PERSPECTIVES

Common disorders are quantitative traits

Robert Plomin, Claire M. A. Haworth and Oliver S. P. Davis

Abstract | After drifting apart for 100 years, the two worlds of genetics — quantitative genetics and molecular genetics — are finally coming together in genome-wide association (GWA) research, which shows that the heritability of complex traits and common disorders is due to multiple genes of small effect size. We highlight a polygenic framework, supported by recent GWA research, in which qualitative disorders can be interpreted simply as being the extremes of quantitative dimensions. Research that focuses on quantitative traits — including the low and high ends of normal distributions — could have far-reaching implications for the diagnosis, treatment and prevention of the problematic extremes of these traits.

After the rediscovery of Mendel's laws of inheritance in the early 1900s, a controversy arose between Mendelians, who searched for qualitative traits that show Mendelian patterns of inheritance, and biometricians, who focused on quantitative traits that are normally distributed. Although Fisher's 1918 paper[1] provided the basis for reconciling the differences between Mendelians and biometricians, these two worlds of genetics drifted apart because of the differing perspectives that follow from thinking qualitatively versus thinking quantitatively. The two worlds are now being brought together by genome-wide association research (GWA research), which shows that the ubiquitous heritability of common disorders is due to multiple genes of small effect size.

As the two worlds of genetics come together, it is timely to reconsider the implications of Fisher's 1918 paper in relation to common disorders. Fisher showed that quantitative traits could be explained by Mendelian inheritance if multiple genes affect the trait. However, most GWA studies are case–control studies that focus on qualitative traits and typically compare allele frequencies for diagnosed cases versus controls[2,3]. But if GWA studies indicate that multiple genes affect these disorders, this

implies that their genetic liability is distributed quantitatively rather than qualitatively. We examine this disconnection between qualitatively diagnosed disorders and their quantitatively distributed polygenic liabilities. We also consider ways in which research into polygenic liabilities for disorders leads to quantitative traits. We do not address the GWA debates about common versus rare variants or about 'missing heritability'[4]. Also, our focus is on common disorders, which are the targets of many current GWA studies, rather than the thousands of rare monogenic disorders.

Thinking quantitatively has practical implications for GWA research as well as far-reaching conceptual and clinical implications. In our opinion, the polygenic liabilities that emerge from GWA research will lead to common disorders being thought of as the extremes of quantitative traits and, ultimately, to a scientific focus on quantitative traits rather than disorders.

The two worlds of genetics
Mendelians versus biometricians. Molecular genetics and quantitative genetics both have their origins in the early 1900s, when Mendel's laws were rediscovered. The split in the world of genetics began with a

decade-long dispute between Mendelians, who were led by Gregory Bateson, and biometricians, who were led by Karl Pearson. Mendelians looked for Mendelian patterns of inheritance in qualitative traits as the hallmark of inheritance, whereas biometricians argued that Mendel's laws could not apply to quantitative traits because such traits showed no simple pattern of inheritance. Both sides were right and both were wrong. The Mendelians were correct in arguing that heredity works the way Mendel said it worked, but they were wrong in assuming that complex traits show simple Mendelian patterns of inheritance. The biometricians were right in arguing that complex traits are distributed quantitatively not qualitatively, but they were wrong in concluding that Mendel's laws of inheritance are particular to pea plants and do not apply to complex traits.

The disagreement between the Mendelians and biometricians was resolved when biometricians realized that Mendel's laws of inheritance of single genes would also apply to complex traits that are influenced by several genes, each of which is inherited according to Mendel's laws. This resolution was formalized in R. A. Fisher's 1918 paper, 'The correlation between relatives on the supposition of Mendelian inheritance'[1]. Fisher began with the hypothesis that if several genes affect a trait, the trait will be normally distributed as a quantitative trait, even though each gene is inherited according to Mendel's laws. Although Fisher referred to "a large number of Mendelian factors", even with just three genes the genotypic distribution begins to approach a normal distribution in the population (BOX 1). This paper became the cornerstone of the field of quantitative genetics[5].

Quantitative genetics versus molecular genetics. Despite Fisher's reconciliation, the two worlds of genetics went their separate ways for most of the twentieth century because of their widely different interests and perspectives. Quantitative geneticists investigated the genetics of naturally occurring phenotypic variation by using methods — such as strain and selection studies in non-human animals and twin and adoption studies in

humans — that estimated the cumulative effect of genetic influence regardless of the number of genes involved or the complexity of their effects. By contrast, the progenitors of molecular genetics studied how genes work rather than particular phenotypes: they focused on rare naturally occurring monogenic effects or experimentally induced ones. Both worlds of genetics assumed that it would not be possible to identify specific genes for complex traits and common disorders because they are likely to be influenced by multiple genes of small effect.

Coming together in genome-wide association research. Since the 1980s, the discovery of millions of DNA markers has made it possible to identify genes for complex quantitative traits that are influenced by multiple genes of small effect, not just qualitative monogenic disorders. This trend accelerated rapidly with the development of microarrays that can genotype millions of DNA markers, which have led to an explosion of GWA studies of common complex disorders and diseases[6]. GWA research finds associations of small effect size with odds ratios of less

than two, which implies that many such genes are needed to account for the heritability of the disorders[7].

Mendel showed how monogenic qualitative effects can explain the inheritance of qualitative traits. Fisher showed how qualitative Mendelian inheritance can explain the inheritance of quantitative traits if multiple genes are involved. Here, we draw attention to the underlying implication that qualitative disorders that are influenced by multiple genes are quantitative traits.

Thinking quantitatively
The finding that multiple DNA variants are associated with common disorders is leading to disorders being thought of in quantitative terms. As multiple DNA variants are identified, they can be aggregated into composites that represent the polygenic liability that underlies common disorders. These polygenic liabilities will hopefully lead us to think about disorders as the extremes of quantitative traits and, ultimately, to focus on quantitative traits rather than disorders. But which quantitative traits underlie the common disorders that we study?

For some disorders, the relevant quantitative traits seem obvious, such as body mass index (BMI) for obesity, blood pressure for hypertension, and mood for depression. However, when the polygenic liabilities of these disorders are investigated empirically, the answers are unlikely to be so simple, as discussed later. The relevant quantitative traits are not at all clear for most disorders — including alcoholism, arthritis, autism, cancers, dementia, diabetes and heart disease.

Thinking quantitatively will be aided by speaking quantitatively — a shift in vocabulary is required so that we start talking about 'dimensions' rather than 'disorders' and about genetic 'variability' rather than genetic 'risk'. Thinking quantitatively also requires familiarity with the statistics of quantitative traits, such as linear regression rather than logistic regression, variance rather than mean differences, and covariance rather than comorbidity.

The polygenic liability for common disorders is normally distributed; the challenge is to determine the extent to which quantitative traits reflect this polygenic liability.

Glossary

Autophagy
The degradation by a cell of its own components. In the immune response, autophagy removes intracellular bacteria and viruses, and enhances adaptive immunity.

Case–control study
Compares cases (a selected group of individuals, for example, those diagnosed with a disorder) with controls (a comparison group of individuals, for example, those who are not diagnosed with the disorder). Genome-wide association case–control studies test whether genetic marker allele frequencies differ between cases and controls.

Comorbidity
The co-occurrence of two or more disorders or diseases in an individual.

Covariance
A statistic that indicates the extent to which two variables are related and vary together.

Crohn's disease
Characterized by chronic intestinal inflammation, which leads to diarrhoea, bleeding, severe abdominal pain and weight loss.

Effect size
The proportion of individual differences for a trait in the population that are accounted for by a particular factor.

Genome-wide association research
A hypothesis-free genetic method that uses hundreds of thousands of DNA markers distributed throughout the chromosomes to identify alleles that are correlated with a trait.

Heritability
The proportion of phenotypic variance in a population that is due to genetic variation.

Linear regression
A statistical method for testing and describing the linear relationship between variables. The regression coefficient describes the slope of the regression line and reflects the amount of variance of the dependent variable that is explained by variation of the independent variable.

Logistic regression
A statistical method for testing and describing the linear relationship between variables when the dependent variable is binary. It relates the log odds of the probability of an event to a linear combination of the predictor variables.

Odds ratio
A measurement of the effect size of an association for binary values. For example, in case–control studies, the odds ratio is calculated as the odds of an allele in cases divided by the odds of the allele in controls. An odds ratio of one indicates that there is no difference in allele frequency between cases and controls.

Pleiotropy
The effect of a single gene on multiple phenotypes.

Population cohort study
A longitudinal study of individuals who are representative of the general population and who are often recruited by their year of birth.

Power
The probability that a statistical test will reject the null hypothesis when the alternative hypothesis is true.

Quantitative genetics
A theory of multiple gene influences that, together with environmental variation, results in quantitative (continuous) distributions of phenotypes. Quantitative genetic methods, such as twin and adoption methods for human analysis, estimate genetic and environmental contributions to phenotypic variance and covariance in a population.

Sensitivity
The proportion of true positives that are accurately identified as such — for example, the percentage of cases that are diagnosed using a questionnaire. A sensitivity of 100% means that all cases are correctly identified.

Specificity
The proportion of true negatives that are classified as negatives. For example, a diagnostic test with specificity of 100% means that all healthy people have been identified as healthy.

Trait
A phenotype that differs between individuals in a species and shows some stability across time and situations. Disorders and diseases are qualitative (dichotomous) traits; quantitative traits are continuously distributed, usually as the bell-shaped curve called the normal distribution.

Variance
A measure of the dispersal of phenotypic scores around the mean.

Web-based testing
Administering online questionnaires and tests using the internet, which allows access to large and geographically diverse samples.

Box 1 | Polygenic traits are quantitative traits

R. A. Fisher's 1918 paper, 'The correlation between relatives on the supposition of Mendelian inheritance', resolved the often bitter conflict between biometricians and Mendelians, which raged for a decade following the rediscovery of Mendel's work. Fisher showed that a complex quantitative trait could be explained by Mendelian inheritance if several genes affect the trait.

Because he crossed true-breeding plants, Mendel's experiments showed that a single locus with two alleles of equal frequency results in three genotypes (see the figure, part **a**). If the allelic effects are additive, the three genotypes produce three phenotypes; in the case of Mendel's qualitative traits, the allelic effects showed complete dominance, so only two phenotypes were observed. However, assuming equal and additive effects, 2 genes yield 9 genotypes and 5 phenotypes (part **b**) and 3 genes yield 27 genotypes and 7 phenotypes (part **c**). With unequal and non-additive allelic effects and some environmental influence, three genes would result in a normal bell-shaped curve of continuous variation (part **d**). This logic assumes common alleles; rare alleles will skew the distribution. Genome-wide association research suggests that many more than three genes affect most traits, which underscores the expectation that polygenic traits are quantitative traits.
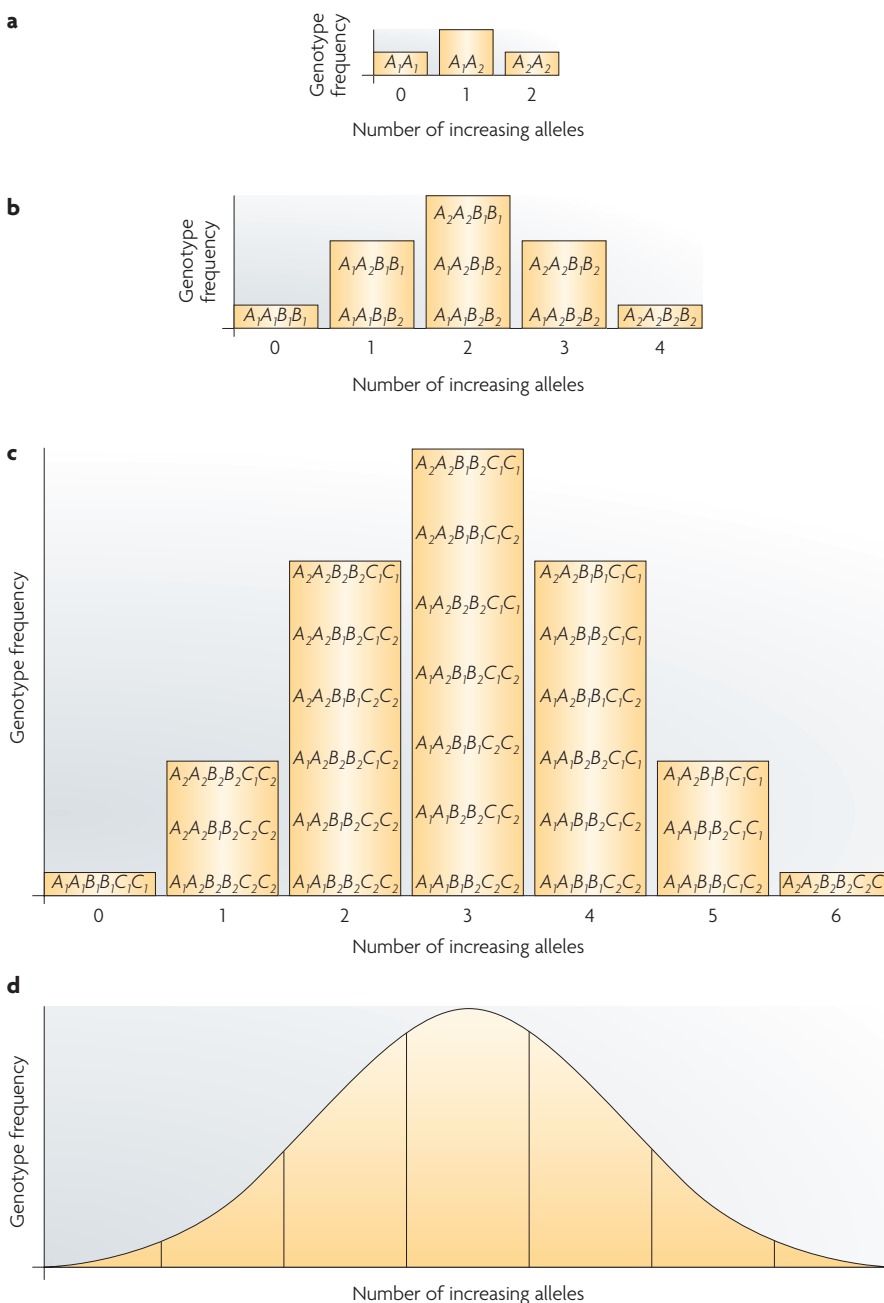
These quantitative traits need not be limited to symptoms of the diagnosed disorder but can occur at any level of analysis, as discussed in the following section.

Identifying quantitative mechanisms

Once multiple genes are found to be associated with a disorder, understanding the mechanisms by which each gene affects the disorder leads to quantitative traits being recognised at all levels of analysis: from gene expression profiles, to other '-omic' levels of analysis, to physiology and often to the structure and function of the brain[8].

For some traits, such as type 2 diabetes (T2D), a quantitative approach has already been embraced, with striking results[9]. Although the first T2D GWA studies were case–control studies (REF. 49, and subsequently other studies, for example, REF. 3), a wave of follow-up studies have focused on quantitative traits that are related to T2D, including levels of fasting glucose[10] and C-reactive protein[11], and glucose tolerance[9]. These studies are leading to refinements in the definition of T2D.

Recent studies of Crohn's disease (CD)[12] have provided less well-known examples of how quantitative traits that are relevant to disease might arise from GWA studies. GWA studies have yielded several genes that are associated with CD susceptibility[13]. The search for the mechanisms by which these genes affect the disorder is leading to quantitative traits, such as inflammatory response[14], bacterial survival and chronic inflammation[15–18]. Recently, GWA research has implicated autophagy as a previously unsuspected quantitative mechanism in CD pathogenesis[19,20].

Other disorders are currently under GWA scrutiny; the mechanisms by which the identified genes affect the disorder are likely to lead to quantitative traits.

Weighting disease genes. GWA studies are revealing that several different quantitative mechanisms underlie disorders and are showing that the sets of variants that are associated with each mechanism sometimes relate to the subtypes of a disease. For example, in the case of CD, nucleotide-binding oligomerization domain-containing 2 (*NOD2*) variants are associated largely with CD of the ileum[21], whereas interleukin 23 receptor (*IL23R*) variants are associated with all CD subphenotypes[22]. This suggests the possibility of using weighted sets of variants to reflect the polygenic liability and to predict clinically useful features[13], which is discussed in the following section on polygenic risk scores.

In addition, emerging evidence suggests that there is considerable and unexpected pleiotropic overlap among the genetic variants that affect different human diseases[2]. For example, one region of chromosome 8q12 is associated with several forms of cancer[23], and a growing number of loci are associated with more than one autoimmune disease[24]. These findings have inspired the concept of the human 'diseasome' — the synthesis of all human genetic disorders (the disease phenome) and all human disease genes (the disease genome)[25]. This approach shows the existence of genetic links not only between subtypes of the same disorder but also between apparently disparate conditions. These links, and their mechanisms, will become clearer as online molecular databases of genome-wide assays mature. In the future, the practice of weighting disease genes in the prediction of quantitative features of a disease phenotype may extend beyond narrow classifications of qualitative diseases to a full quantitative understanding of the multivariate diseasome (BOX 2).

The practical use of polygenic risk scores
Predictive testing. Above we have discussed investigating the pleiotropic effects of each DNA variant independently. However, polygenic liability can be indexed empirically by a set of DNA variants that are associated with the disorder; such a set aggregates the small effects of each DNA variant[26]. We use the term 'polygenic risk score' to refer specifically to the set of multiple DNA variants that are associated with a disorder. (Such composites have also been called polygenic susceptibility scores[27], genomic profiles[28], SNP sets[29], genetic risk scores[30] and aggregate risk scores[31].) Polygenic risk scores are beginning to be used to predict the population-wide genetic risk for common disorders, such as breast cancer[32], atherosclerosis[30], coronary heart disease[33], age-related macular degeneration[34], recurrent venous thrombosis[35] and T2D[36].

In addition to using them to predict diagnosed disorders, polygenic risk scores can be used in research, unencumbered by the demanding clinical thresholds for sensitivity and specificity. We suggest that the research applications of polygenic risk scores will inevitably lead to quantitative traits, partly because polygenic risk scores are themselves normally distributed quantitative traits.

Sensitive genotypic selection. Another use is for studying individuals at high polygenic risk by using polygenic risk scores to select individuals based on their genotype rather than their phenotype. Because many individuals at high polygenic risk will not reach the diagnostic criteria for the disorder, more subtle and sensitive quantitative trait measures will be needed to investigate the multivariate quantitative traits that reflect polygenic risk. Genotypic selection for research into high polygenic risk will be especially useful in intensive and expensive research, such as neuroimaging, which is limited to relatively small sample sizes whereas genetic research requires large sample sizes owing to the small effect sizes of the individual DNA variants.

Thinking positively. The most innovative implication of thinking quantitatively about polygenic liability is that it leads to thinking positively. Instead of focusing only on the vulnerabilities of individuals with high polygenic risk scores, a new direction for research is to consider the resilience of individuals with low polygenic risk scores — the individuals at the neglected 'other end' of the continuum of polygenic liability.

Qualitative thinking about disorders leads to a medical model mindset in which the goal is to fix the problem so that the individual can be returned to normality. By contrast, thinking positively suggests that we should investigate mechanisms that push beyond normality. For example, individuals who are dealt a hand of 'low-risk' alleles for obesity may not just be at low risk for obesity — the 'healthily slim' might also be invulnerable to the temptations of our obesogenic environment. Investigating the positive end of polygenic liability could lead to mechanisms that promote healthy outcomes, which might differ from mechanisms that help us to avoid unhealthy outcomes. Moreover, interventions that are aimed at making people healthier might be more effective at a public health level than those that aim to make people avoid harm.

Box 2 | Pleiotropic relationships and quantitative traits: the example of *FTO*

The robust association between variants in fat mass and obesity associated (*FTO*) and obesity is one of the great success stories of recent genome-wide association studies[48]. The original article also reported the association of *FTO* with body mass index, a quantitative measure. However, as the literature surrounding the pattern of associations matures, it is becoming clear that the association also holds true for quantitative traits at sev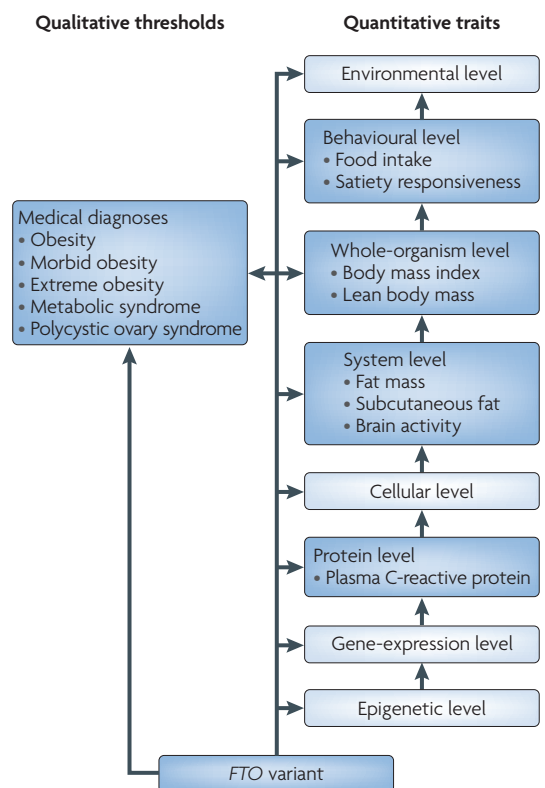eral different levels of analysis. Quantitative associations have already been reported for some of the many possible levels of analysis (indicated by the dark blue boxes in the figure). We predict that quantitative associations will also become apparent at the other levels depicted here, including at the environmental level through gene–environment correlations (light blue boxes indicate the theoretical levels that have not yet been explored in the case of *FTO*). Although medical diagnoses (such as obesity) provide a convenient pragmatic framework for the initial discovery of genetic variants, in scientific terms there are no real 'genes for disorders'. On the contrary, the genetic variants that are implicated in complex traits are associated with quantitative traits at every level of analysis. Thinking and researching quantitatively will provide a much richer picture of the complex biological pathways that lead from genes to disorders and will help us to generate biologically meaningful models of disease aetiology. The medical model is a useful benchmark for translating research into practice, but we must be careful that our diagnoses follow from our science, and not the other way around.

Qualitative thresholds

Medical diagnoses
• Obesity
• Morbid obesity
• Extreme obesity
• Metabolic syndrome
• Polycystic ovary syndrome

FTO variant

Quantitative traits

Environmental level

Behavioural level
• Food intake
• Satiety responsiveness

Whole-organism level
• Body mass index
• Lean body mass

System level
• Fat mass
• Subcutaneous fat
• Brain activity

Cellular level

Protein level
• Plasma C-reactive protein

Gene-expression level

Epigenetic level

Box 3 | **Towards quantitative thinking — phenotypically and genotypically**
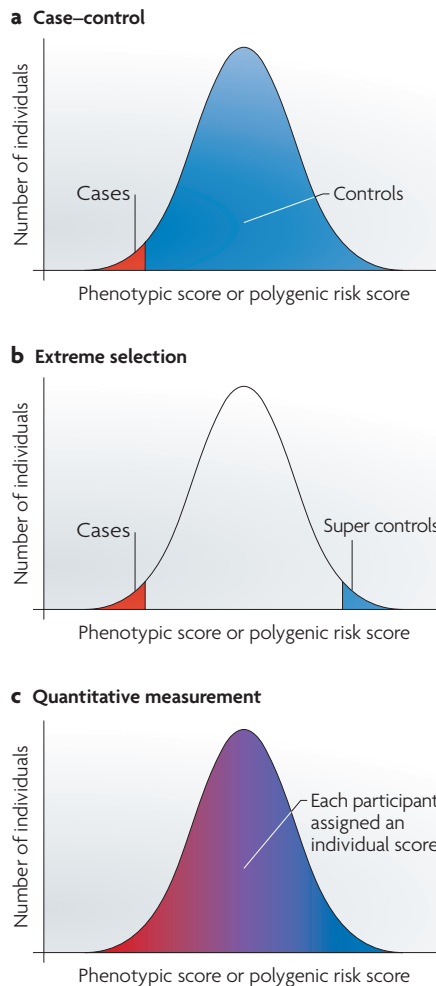
In a traditional case–control study, controls are often selected from among individuals without the disorder, or even from among members of the population irrespective of disease status. This means that in terms of normally distributed polygenic risk, the controls can be nearly (or actually) cases. However, it is often also possible to characterize disorders in terms of quantitative traits and, analogously, to consider the quantitative polygenic risk. Thinking quantitatively in the design of genome-wide association (GWA) studies for common disorders and complex traits can increase the statistical power, help to uncover biological pathways from gene to disorder, and harness the potential of population-based cohort studies that have already collected array data and phenotypic data on multiple quantitative dimensions.

The three risk distributions shown in the figure — 'case–control', 'extreme selection' and 'quantitative measurement' — can be interpreted from both phenotypic and genotypic viewpoints, with the *x* axis representing either phenotypic scores on a quantitative trait or polygenic risk scores, as described in the main text. Phenotypically, the first distribution ('case–control'; see the figure, part **a**) represents the selection of cases and controls in a GWA study of a common disorder. If the underlying disorder is interpreted as a quantitative trait, then many of the so-called 'controls' are phenotypically nearly cases. In studies that use an unselected control sample, some of the controls would indeed meet diagnostic criteria. Similarly, if the distribution is taken to represent polygenic risk scores, then many of the controls are as extreme as the cases in terms of genetic risk.

The second distribution ('extreme selection'; part **b**) represents an alternative selection strategy. Phenotypically, if the common disorder is interpreted as a quantitative trait, 'super controls' can be drawn from the opposite end of the distribution to maximize the statistical power for a given number of arrays by achieving the maximum possible phenotypic separation between the cases and controls. Genotypically, genetic associations from a conventional case–control analysis can be used to form a normally distributed polygenic risk score. The cases are already known to score low on this index, but what about the opposite extreme of the polygenic risk, the individuals with all the 'healthy' alleles? The potential of genomically selecting these individuals on the basis of polygenic risk and investigating them phenotypically is explored in the main text.

The apogee of quantitative thinking is an unselected population sample ('quantitative measurement'; part **c**) in which participants are assigned their own score based on a phenotypic quantitative dimension and researchers adopt a different set of statistical tools to analyse the data. Using a population cohort with multiple quantitative traits that are measured in this way opens up other possibilities for genotypic analysis, such as exploring the genetic overlap among quantitative traits that contribute to a disease and charting the relative contributions of particular genetic variants to build up aetiological profiles that may change throughout development and in the context of different environmental influences.

a Case–control

Cases — Controls

Phenotypic score or polygenic risk score

b Extreme selection

Cases — Super controls

Phenotypic score or polygenic risk score

c Quantitative measurement

Each participant assigned an individual score

Phenotypic score or polygenic risk score

However, this 'other end' of the normal distribution of polygenic liability for disorders is uncharted territory. It is possible that the 'positive' end of the polygenic liability is not so positive — the lowest polygenic risk scores could have problems of their own. For example, being at the lowest end of the polygenic liability for obesity might entail other risks, such as greater risk for not eating well-balanced meals or greater susceptibility to eating disorders. From an evolutionary perspective, averageness might be an adaptive trade-off against the mixture of costs and benefits of more extreme polygenic liabilities, especially given the fluctuating nature of selection[37]. An adaptive edge for averageness also suggests an evolutionary mechanism that would maintain genetic variation in the population and therefore account for the ubiquitous high heritabilities of traits across the life sciences.

Studying the full range of variation. In addition to investigating the low and high extremes of the normal distribution of polygenic liability, polygenic risk scores will lead to research on the full range of normal quantitative trait variation. Studying populations rather than probands might lead to multivariate quantitative traits that reflect polygenic risk scores better than do the disorders themselves. Sometimes multivariate research based on polygenic risk scores will point to a network of related quantitative traits. For example, five SNPs associated with coronary heart disease all seem to be involved in intracellular vesicle trafficking[33]. Most often, however, polygenic risk scores are likely to incorporate DNA variants with differing functions, at least in terms of what is currently known about their function. For example, SNPs that are associated with breast cancer include genes in the cell cycle control pathway and genes that are involved in steroid hormone metabolism and signalling[32]. Nonetheless, their joint association with a disorder indicates that they do in fact overlap in their downstream function. In this way, polygenic risk scores that consist of DNA variants of disparate function could lead to a systems approach that is anchored by their common association with a disorder[38]. Research on the full range of normal variation in a population also has implications for GWA studies, as described in the following section.

Implications for GWA studies

Quantitative traits that reflect the polygenic liability for disorders have practical implications for GWA research.

First, for common disorders, statistical power can be greatly enhanced by conducting GWA studies that compare the low and high extremes of quantitative traits or by studying the entire distribution, rather than dichotomizing the same distribution into cases and controls[39]. Case–control studies of rare disorders already achieve a form of selection of one population extreme by the overrepresentation of cases compared with 'normal' controls; however, as disorders become more common, a quantitative

approach becomes more powerful because the cases become less extreme and the control group becomes increasingly contaminated by 'near cases' — that is, individuals who nearly reach the diagnostic thresholds for the disorder[3].

Second, identifying the quantitative dimensions of complex traits can be beneficial even for dissecting rare disorders, the study of which can be enriched by population cohort studies. The latest GWA studies are collecting genome-wide genotype information on large population cohorts, which often include extensive data on multiple quantitative phenotypes and how they change over time[2]. These data will be invaluable for further exploration of how disorders develop and relate to each other in aetiological terms[7]. Population cohorts with relevant quantitative phenotypes will be a useful complement to the current prevalence of case–control GWA studies (BOX 3).

Third, the increasing emphasis on population cohort studies opens the door for a new wave of phenotyping; the aetiological data already available make these cohorts the ideal sampling frame for new data collection. Now that much of the genotyping has been done, one of the greatest challenges of the next decade will be to add value to our existing population resources. Advances in technology, such as web-based testing[40] and the digitalization of patient records, will allow us to accumulate and coalesce massive amounts of quantitative information on population cohorts. This will lead us towards a holistic understanding of the origins of human disease, thereby completing the work that has already begun in the field of phenomics[41,42].

Perspectives

When several genes affect a disorder, the polygenic liability is continuously distributed as a normal bell-shaped curve, as Fisher noted in 1918. Although qualitative diagnoses of common disorders are the focus of most GWA research, GWA results indicate that multiple genes affect these disorders, which means that what we call common disorders are, in fact, the quantitative extremes of continuous distributions of genetic risk. The obvious test of this hypothesis is that polygenic risk scores will be associated not only with differences between cases and controls but also with individual differences throughout the entire range of variation. For example, genes that are found to be associated with hypertension in case–control studies are predicted to be correlated with the entire range of

variation in blood pressure, just as genes that are found to be associated with obesity in case–control studies are correlated with the entire range of variation in BMI[43].

The problem is that for most disorders, we do not know what the relevant quantitative traits are. Even when we think we know the relevant quantitative traits, we are likely to be surprised, as research on polygenic risk scores considers the full range of normal variation in population cohort studies and captures the diffuse pleiotropic effects of genes. One limitation is measurement: in many domains, more sensitive measures than simply discriminating between cases and controls are needed to assess individual differences throughout the normal distribution. Another limitation is the need for large samples for investigating variation throughout the population, although this constraint will be mitigated by the use of polygenic risk scores, which aggregate the small effect sizes of individual DNA variants.

Clinical implications. Polygenic risk scores are already being used to predict individuals at high polygenic risk, although even in aggregate the effect sizes of associations are not yet large enough to yield predictions that reach the levels of sensitivity and specificity that are required for clinical utility. As mentioned above, there is a more general and novel clinical repercussion: thinking quantitatively about the full range of normal variation, including the neglected 'other end' of the normal distribution of polygenic risk.

We predict that qualitative diagnoses will give way to quantitative dimensions as a consequence of research on polygenic risk. Independently of genetics, this trend towards thinking quantitatively can already be seen in the area of mental disorders, in which new diagnostic procedures include dimensional approaches[44,45], although debates about diagnoses versus dimensions span the entire breadth of medicine[46].

In addition to its impact on diagnosis, research on the polygenic liability for disorders will affect interventions. Most notably, thinking quantitatively leads to a public health model that attempts to evaluate the population's risk quantitatively and to focus on prevention rather than just treating cases[47]. Improved prediction is necessary for effective prevention; primarily this will happen when we study these complex traits as multivariate continuous dimensions, rather than being limited by clinical diagnoses and definitions.

Conclusions

We predict that research on polygenic liabilities will eventually lead to a focus on quantitative dimensions rather than qualitative disorders. The extremes of the distribution are important medically and socially, but we see no scientific advantage in reifying diagnostic constructs that have evolved historically on the basis of symptoms rather than aetiology. A more provocative way to restate our argument is that from the perspective of polygenic liability, there are no common disorders — just the extremes of quantitative traits.

Robert Plomin, Claire M. A. Haworth and Oliver S. P. Davis are at the Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London SE5 8AF, UK.
Correspondence to R.P.
e-mail: Robert.Plomin@kcl.ac.uk

1. Fisher, R. A. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* **52**, 399–433 (1918).
2. Donnelly, P. Progress and challenges in genome-wide association studies in humans. *Nature* **456**, 728–731 (2008).
3. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
4. Maher, B. Personal genomes: the case of the missing heritability. *Nature* **456**, 18–21 (2008).
5. Falconer, D. S. & MacKay, T. F. C. *Introduction to Quantitative Genetics* (Longman, Harlow, 1996).
6. Kruglyak, L. The road to genome-wide association studies. *Nature Rev. Genet.* **9**, 314–318 (2008).
7. McCarthy, M. I. *et al.* Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Rev. Genet.* **9**, 356–369 (2008).
8. MacKay, T. F. C., Stone, E. A. & Ayroles, J. F. The genetics of quantitative traits: challenges and prospects. *Nature Rev. Genet.* **10**, 565–577 (2009).
9. Rung, J. *et al.* Genetic variant near *IRS1* is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. *Nature Genet.* **41**, 1110–1115 (2009).
10. Bouatia-Naji, N. *et al.* A polymorphism within the *G6PC2* gene is associated with fasting plasma glucose levels. *Science* **320**, 1085–1088 (2008).
11. Elliott, P. *et al.* Genetic loci associated with C-reactive protein levels and risk of coronary heart disease. *JAMA* **302**, 37–48 (2009).
12. Podolsky, D. K. Inflammatory bowel disease. *N. Engl. J. Med.* **347**, 417–429 (2002).
13. Mathew, C. G. New links to the pathogenesis of Crohn disease provided by genome-wide association scans. *Nature Rev. Genet.* **9**, 9–14 (2008).
14. Fritz, J. H., Ferrero, R. L., Philpott, D. J. & Girardin, S. E. Nod-like proteins in immunity, inflammation and disease. *Nature Immunol.* **7**, 1250–1257 (2006).
15. Chamaillard, M. *et al.* Gene–environment interaction modulated by allelic heterogeneity in inflammatory diseases. *Proc. Natl Acad. Sci. USA* **100**, 3455–3460 (2003).
16. Ogura, Y. *et al.* A frameshift mutation in *NOD2* associated with susceptibility to Crohn's disease. *Nature* **411**, 603–606 (2001).
17. Neurath, M. F. IL-23: a master regulator in Crohn disease. *Nature Med.* **13**, 26–28 (2007).
18. Xavier, R. J. & Podolsky, D. K. Unravelling the pathogenesis of inflammatory bowel disease. *Nature* **448**, 427–434 (2007).

19. Rioux, J. D. *et al.* Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nature Genet.* **39**, 596–604 (2007).
20. Singh, S. B., Davis, A. S., Taylor, G. A. & Deretic, V. Human IRGM induces autophagy to eliminate intracellular mycobacteria. *Science* **313**, 1438–1441 (2006).
21. Economou, M., Trikalinos, T. A., Loizou, K. T., Tsianos, E. V. & Ioannidis, J. P. A. Differential effects of NOD2 variants on Crohn's disease risk and phenotype in diverse populations: a metaanalysis. *Am. J. Gastroenterol.* **99**, 2393–2404 (2004).
22. Tremelling, M. *et al.* IL23R variation determines susceptibility but not disease phenotype in inflammatory bowel disease. *Gastroenterology* **132**, 1657–1664 (2007).
23. Stratton, M. R. & Rahman, N. The emerging landscape of breast cancer susceptibility. *Nature Genet.* **40**, 17–22 (2008).
24. Barrett, J. C. *et al.* Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nature Genet.* **40**, 955–962 (2008).
25. Goh, K. I. *et al.* The human disease network. *Proc. Natl Acad. Sci. USA* **104**, 8685–8690 (2007).
26. Wray, N. R., Goddard, M. E. & Visscher, P. M. Prediction of individual genetic risk of complex disease. *Curr. Opin. Genet. Dev.* **18**, 257–263 (2008).
27. Pharoah, P. D. P. *et al.* Polygenic susceptibility to breast cancer and implications for prevention. *Nature Genet.* **31**, 33–36 (2002).
28. Khoury, M. J., Yang, Q., Gwinn, M., Little, J. & Flanders, W. D. An epidemiologic assessment of genomic profiling for measuring susceptibility to common diseases and targeting interventions. *Genet. Med.* **6**, 38–47 (2004).
29. Harlaar, N., Butcher, L., Meaburn, E., Craig, I. W. & Plomin, R. A behavioural genomic analysis of DNA markers associated with general cognitive ability in 7-year-olds. *J. Child Psychol. Psychiatry* **46**, 1097–1107 (2005).
30. Morrison, A. C. *et al.* Prediction of coronary heart disease risk using a genetic risk score: the Atherosclerosis Risk in Communities Study. *Am. J. Epidemiol.* **166**, 28–35 (2007).
31. The International Schizophrenia Consortium. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752 (2009).
32. Pharoah, P. D. *et al.* Association between common variation in 120 candidate genes and breast cancer risk. *PLoS Genet.* **3**, e42 (2007).
33. Bare, L. A. *et al.* Five common gene variants identify elevated genetic risk for coronary heart disease. *Genet. Med.* **9**, 682–689 (2007).
34. Maller, J. *et al.* Common variation in three genes, including a noncoding variant in *CFH*, strongly influences risk of age-related macular degeneration. *Nature Genet.* **38**, 1055–1059 (2006).
35. van Hylckama Vlieg, A., Baglin, C. A., Bare, L. A., Rosendaal, F. R. & Baglin, T. P. Proof of principle of potential clinical utility of multiple SNP analysis for prediction of recurrent venous thrombosis. *J. Thromb. Haemost.* **6**, 751–754 (2008).
36. Lyssenko, V. *et al.* Genetic prediction of future type 2 diabetes. *PLoS Med.* **2**, e345 (2005).
37. Nettle, D. The evolution of personality variation in humans and other animals. *Am. Psychol.* **61**, 622–631 (2006).
38. Kitano, H. Systems biology: a brief overview. *Science* **295**, 1662–1664 (2002).
39. Potkin, S. G. *et al.* Genome-wide strategies for discovering genetic influences on cognition and cognitive disorders: methodological considerations. *Cogn. Neuropsychiatry* **14**, 391–418 (2009).
40. Haworth, C. M. A. *et al.* Internet cognitive testing of large samples needed in genetic research. *Twin Res. Hum. Genet.* **10**, 554–563 (2007).
41. Bilder, R. M. *et al.* Phenomics: the systematic study of phenotypes on a genome-wide scale. *Neuroscience* **164**, 30–42 (2009).
42. Freimer, N. & Sabatti, C. The human phenome project. *Nature Genet.* **34**, 15–21 (2003).
43. Walley, A. J., Asher, J. E. & Froguel, P. The genetic contribution to non-syndromic human obesity. *Nature Rev. Genet.* **10**, 431–442 (2009).
44. First, M. B. *et al.* Clinical utility as a criterion for revising psychiatric diagnoses. *Am. J. Psychiatry* **161**, 946–954 (2004).
45. The Psychiatric GWAS Consortium Steering Committee. A framework for interpreting genome-wide association studies of psychiatric disorders. *Mol. Psychiatry* **14**, 10–17 (2009).
46. Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B. & Richardson, W. S. Evidence based medicine: what it is and what it isn't. *BMJ* **312**, 71–72 (1996).
47. Rose, G., Khaw, K. T. & Marmot, M. *Rose's Strategy of Preventive Medicine* (Oxford Univ. Press, 2008).
48. Frayling, T. M. *et al.* A common variant in the *FTO* gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* **316**, 889–894 (2007).
49. Sladek, R. *et al.* A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* **445**, 881–885 (2007).

FURTHER INFORMATION
Nature Reviews Genetics series on Genome-Wide Association Studies:
http://www.nature.com/nrg/series/gwas/index.html

ALL LINKS ARE ACTIVE IN THE ONLINE PDF

Common disorders are quantitative traits

Robert Plomin, Claire M. A. Haworth and Oliver S. P. Davis

An incorrect version of this article was previously published online (publication date 27 October 2009). In the second paragraph of the 'Identifying quantitative mechanisms' section on page 874 in this article, the history of genome-wide association (GWA) studies for type 2 diabetes was incorrectly described and a key reference was omitted. The corrected paragraph is shown below.

The authors apologize for this error.

For some traits, such as type 2 diabetes (T2D), a quantitative approach has already been embraced, with striking results[9]. Although the first T2D GWA studies were case–control studies (REF. 49, and subsequently other studies, for example, REF. 3), a wave of follow-up studies have focused on quantitative traits that are related to T2D, including levels of fasting glucose[10] and C-reactive protein[11], and glucose tolerance[9]. These studies are leading to refinements in the definition of T2D.

Reference 49 has now been added to the reference list.