, will stimulate follow-on research that will lead to better understanding of the illnesses and ultimately improvement of prevention and intervention for individuals with eating disorders. In the long run, the confident identification of genes and pathways implicated in eating disorders etiology will enable the repositioning of available medications and even the development of novel medications to alleviate symptoms or target causal or maintaining mechanisms in eating disorders (Breen et al., 2016).

In addition, being able to quantify genetic risk through the use of PRS and other techniques not yet developed, may improve our ability to determine individual risk, compare environmental effects between high- or low-risk individuals with and without eating disorders to identify triggering and/or protective environmental effects, offer preventative interventions (if appropriate), be more vigilant for early symptoms in high-risk individuals, and inform personalization of treatment. We underscore that genetic risk is only one factor in determining whether an individual develops a given illness. Even someone with a high PRS for AN may

never develop the illness (e.g., due to other resilience or protective genes and/or under favorable environmental conditions). Likewise, someone with low genetic risk could develop AN for other reasons (e.g., other genetic risk factors or high burden of adverse environmental exposures). Ultimately, the goals of genetic research are to deliver rigorous "actionable" findings for eating disorders, defined as genomic results that 1) reveal the fundamental biology of the illnesses, 2) inform clinical practice, and 3) deliver new therapeutic targets (Sullivan et al., 2018).

References

1000 Genomes Project Consortium, Abecasis, G. R., Auton, A., Brooks, L. D., DePristo, M. A., Durbin, R. M., … McVean, G. A. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature*, *491*(7422), 56–65.

Agerbo, E., Sullivan, P. F., Vilhjálmsson, B. J., Pedersen, C. B., Mors, O., Børglum, A. D., … Mortensen, P. B. (2015). Polygenic risk score, parental socioeconomic status, family history of psychiatric disorders, and the risk for schizophrenia: A Danish population-based study and meta-analysis. *JAMA Psychiatry*, *72*(7), 635–641.

Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Bras, J., Duncan, L., Escott-Price, V., … Neale, B. (2016). *Analysis of shared heritability in common disorders of the brain*. *bioRxiv*. doi.org/10.1101/048991

Astle, W., & Balding, D. J. (2009). Population structure and cryptic relatedness in genetic association studies. *Statistical Science*, *24*(4), 451–471.

Bacanu, S.-A., Bulik, C. M., Klump, K. L., Fichter, M. M., Halmi, K. A., Keel, P., … Devlin, B. (2005). Linkage analysis of anorexia and bulimia nervosa cohorts using selected behavioral phenotypes as quantitative traits or covariates. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics, 139B*(1), 61–68.

Baker, J. H., Mitchell, K. S., Neale, M. C., & Kendler, K. S. (2010). Eating disorder symptomatology and substance use disorders: Prevalence and shared risk in a population based twin sample. *International Journal of Eating Disorders*, *43*(7), 648–658.

Bartel, D. P. (2009). MicroRNAs: Target recognition and regulatory functions. *Cell*, *136*(2), 215–233.

Bergen, A. W., van den Bree, M. B. M., Yeager, M., Welch, R., Ganjei, J. K., Haque, K., … Kaye, W. H. (2003). Candidate genes for anorexia nervosa in the 1p33-36 linkage region: Serotonin 1d and delta opioid receptor loci exhibit significant association to anorexia nervosa. *Molecular Psychiatry*, *8*(4), 397–406.

Boraska, V., Franklin, C. S., Floyd, J. A. B., Thornton, L. M., Huckins, L. M., Southam, L., … Bulik, C. M. (2014). A genome-wide association study of anorexia nervosa. *Molecular Psychiatry*, *19*(10), 1085–1094.

Brandys, M. K., de Kovel, C. G. F., Kas, M. J., van Elburg, A. A., & Adan, R. A. H. (2015). Overview of genetic research in anorexia nervosa: The past, the present and the future. *International Journal of Eating Disorders*, *48*(7), 814–825.

Breen, G., Li, Q., Roth, B. L., O'Donnell, P., Didriksen, M., Dolmetsch, R., … Edenberg, H. J. (2016). Translating genome-wide association findings into new therapeutics for psychiatry. *Nature Neuroscience*, *19*(11), 1392–1396.

Breithaupt, L., Hubel, C., & Bulik, C. M. (2018). Updates on genome-wide association findings in eating disorders and future application to precision medicine. *Current Neuropharmacology*. doi.org/10.2174/1570159X16666180222163450

Bruder, C. E. G., Piotrowski, A., Gijsbers, A. A. C. J., Andersson, R., Erickson, S., Diaz de Ståhl, T., … Dumanski, J. P. (2008). Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. *American Journal of Human Genetics*, *82*(3), 763–771.

Bulik, C. M., Devlin, B., Bacanu, S.-A., Thornton, L., Klump, K. L., Fichter, M. M., … Kaye, W. H. (2003). Significant linkage on chromosome 10p in families with bulimia nervosa. *American Journal of Human Genetics*, *72*(1), 200–207.

Bulik, C. M., Sullivan, P. F., & Kendler, K. S. (1998). Heritability of binge-eating and broadly defined bulimia nervosa. *Biological Psychiatry*, *44*(12), 1210–1218.

Bulik, C. M., Sullivan, P. F., Tozzi, F., Furberg, H., Lichtenstein, P., & Pedersen, N. L. (2006). Prevalence, heritability, and prospective risk factors for anorexia nervosa. *Archives of General Psychiatry*, *63*(3), 305–312.

Bulik, C. M., Thornton, L. M., Root, T. L., Pisetsky, E. M., Lichtenstein, P., & Pedersen, N. L. (2010). Understanding the relation between anorexia nervosa and bulimia nervosa in a Swedish national twin sample. *Biological Psychiatry*, *67*(1), 71–77.

Bulik-Sullivan, B. K., Loh, P. R., Finucane, H. K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics Consortium, … Neale, B. M. (2015). LD score regression distinguishes confounding from polygenicity in genome-wide association

studies. *Nature Genetics*, *47*(3), 291–295.

Castellini, G., Lo Sauro, C., Mannucci, E., Ravaldi, C., Rotella, C. M., Faravelli, C., & Ricca, V. (2011). Diagnostic crossover and outcome predictors in eating disorders according to DSM-IV and DSM-V proposed criteria: A 6-year follow-up study. *Psychosomatic Medicine*, *73*(3), 270–279.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*, 7.

Charney, E. (2012). Behavior genetics and postgenomics. *Behavioral and Brain Sciences*, *35*(5), 331–358.

Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics - Simulation and Computation*, *39*(4), 860–864.

Claussnitzer, M., Dankel, S. N., Kim, K.-H., Quon, G., Meuleman, W., Haugen, C., … Kellis, M. (2015). FTO obesity variant circuitry and adipocyte browning in humans. *New England Journal of Medicine*, *373*(10), 895–907.

Coleman, J. R. I., Euesden, J., Patel, H., Folarin, A. A., Newhouse, S., & Breen, G. (2016). Quality control, imputation and analysis of genome-wide genotyping data from the Illumina HumanCoreExome microarray. *Briefings in Functional Genomics*, *15*(4), 298–304.

Coleman, J. R. I., Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, UK Biobank Mental Health Consortium, Eley, T. C., & Breen, G. (2018). *Genome-wide gene-environment analyses of depression and reported lifetime traumatic experiences in uk biobank*. *bioRxiv*. doi.org/10.1101/247353

Conley, D., Rauscher, E., Dawes, C., Magnusson, P. K. E., & Siegal, M. L. (2013). Heritability

and the equal environments assumption: Evidence from multiple samples of misclassified twins. *Behavior Genetics*, *43*(5), 415–426.

Cross-Disorder Group of the Psychiatric Genomics Consortium, Lee, S. H., Ripke, S., Neale, B. M., Faraone, S. V., Purcell, S. M., … International Inflammatory Bowel Disease Genetics Consortium (IIBDGC). (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics*, *45*(9), 984–994.

Davis, C. (2015). The epidemiology and genetics of binge eating disorder (BED). *CNS Spectrums*, *20*(6), 522–529.

de Leeuw, C. A., Neale, B. M., Heskes, T., & Posthuma, D. (2016). The statistical properties of gene-set analysis. *Nature Reviews Genetics*, *17*(6), 353–364.

Dellava, J. E., Thornton, L. M., Lichtenstein, P., Pedersen, N. L., & Bulik, C. M. (2011). Impact of broadening definitions of anorexia nervosa on sample characteristics. *Journal of Psychiatric Research*, *45*(5), 691–698.

Devlin, B., Bacanu, S.-A., Klump, K. L., Bulik, C. M., Fichter, M. M., Halmi, K. A., … Kaye, W. H. (2002). Linkage analysis of anorexia nervosa incorporating behavioral covariates. *Human Molecular Genetics*, *11*(6), 689–696.

Dima, D., & Breen, G. (2015). Polygenic risk scores in imaging genetics: Usefulness and applications. *Journal of Psychopharmacology*, *29*(8), 867–871.

D'Onofrio, B. M., Lahey, B. B., Turkheimer, E., & Lichtenstein, P. (2013). Critical need for family-based, quasi-experimental designs in integrating genetic and social science research. *American Journal of Public Health*, *103 Suppl 1*, S46–55.

Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS Genetics*, *9*(3), e1003348.

Dudbridge, F. (2016). Polygenic epidemiology. *Genetic Epidemiology*, *40*(4), 268–272.

Dudbridge, F., & Gusnanto, A. (2008). Estimation of significance thresholds for genomewide association scans. *Genetic Epidemiology*, *32*(3), 227–234.

Duncan, L., & Keller, M. C. (2011). A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *American Journal of Psychiatry*, *168*(10), 1041–1049.

Duncan, L., Yilmaz, Z., Gaspar, H., Walters, R., Goldstein, J., Anttila, V., … Bulik, C. M. (2017). Significant locus and metabolic genetic correlations revealed in genome-wide association study of anorexia nervosa. *American Journal of Psychiatry*, *174*(9), 850–858.

Dunn, E. C., Wiste, A., Radmanesh, F., Almli, L. M., Gogarten, S. M., Sofer, T., … Smoller, J. W. (2016). Genome-wide association study (GWAS) and genome-wide by environment interaction study (GWEIS) of depressive symptoms in African American and Hispanic/Latina women. *Depression and Anxiety*, *33*(4), 265–280.

Eddy, K. T., Dorer, D. J., Franko, D. L., Tahilani, K., Thompson-Brenner, H., & Herzog, D. B. (2008). Diagnostic crossover in anorexia nervosa and bulimia nervosa: Implications for DSM-V. *American Journal of Psychiatry*, *165*(2), 245–250.

Ekeroth, K., Clinton, D., Norring, C., & Birgegård, A. (2013). Clinical characteristics and distinctiveness of DSM5 eating disorder diagnoses: Findings from a large naturalistic clinical database. *Journal of Eating Disorders*, *1*, 31.

Euesden, J., Lewis, C. M., & O'Reilly, P. F. (2015). PRSice: polygenic risk score software. *Bioinformatics*, *31*(9), 1466–1468.

Evangelou, E., & Ioannidis, J. P. A. (2013). Meta-analysis methods for genome-wide association studies and beyond. *Nature Reviews Genetics*, *14*(6), 379–389.

Felson, J. (2014). What can we learn from twin studies? A comprehensive evaluation of the equal environments assumption. *Social Science Research*, *43*, 184–199.

Gaspar, H. A., & Breen, G. (2017). Drug enrichment and discovery from schizophrenia genome-wide association results: An analysis and visualisation approach. *Scientific Reports*, *7*(1), 12460.

Gelernter, J. (2015). Genetics of complex traits in psychiatry. *Biological Psychiatry*, *77*(1), 36–42.

González, F., Boué, S., & Izpisúa Belmonte, J. C. (2011). Methods for making induced pluripotent stem cells: Reprogramming à la carte. *Nature Reviews Genetics*, *12*(4), 231–242.

Goodman, A., Heshmati, A., Malki, N., & Koupil, I. (2014). Associations between birth characteristics and eating disorders across the life course: Findings from 2 million males and females born in Sweden, 1975-1998. *American Journal of Epidemiology*, *179*(7), 852–863.

Grice, D. E., Halmi, K. A., Fichter, M. M., Strober, M., Woodside, D. B., Treasure, J. T., … Berrettini, W. H. (2002). Evidence for a susceptibility gene for anorexia nervosa on chromosome 1. *American Journal of Human Genetics*, *70*(3), 787–792.

Hawrylycz, M. J., Lein, E. S., Guillozet-Bongaarts, A. L., Shen, E. H., Ng, L., Miller, J. A., … Jones, A. R. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, *489*(7416), 391–399.

Hellwege, J. N., Keaton, J. M., Giri, A., Gao, X., Velez Edwards, D. R., & Edwards, T. L. (2017). Population stratification in genetic association studies. *Current Protocols in Human Genetics*, *95*, 1.22.1–1.22.23.

Henriksen, M. G., Nordgaard, J., & Jansson, L. B. (2017). Genetics of schizophrenia: Overview of methods, findings and limitations. *Frontiers in Human Neuroscience*, *11*, 322.

Hinney, A., & Volckmar, A.-L. (2013). Genetics of eating disorders. *Current Psychiatry*

Reports, *15*(12), 423.

Hübel, C., Marzi, S., Breen, G., & Bulik, C. M. (2017). Epigenetics in eating disorders: a systematic review. Manuscript submitted for publication.

Hudson, J. I., Lalonde, J. K., Berry, J. M., Pindyck, L. J., Bulik, C. M., Crow, S. J., … Pope, H. G., Jr. (2006). Binge-eating disorder as a distinct familial phenotype in obese individuals. *Archives of General Psychiatry*, *63*(3), 313–319.

Hudson, J. I., Pope, H. G., Jr, Jonas, J. M., & Yurgelun-Todd, D. (1983). Family history study of anorexia nervosa and bulimia. *British Journal of Psychiatry*, *142*, 133–138.

Hudson, J. I., Pope, H. G., Jr, Jonas, J. M., Yurgelun-Todd, D., & Frankenburg, F. R. (1987). A controlled family history study of bulimia. *Psychological Medicine*, *17*(4), 883–890.

Hyman, S. (2014). Mental health: Depression needs large human-genetics studies. *Nature*, *515*(7526), 189–191.

Insel, T. R. (2015). The anatomy of NIMH funding. White paper downloaded from: http://www.nimh.nih.gov/funding/funding-strategy-for-research-grants/white-paper_149362.pdf.

International HapMap Consortium. (2005). A haplotype map of the human genome. *Nature*, *437*(7063), 1299–1320.

Javaras, K. N., Laird, N. M., Reichborn-Kjennerud, T., Bulik, C. M., Pope, H. G., Jr, & Hudson, J. I. (2008). Familiality and heritability of binge eating disorder: Results of a case-control family study and a twin study. *International Journal of Eating Disorders*, *41*(2), 174–179.

Kao, P. Y. P., Leung, K. H., Chan, L. W. C., Yip, S. P., & Yap, M. K. H. (2017). Pathway analysis of complex diseases for GWAS, extending to consider rare variants, multi-omics and interactions. *Biochimica et Biophysica Acta*, *1861*(2), 335–353.

Kassett, J. A., Gershon, E. S., Maxwell, M. E., Guroff, J. J., Kazuba, D. M., Smith, A. L., …

Jimerson, D. C. (1989). Psychiatric disorders in the first-degree relatives of probands with bulimia nervosa. *American Journal of Psychiatry*, *146*(11), 1468–1471.

Keck, P. E., Jr, Pope, H. G., Jr, Hudson, J. I., McElroy, S. L., Yurgelun-Todd, D., & Hundert, E. M. (1990). A controlled study of phenomenology and family history in outpatients with bulimia nervosa. *Comprehensive Psychiatry*, *31*(4), 275–283.

Klump, K. L., Miller, K. B., Keel, P. K., McGue, M., & Iacono, W. G. (2001). Genetic and environmental influences on anorexia nervosa syndromes in a population-based twin sample. *Psychological Medicine*, *31*(04), 737–740.

Klump, K. L., Suisman, J. L., Burt, S. A., McGue, M., & Iacono, W. G. (2009). Genetic and environmental influences on disordered eating: An adoption study. *Journal of Abnormal Psychology*, *118*(4), 797–805.

Kortegaard, L. S., Hoerder, K., Joergensen, J., Gillberg, C., & Kyvik, K. O. (2001). A preliminary population-based twin study of self-reported eating disorder. *Psychological Medicine*, *31*(02), 361–365.

Lichtenstein, P., Yip, B. H., Björk, C., Pawitan, Y., Cannon, T. D., Sullivan, P. F., & Hultman, C. M. (2009). Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: A population-based study. *Lancet*, *373*(9659), 234–239.

Lilenfeld, L. R., Kaye, W. H., Greeno, C. G., Merikangas, K. R., Plotnicov, K., Pollice, C., … Nagy, L. (1998). A controlled family study of anorexia nervosa and bulimia nervosa: Psychiatric disorders in first-degree relatives and effects of proband comorbidity. *Archives of General Psychiatry*, *55*(7), 603–610.

Lin, D. Y., & Zeng, D. (2010). Meta-analysis of genome-wide association studies: No efficiency gain in using individual participant data. *Genetic Epidemiology*, *34*(1), 60–66.

Liu, C., Molenaar, P. C. M., & Neiderhiser, J. M. (2017). The impact of variation in twin

relatedness on estimates of heritability and environmental influences. *Behavior Genetics*. doi.org/10.1007/s10519-017-9875-x

Maherali, N., & Hochedlinger, K. (2008). Guidelines and techniques for the generation of induced pluripotent stem cells. *Cell Stem Cell*, *3*(6), 595–605.

Maier, R. M., Visscher, P. M., Robinson, M. R., & Wray, N. R. (2017). Embracing polygenicity: A review of methods and tools for psychiatric genetics research. *Psychological Medicine*, 1–19.

Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Ripke, S., Wray, N. R., Lewis, C. M., Hamilton, S. P., Weissman, M. M., … Sullivan, P. F. (2013). A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular Psychiatry*, *18*(4), 497–511.

Mangweth, B., Hudson, J. I., Pope, H. G., Hausmann, A., De Col, C., Laird, N. M., … Tsuang, M. T. (2003). Family study of the aggregation of eating disorders and mood disorders. *Psychological Medicine*, *33*(7), 1319–1323.

Marchini, J., Cardon, L. R., Phillips, M. S., & Donnelly, P. (2004). The effects of human population structure on large genetic association studies. *Nature Genetics*, *36*(5), 512–517.

Mazzeo, S. E., Mitchell, K. S., Bulik, C. M., Aggen, S. H., Kendler, K. S., & Neale, M. C. (2010). A twin study of specific bulimia nervosa symptoms. *Psychological Medicine*, *40*(7), 1203–1213.

Mazzeo, S. E., Mitchell, K. S., Bulik, C. M., Reichborn-Kjennerud, T., Kendler, K. S., & Neale, M. C. (2009). Assessing the heritability of anorexia nervosa symptoms using a marginal maximal likelihood approach. *Psychological Medicine*, *39*(3), 463–473.

McAdams, T. A., Neiderhiser, J. M., Rijsdijk, F. V., Narusyte, J., Lichtenstein, P., & Eley, T. C. (2014). Accounting for genetic and environmental confounds in associations between parent

and child characteristics: A systematic review of children-of-twins studies. *Psychological Bulletin*, *140*(4), 1138–1173.

McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A. R., Teumer, A., … Haplotype Reference Consortium. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nature Genetics*, *48*(10), 1279–1283.

McGuffin, P., & Burke, A. (2014). Family study and relative risk. In *Wiley StatsRef: Statistics Reference Online*. John Wiley & Sons, Ltd.

Milos, G., Spindler, A., Schnyder, U., & Fairburn, C. G. (2005). Instability of eating disorder diagnoses: prospective study. *British Journal of Psychiatry*, *187*, 573–578.

Mitchell, K. S., Neale, M. C., Bulik, C. M., Aggen, S. H., Kendler, K. S., & Mazzeo, S. E. (2010). Binge eating disorder: A symptom-level investigation of genetic and environmental influences on liability. *Psychological Medicine*, *40*(11), 1899–1906.

Munn-Chernoff, M. A., Keel, P. K., Klump, K. L., Grant, J. D., Bucholz, K. K., Madden, P. A. F., … Duncan, A. E. (2015). Prevalence of and familial influences on purging disorder in a community sample of female twins. *International Journal of Eating Disorders*, *48*(6), 601–606.

Murray, S. B., Pila, E., Griffiths, S., & Le Grange, D. (2017). When illness severity and research dollars do not align: Are we overlooking eating disorders? *World Psychiatry*, *16*(3), 321.

Narusyte, J., Neiderhiser, J. M., D'Onofrio, B. M., Reiss, D., Spotts, E. L., Ganiban, J., & Lichtenstein, P. (2008). Testing different types of genotype-environment correlation: An extended children-of-twins model. *Developmental Psychology*, *44*(6), 1591–1603.

Nelson, M. R., Tipney, H., Painter, J. L., Shen, J., Nicoletti, P., Shen, Y., … Sanseau, P. (2015). The support of human genetic evidence for approved drug indications. *Nature Genetics*, *47*(8), 856–860.

Novembre, J., Johnson, T., Bryc, K., Kutalik, Z., Boyko, A. R., Auton, A., … Bustamante, C. D. (2008). Genes mirror geography within Europe. *Nature*, *456*(7218), 98–101.

Park, J.-H., Gail, M. H., Weinberg, C. R., Carroll, R. J., Chung, C. C., Wang, Z., … Chatterjee, N. (2011). Distribution of allele frequencies and effect sizes and their interrelationships for common genetic susceptibility variants. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(44), 18026–18031.

Peers, I. S., South, M. C., Ceuppens, P. R., Bright, J. D., & Pilling, E. (2014). Can you trust your animal study data? *Nature Reviews Drug Discovery*, *13*(7), 560.

Phipson, B., & Smyth, G. K. (2010). Permutation p-values should never be zero: calculating exact p-values when permutations are randomly drawn. *Statistical Applications in Genetics and Molecular Biology*, *9*, Article39.

Plomin, R., Haworth, C. M., & Davis, O. S. (2009). Common disorders are quantitative traits. *Nature Reviews Genetics*, *10*(12), 872–878.

Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, *47*(7), 702–709.

Price, A. L., Zaitlen, N. A., Reich, D., & Patterson, N. (2010). New approaches to population stratification in genome-wide association studies. *Nature Reviews Genetics*, *11*(7), 459–463.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Psychiatric GWAS Consortium Bipolar Disorder Working Group. (2011). Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near odz4. *Nature Genetics*, *43*(10), 977–983.

Pulit, S. L., Voight, B. F., & de Bakker, P. I. W. (2010). Multiethnic genetic association studies

improve power for locus discovery. *PloS One*, *5*(9), e12600.

Purcell, S., Cherny, S. S., & Sham, P. C. (2003). Genetic power calculator: Design of linkage and association genetic mapping studies of complex traits. *Bioinformatics*, *19*(1), 149–150.

Richards, A. L., Leonenko, G., Walters, J. T., Kavanagh, D. H., Rees, E. G., Evans, A., … O'Donovan, M. C. (2016). Exome arrays capture polygenic rare variant contributions to schizophrenia. *Human Molecular Genetics*, *25*(5), 1001–1007.

Rijsdijk, F. V., & Sham, P. C. (2002). Analytic approaches to twin data using structural equation models. *Briefings in Bioinformatics*, *3*(2), 119–133.

Risch, N., & Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science*, *273*(5281), 1516–1517.

Rivinus, T. M., Biederman, J., Herzog, D. B., Kemper, K., Harper, G. P., Harmatz, J. S., & Houseworth, S. (1984). Anorexia nervosa and affective disorders: A controlled family history study. *American Journal of Psychiatry*, *141*(11), 1414–1418.

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.

Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. (2011). Genome-wide association study identifies five new schizophrenia loci. *Nature Genetics*, *43*(10), 969–976.

Schizophrenia Working Group of the Psychiatric Genomics Consortium, Ripke, S., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K.-H., … O'Donovan, M. C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, *511*(7510), 421–427.

Schwarze, K., Buchanan, J., Taylor, J. C., & Wordsworth, S. (2018). Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature.

Genetics in Medicine. https://doi.org/10.1038/gim.2017.247

Sekar, A., Bialas, A. R., de Rivera, H., Davis, A., Hammond, T. R., Kamitaki, N., … McCarroll, S. A. (2016). Schizophrenia risk from complex variation of complement component 4. *Nature*, *530*(7589), 177–183.

Sham, P. C., & Purcell, S. M. (2014). Statistical power and significance testing in large-scale genetic studies. *Nature Reviews Genetics*, *15*(5), 335–346.

Smemo, S., Tena, J. J., Kim, K.-H., Gamazon, E. R., Sakabe, N. J., Gómez-Marín, C., … Nóbrega, M. A. (2014). Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature*, *507*(7492), 371–375.

Song, J., Bergen, S. E., Kuja-Halkola, R., Larsson, H., Landén, M., & Lichtenstein, P. (2015). Bipolar disorder and its relation to major psychiatric disorders: A family-based study in the Swedish population. *Bipolar Disorders*, *17*(2), 184–193.

St Clair, D., Blackwood, D., Muir, W., Carothers, A., Walker, M., Spowart, G., … Evans, H. J. (1990). Association within a family of a balanced autosomal translocation with major mental illness. *Lancet*, *336*(8706), 13–16.

Stein, D., Lilenfeld, L. R., Plotnicov, K., Pollice, C., Rao, R., Strober, M., & Kaye, W. H. (1999). Familial aggregation of eating disorders: Results from a controlled family study of bulimia nervosa. *International Journal of Eating Disorders*, *26*(2), 211–215.

Steinhausen, H.-C., Jakobsen, H., Helenius, D., Munk-Jørgensen, P., & Strober, M. (2015). A nation-wide study of the family aggregation and risk factors in anorexia nervosa over three generations. *International Journal of Eating Disorders*, *48*(1), 1–8.

Strober, M., Freeman, R., Lampert, C., Diamond, J., & Kaye, W. (2000). Controlled family study of anorexia nervosa and bulimia nervosa: Evidence of shared liability and transmission of partial syndromes. *American Journal of Psychiatry*, *157*(3), 393–401.

Strober, M., Freeman, R., Lampert, C., Diamond, J., & Kaye, W. (2001). Males with anorexia nervosa: A controlled study of eating disorders in first-degree relatives. *International Journal of Eating Disorders*, *29*(3), 263–269.

Strober, M., Lampert, C., Morrell, W., Burroughs, J., & Jacobs, C. (1990). A controlled family study of anorexia nervosa: evidence of familial aggregation and lack of shared transmission with affective disorders. *The International Journal of Eating Disorders*, *9*(3), 239–253.

Strober, M., Morrell, W., Burroughs, J., Salkin, B., & Jacobs, C. (1985). A controlled family study of anorexia nervosa. *Journal of Psychiatric Research*, *19*(2-3), 239–246.

Sullivan, P. F., Agrawal, A., Bulik, C. M., Andreassen, O. A., Børglum, A. D., Breen, G., … Psychiatric Genomics Consortium. (2018). Psychiatric genomics: An update and an agenda. *American Journal of Psychiatry*, *175*(1), 15–27.

Sullivan, P. F., Kendler, K. S., & Neale, M. C. (2003). Schizophrenia as a complex trait: Evidence from a meta-analysis of twin studies. *Archives of General Psychiatry*, *60*(12), 1187–1192.

Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression: Review and meta-analysis. *The American Journal of Psychiatry*, *157*(10), 1552–1562.

Sung, Y. J., Schwander, K., Arnett, D. K., Kardia, S. L. R., Rankinen, T., Bouchard, C., … Rao, D. C. (2014). An empirical comparison of meta-analysis and mega-analysis of individual participant data for identifying gene-environment interactions. *Genetic Epidemiology*, *38*(4), 369–378.

Thaler, L., & Steiger, H. (2017). Eating disorders and epigenetics. *Advances in Experimental Medicine and Biology*, *978*, 93–103.

Thapar, A., & Cooper, M. (2013). Copy number variation: What is it and what has it told us

about child psychiatric disorders? *Journal of the American Academy of Child and Adolescent Psychiatry*, *52*(8), 772–774.

Timpson, N. J., Greenwood, C. M. T., Soranzo, N., Lawson, D. J., & Richards, J. B. (2018). Genetic architecture: The shape of the genetic contribution to human traits and disease. *Nature Reviews Genetics*, *19*(2), 110–124.

Trace, S. E., Thornton, L. M., Baker, J. H., Root, T. L., Janson, L. E., Lichtenstein, P., … Bulik, C. M. (2013). A behavioral-genetic investigation of bulimia nervosa and its relationship with alcohol use disorder. *Psychiatry Research*, *208*(3), 232–237.

Vinkhuyzen, A. A. E., & Wray, N. R. (2015). Novel directions for G × E analysis in psychiatry. *Epidemiology and Psychiatric Sciences*, *24*(1), 12–19.

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era--concepts and misconceptions. *Nature Reviews Genetics*, *9*(4), 255–266.

Wade, T. D., Bulik, C. M., Neale, M., & Kendler, K. S. (2000). Anorexia nervosa and major depression: Shared genetic and environmental risk factors. *American Journal of Psychiatry*.

Wang, K., Zhang, H., Bloss, C. S., Duvvuri, V., Kaye, W., Schork, N. J., … Price Foundation Collaborative Group. (2011). A genome-wide association study on common SNPs and rare CNVs in anorexia nervosa. *Molecular Psychiatry*, *16*(9), 949–959.

Wang, W. Y. S., Barratt, B. J., Clayton, D. G., & Todd, J. A. (2005). Genome-wide association studies: Theoretical and practical concerns. *Nature Reviews Genetics*, *6*(2), 109–118.

Wildes, J. E., & Marcus, M. D. (2015). Application of the research domain criteria (RDoC) framework to eating disorders: Emerging concepts and research. *Current Psychiatry Reports*, *17*(5), 30.

Wong, C. C. Y., Caspi, A., Williams, B., Craig, I. W., Houts, R., Ambler, A., … Mill, J. (2010). A longitudinal study of epigenetic variation in twins. *Epigenetics*, *5*(6), 516–526.

Wray, N. R., & Gottesman, I. I. (2012). Using summary data from the Danish national registers to estimate heritabilities for schizophrenia, bipolar disorder, and major depressive disorder. *Frontiers in Genetics*, *3*, 118.

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, and Allied Disciplines*, *55*(10), 1068–1087.

Wray, N. R., & Visscher, P. (2008). Estimating trait heritability. *Nature Education*, *1*(1), 161.

Wray, N. R., Yang, J., Hayes, B. J., Price, A. L., Goddard, M. E., & Visscher, P. M. (2013). Pitfalls of predicting complex traits from SNPs. *Nature Reviews Genetics*, *14*(7), 507–515.

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, *88*(1), 76–82.

Yang, J., Zeng, J., Goddard, M. E., Wray, N. R., & Visscher, P. M. (2017). Concepts, estimation and interpretation of SNP-based heritability. *Nature Genetics*, *49*(9), 1304–1310.

Yilmaz, Z., Hardaway, J. A., & Bulik, C. M. (2015). Genetics and epigenetics of eating disorders. *Advances in Genomics and Genetics*, *5*, 131–150.

Yilmaz, Z., Szatkiewicz, J. P., Crowley, J. J., Ancalade, N., Brandys, M. K., van Elburg, A., … Bulik, C. M. (2017). Exploration of large, rare copy number variants associated with psychiatric and neurodevelopmental disorders in individuals with anorexia nervosa. *Psychiatric Genetics*, *27*(4), 152–158.

Young-Pearse, T. L., & Morrow, E. M. (2016). Modeling developmental neuropsychiatric disorders with iPSC technology: Challenges and opportunities. *Current Opinion in Neurobiology*, *36*, 66–73.

Zhang, J., & Stram, D. O. (2014). The role of local ancestry adjustment in association studies using admixed populations. *Genetic Epidemiology*, *38*(6), 502–515.

Zhang, Y., Shen, X., & Pan, W. (2013). Adjusting for population stratification in a fine scale with principal components and sequencing data. *Genetic Epidemiology*, *37*(8), 787–801.

Table

Table 1. Heritability (h^2) estimates derived from different genetic study designs for other psychiatric disorders and eating disorders (Sullivan et al., 2018; Yilmaz et al., 2015).

Disorder	Twin-h^2	Family-h^2	SNP-h^2
Other psychiatric disorders			
Schizophrenia	81% (Sullivan, Kendler, & Neale, 2003)	64% (Lichtenstein et al., 2009)	45% (Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2014)
Bipolar disorder	68% (Polderman et al., 2015)	55% (Song et al., 2015)	21% (Psychiatric GWAS Consortium Bipolar Disorder Working Group, 2011)
Major depressive disorder	37% (Sullivan, Neale, & Kendler, 2000)	32% (Wray & Gottesman, 2012)	8.9% (Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium et al., 2013)

Eating disorders			
Anorexia nervosa	57% (Dellava et al., 2011)	64% (Strober et al., 1990)	~20% (Duncan et al., 2017)
Bulimia nervosa	62% (Bulik et al., 2010)	Not available	Not available
Binge-eating disorder	39% (Javaras et al., 2008)	57% (Javaras et al., 2008)	Not available
Purging disorder	Not available†	Not available	Not available
ARFID	Not available	Not available	Not available

Abbreviations: h^2=narrow-sense heritability; ARFID=avoidant and restrictive food intake disorder

†Munn-Chernoff et al. (2015) conducted a twin study of purging disorder and identified that 44% of the variance attributable to familial factors, but were unable to distinguish between additive genetic and shared environmental factors.

Figure legends

Figure 1. Standard ACE model of the classical twin design. Latent underlying factors: A=additive genetic factors, C=common/shared environmental factors, E=nonshared environmental factors and measurement error. Squares of paths: a^2=heritability, which is in some depictions interchangeably with h^2, c^2=variance of common environmental factors, e^2=nonshared environmental factors.

Figure 2. Different degrees of genetic relatedness of family members. *Niece scientifically and linguistically correct named as first cousin once removed.

Figure 3. Manhattan plot of genome-wide association study (GWAS) of body fat in UK Biobank. The x-axis displays all 22 autosomes and the y-axis the -log10 of the p-values of the association of every genetic marker (i.e., single nucleotide polymorphism, SNPs). The magenta line marks the genome-wide significance level of $5x10^{-8}$ and the blue area underneath indicates SNPs that did not reach genome-wide significance but may carry genetic information which can be summarized in polygenic risk scores (PRS).

Figure 4. Population structure within Europe. Displayed is the genetic data of 1,387 Europeans based on principal components (PC1 vs. PC2). The PC axes are rotated to emphasize the similarity to the geographic map of Europe. AL, Albania; AT, Austria; BA, Bosnia-Herzegovina; BE, Belgium; BG, Bulgaria; CH, Switzerland; CY, Cyprus; CZ, Czech Republic; DE, Germany; DK, Denmark; ES, Spain; FI, Finland; FR, France; GB, United Kingdom; GR, Greece; HR, Croatia; HU, Hungary; IE, Ireland; IT, Italy; KS, Kosovo; LV, Latvia; MK, Macedonia; NO, Norway; NL, Netherlands; PL, Poland; PT, Portugal; RO, Romania; RS, Serbia and Montenegro; RU, Russia, Sct, Scotland; SE, Sweden; SI, Slovenia; SK, Slovakia; TR, Turkey; UA, Ukraine; YG, Yugoslavia. Reprinted by permission from RightsLink: Springer Nature, Nature, Genes mirror geography within Europe, Novembre et. al. © (2008).

Figure 5. Hypothetical relation between numbers of cases and genome-wide significant associations for a human complex disease or trait. After an initial dead zone with no genome-wide significant associations an inflection point will be reached. At this inflection point the first genome-wide significant associations are detected and mostly their number increases linearly. The initial dead zone is dependent on the velocity of sample collection and effect sizes (adapted from Sullivan et al., 2018).