Heritability: What's the point? What is it not for? A human genetics perspective

Nicolas Robette, Françoise Clerget-Darpoux, Emmanuelle Génin

▶ **To cite this version:**

Nicolas Robette, Françoise Clerget-Darpoux, Emmanuelle Génin. Heritability: What's the point? What is it not for? A human genetics perspective. Genetica, 2022, Online ahead of print. 10.1007/s10709-022-00149-7 . inserm-03598644

HAL Id: inserm-03598644

https://inserm.hal.science/inserm-03598644

Submitted on 5 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Heritability: What's the point? What is it not for? A human genetics perspective

Nicolas Robette[1], Emmanuelle Génin[2], Françoise Clerget-Darpoux[3]

1 – ENSAE-CREST (UMR 9194), Institut Polytechnique de Paris
2 – Inserm, Univ Brest, EFS, CHU Brest, UMR 1078, GGB, Brest, France
3 – Imagine Institute, INSERM UMR1163, Univ Paris Descartes, France

ORCID numbers
Nicolas Robette 0000-0003-3932-5964
Emmanuelle Génin 0000-0003-4117-2813
Françoise Clerget-Darpoux 0000-0003-4816-6619

Key-words
Heritability, polygenic additive model, twin, correlation, causation, association

Address for correspondence
Emmanuelle Génin
UMR 1078 GGB
Faculté de Médecine – IBRBS
22 avenue Camille Desmoulins
F-29238 Brest Cedex 3
France

Abstract

In this paper, we explain the concept of heritability and describe the different methods and the genotype-phenotype correspondences used to estimate heritability in the specific field of human genetics.

Heritability studies are conducted on extremely diverse human traits: quantitative traits (physical, biological, but also cognitive and behavioral measurements) and binary traits (as is the case of most human diseases). Instead of variables such as education and socio-economic status as covariates in genetic studies, they are now the direct object of genetic analysis.

We make a review of the different assumptions underlying heritability estimates and dispute the validity of most of them. Moreover, and maybe more importantly, we show that they are very often misinterpreted. These erroneous interpretations lead to a vision of a genetic determinism of human traits. This vision is currently being widely disseminated not only by the mass media and the mainstream press, but also by the scientific press. We caution against the dangerous implication it has both medically and socially.

Contrarily to the field of animal and plant genetics for which the polygenic model and the concept of heritability revolutionized selection methods, we explain why it does not provide answer in human genetics.

Introduction

Understanding how genes contribute to complex traits has been at the center of many researches in the last century and remains a major question today. This question was at the origin of the field of quantitative genetics that developed, a century ago, after the seminal work of Ronald Fisher. In this work, Fisher (1918) aimed at explaining the observations of Biometricians on trait measure correlations between relatives by the effects of a large number of Mendelian genetic factors. He introduced the concept of variance and its decomposition into a genetic and an environmental component that is at the basis of the concept of heritability. Indeed, genetic and environmental variations both contribute to the phenotypic variation of individuals. The phenotypic correlation between parents and their offspring only depends on the importance of the genetic variation part which is the so-called heritability. To estimate this heritability, it is necessary to specify the correspondence between genotypes and phenotype and this is what Fisher proposed in his work by suggesting that quantitative traits could be explained by a polygenic model. This polygenic model and the related mathematical concept of heritability revolutionized animal and plant selective breeding, previously made on a purely empirical basis. Indeed, the response to selection of a trait was shown to depend on the heritability through what is called the breeder's equation proposed by Lush in 1937 (Lush 1937).

With 20,911 occurrences in articles from PubMed since 1946 (search performed on February 2021), including 6,860 over the last five years, the term "heritability" is now commonly used in many studies although not always correctly. The concept is indeed a statistical concept that is often not well understood and misused. This is especially true in the field of human genetics where several authors have already warned against its misusage (see for example, Feldman and Lewontin 1975). Recently, with the development of novel technologies to characterize genome variations and the possibility to perform association tests at the scale of the entire genome, heritability computations have become even more used in human genetics. Novel methods have been developed to estimate heritability from genetic data in order to determine how much of the phenotypic variation could be explained by the associated loci that were discovered by Genome-Wide Association Studies (GWAS). The game then consisted in comparing these genetically derived estimates against estimates of heritability based on pedigrees or twin studies to determine how much additional efforts were needed to capture all the genetic component of trait variability. The hope was that by finding the genetic factors contributing to this genetic component, it would be possible to make some predictions on expected traits values in individuals.

Birth of the concept of heritability

History and definition

The origin of the word heritability is difficult to trace. It is often attributed to Lush and his book entitled "Animal Breeding Plans" (Lush 1937) is referred. However, the word is not used in the first edition of Lush's book. As explained by Bell (1977), the usage of the term has evolved from an initial stage in the middle of the nineteenth century when it was used to denote "the hereditary transmission of characteristics or material things, simply having the capability (legally of biologically) of being inherited". Then, at the beginning of the twentieth century, Johannsen introduced the terms "gene", "genotype" and "phenotype" and described phenotypic variation as arising from environmental and genotypic fluctuations (see Johannsen 1911). Lush himself suggested that Johannsen was probably the first who captured the idea of heritability. Lush wrote to Bell (see Bell 1977) "Wilhelm Johannsen deserves credit for distinguishing clearly between variance caused by differences among the individuals in their genotypes and variance due to differences in the environments under which they grew. This is close to what I call "heritability in the broad sense". I take it that "Erblichkeit" can be translated fairly as heritability". Although the term is not used in the first edition of his book "Animal Breeding Plan" published in 1937, Lush was the first one to define heritability in its modern-day usage. In the second version of the book, published in 1943, the term "heritability" is used in several places and even appears in the index (Lush 1943).

An important step in the formulation of the concept of heritability was the work of Ronald Fisher and his famous article, "The Correlation between Relatives on the Supposition of Mendelian Inheritance", published in 1918, that marked a decisive step in the development of quantitative genetics. From a mathematical point of view, Fisher proposed to decompose the phenotypic variance *P* (note that Fisher introduced the concept of variance at that time) into the sum of the genotypic variance *G* and the environmental variance *E*:

$$var(P) = var(G) + var(E) \qquad \text{(Eq 1)}$$

Fisher did not use the word "heritability" in his 1918 paper but pointed out the importance of the ratio *var(G)/var(P)* that is precisely what we call the broad-sense heritability H^2. Thus,

$$H^2 = \frac{var(G)}{var(P)} \qquad \text{(Eq 2).}$$

H^2 measures how much of the phenotypic variance is attributable to genotypic variance.

The genotypic variance *G* can be further decomposed into its additive *A*, dominance *D* and epistasis *Ep* components. The ratio of the additive genetic variance (which

corresponds to the addition of the average effects of the two alleles of each genetic locus) over the phenotypic variance is called "narrow-sense heritability", and noted h^2:

$$h^2 = \frac{var(A)}{var(P)} \qquad \text{(Eq 3).}$$

The first uses of heritability for animal and plant selection

The ratio of the variance of the additive genetic effects over the phenotypic variance was shown by Fisher (1918) to be directly related to the correlation of phenotypes between relatives. For the special case of parent-offspring for example, he found that the phenotype correlation is simply ½ h^2 and he derived similar equations for different degrees of kinship. Thus, the narrow-sense heritability measures the resemblance between relatives, and hence is also a way to predict the response to selection.

Indeed, if we denote by *R* the response to selection defined as the difference in mean phenotype between offspring and parental generations, then, it could be shown that *R* depends on h^2. This was first shown by Fisher (1930) when considering the evolution of fitness under natural selection, leading to the so-called "Fundamental Theorem of Natural Selection". The same is also true under artificial selection with *R* being a simple function of h^2 and the selection differential *S* defined as the difference between the phenotype mean in the parental population and the phenotype mean of the population that reproduces (among those that have been selected):

$$R = h^2 S \qquad \text{(Eq 4).}$$

This formula was first derived by Lush (1937) and later called "the breeder's equation" (see Ollivier 2008). It can predict how successful artificial selection will be given the mean value of the selected individuals.

The genotype-phenotype relationship model underlying the concept of heritability

Underlying the concept of heritability is the idea that any quantitative trait can be described as the sum of a genetic and a non-genetic (or environmental) component, and that the genetic component involves a large number of Mendelian factors with additive effects. This model referred to as the "infinitesimal model" or "the polygenic additive model" has its roots in the observations made by Galton (1877), and their analysis by Pearson (1898), followed by the Mendelian interpretation made by Fisher (1918). Under this model, each locus makes an infinitesimal contribution to the genotypic variance and thus this latter variance remains constant over time even when natural selection is occurring.

The infinitesimal model was derived for quantitative traits such as height, weight or any other trait that is measurable and normally distributed in the population. Several human traits and in particular diseases are not quantitative but dichotomous (or binary) - e.g., affected or unaffected by the disease of interest. Some of these dichotomous traits can be explained by simple Mendelian models. It is for example the case for rare diseases that segregate within family. For most common diseases, however, this is not the case and the observed familial aggregation cannot be explained by a simple monogenic model. To study such traits, Carter (1961) proposed the so-called "liability model" that assumes that, underlying these binary traits, there is an unobserved normally distributed quantitative variable, the liability, that measures individual's susceptibility to the disease. When the liability value exceeds a given threshold, the individual is affected and otherwise, the individual is unaffected.

The uses of heritability in human genetics

A short history

Fisher's model was first applied to human measurable physical traits such as height or weight but also to another quantitative trait, intelligence quotient (IQ), using the results of IQ tests. The original test proposed by Binet and Simon as a measure of child's mental age was standardized and published in 1916 to become the standard intelligence test used in the U.S. It is from the end of the 1960s that the heritability of IQ becomes widely debated, in particular after publication of the work of Jensen (1969). Using data on IQ collected in different studies, Jensen estimated that the heritability of IQ was about 80%. He concluded that differences in intelligence between social groups are largely genetic in origin and that educational policies aiming at reducing inequalities would therefore be ineffective. This type of reasoning is notoriously extended in Herrnstein and Murray's *The Bell Curve* (1994) or in the work of Robert Plomin (2018). At the same time, the study of the heritability of IQ was the subject of a great deal of theoretical, methodological, moral and political criticism (see, among others, Kempthorne 1978; Jacquard 1978; Lewontin et al 1984).

Studies of the heritability of cognitive and cultural traits have multiplied since the 1970s, notably through the development of a new research specialty, behavior genetics (see Panofsky 2011, 2014), which aims to study the nature and origin of individual behavioral differences. Heavily invested by psychologists, behavior genetics is interested, for example, in personality traits, social attitudes and mental illnesses. More recently, the study of the heritability of behavioral traits has been taken up in other disciplines of the human sciences, with the development of research currents in criminology ("biosocial criminology", see Larrègue 2016, 2017, 2018a), political science ("genopolitics", see Larrègue, 2018b), economics ("genoeconomics", see Benjamin et al 2012) or sociology

("sociogenomics" or "social science genetics", see Robette 2021), and the study of traits as diverse as delinquency, electoral behavior, income, educational attainment, social mobility or fertility.

Apart from quantitative traits, there have also been considerable interest in estimating the heritability of different common human diseases. Diabetes is an example of such disease that has been at the center of many studies to understand the genetic contribution and estimate heritability (for a review, see Genin and Clerget-Darpoux 2015b). As explained above, for these disease traits that are dichotomous, rather than the heritability of the binary trait, it is the heritability of this liability that is estimated (Falconer 1965).

The evolution of the data used and the estimation method

Human genetic studies on heritability are based on a variety of data and methods that have evolved with scientific and technological advances.

Use of phenotypic covariance between relatives: twin studies

Following on from Fisher (1918), Falconer shows in his book *Introduction to Quantitative Genetics* (1960) that the phenotypic covariance between relatives is a function of the additive variance according to the degree of kinship. The first heritability estimates were derived from empirical data on phenotype correlations between relatives and different approaches were proposed that compare these correlations between different types of relatives (see Tenesa and Haley 2013 for a review). Among the different types of relatives, the most exploitable in human genetics is the comparison of identical and fraternal twins. In 1876, Galton proposed the use of twins to distinguish between genetic and environmental factors in the expression of a trait. But it was not until the beginning of the twentieth century that the idea emerged that there are two kinds of twins: monozygotic (MZ) and dizygotic (DZ). In 1924, Siemens published the first study comparing the similarity between MZ and DZ twins and, in 1929, Holzinger gave formulas for studying the relative effect of nature and nurture upon mean twin differences and their variability.

In 1960, Falconer showed that heritability may be simply estimated from the difference between MZ and DZ concordance rates, provided the following assumptions hold:

1) environmental variance is identical for MZ and DZ twins and remains the same throughout life,

2) variance of interaction could be neglected.

Falconer first applied the method to quantitative traits such as height, weight and IQ using data on 50 MZ and DZ pairs from Newman et al (1937). Shortly afterwards,

Falconer (1965) proposed to calculate the heritability of diseases that do not have a simple monogenic determinism by relying on the polygenic additive model for liability.

Christian et al (1974) recalls that in practice, heritability estimates cannot be done without simplifying assumptions : the most common of which are: (1) the effect of environmental influences on the trait are similar for the two types of twins; (2) hereditary and environmental influences are neither correlated in the same individual nor between members of a twin set; (3) there is no correlation between parents due to assortative mating; and (4) the trait in question is continuously distributed with no dominance and no epistatic effect (narrow sense heritability).

Twin studies are widely used because of the wide availability of data from twin registries (see van Dongen et al 2012 for an inventory). Controversies about the heritability of intelligence have largely been based on twin studies since the work of Holzinger (1929). Social scientists have applied - and still apply - the same methods to many behavioural traits. To cite just two recent examples, using data on twin girls in the UK, Tropf et al (2015) find that 26% of the variation in age at first child is explained by genetic predisposition, 14% by the environment shared by siblings, and 60% by the non-shared environment or measurement error. Baier and Van Winkle (2020) find that the heritability of school performance is lower for children of separated parents and conclude that educational policies could specifically target children of separated parents to help them fulfill their genetic potential.

Twin studies are also ubiquitous in the field of disease heritability. It is impossible to give even a very partial view of this work in the context of this article. However, we can illustrate it by mentioning the work on schizophrenia, where a meta-analysis concludes that the heritability of the liability is estimated to be around 80% (Sullivan et al 2003). Some authors even argue that twin studies are also "valuable for investigating the etiological relationships between schizophrenia and other disorders, and the genetic basis of clinical heterogeneity within schizophrenia." (Cardno and Gottesman 2000). Similarly, a meta-analysis of a sample of 34,166 twin pairs from the International Twin Registers concluded that the heritability of the liability of type II diabetes was 72% (Willemsen et al 2015).

From correlation between relatives to the use of genetic markers

With the development of molecular techniques to detect variations in the human genome, methods were developed to gain information on individual relatedness from observed genotypes at genetic markers. Building on this idea, Ritland (1996,2000) first proposed a marker-based method for estimating heritability of quantitative traits from genetic and phenotypic data on individuals of unknown relationship. The method however could not really be used at that time as there were not enough genetic markers covering the human genome. After the first sequence of the human genome was released (McPherson et al

2001), efforts were put on characterizing the distribution of common genetic variants in different worldwide populations, their frequency and correlation patterns - see Hapmap project launched in 2002 (Couzin 2002). Hundreds of thousands of Single Nucleotide Polymorphisms (SNP) spanning the entire human genome were characterized and the first SNP-arrays were developed to easily genotype them. These SNP-arrays were used in large samples of cases and controls to perform Genome-Wide Association Studies (GWAS) and to find the genetic risk factors involved in common diseases. Using these GWAS data, it was possible to directly quantify the contribution of genetic variants to phenotypic variance. This was first done by Visscher et al (2006) for the case of sibpairs where instead of using the expected identity-by-descent (IBD) sharing, they used the observed IBD sharing among the sibs to estimate the heritability of height. The method was further extended to allow estimation of heritability from unrelated individuals by measuring the proportion of phenotype variance that can be explained by a linear regression on the set of genetic markers found significantly associated with the phenotype and used as explanatory variables. This was done for human height after the first GWAS were performed. The proportion of human height variance explained by the 54 genome-wide significant SNPs however, was very low, around 5%, much smaller than the 80% heritability found from family and twin studies (see Visscher 2008). There was thus a problem of "missing heritability" (Maher 2008; Manolio et al 2009) that led investigators to suggest different explanations for this missing heritability (see Genin 2020 for a review). Among these different explanations was the possibility that there were some other loci contributing to the heritability than the GWAS top signals. This led Yang et al. (2010) to propose to estimate heritability based on the information from all SNPs present on the SNP-array using a mixed linear model where SNP effects are treated as random variable. This method was latter implemented in the GTCA software (Yang et al. 2011) and referred to as Genomic Relatedness matrix restricted Maximum Likelihood (GREML). It was then extended to estimate heritability for dichotomous traits (Lee at al 2011) and extensively used to estimate the so-called SNP-based heritability of many different common diseases (see Yang et al. 2017 for a review of the concept of SNP-based heritability and its estimation methods).

Errors of interpretation and limitations of the models and methods

While the concept of heritability may seem relatively simple, in practice it is subject of misuse and misinterpretation. It also relies on strong assumptions that are not always met in natural populations.

Errors of interpretation of heritability

First, the terminology itself is misleading. Indeed, as discussed by Stoltenberg (1997), terms such as "heritable", "inherited" or "heredity" have folk meanings that are different

from the scientific notions they are supposed to represent. This contributes to misinterpretations. In particular, the term "heritability" is used in the common language as a synonym of "inheritance", with the idea that something heritable is something that passed from parents to offspring. As we have seen, heritability is not an individual characteristic but a population measure. It does not tell anything on the genetic determinism of the trait under study. To avoid any ambiguity and insist on the fact that the primary use of heritability estimates is to predict the results of selective breeding, Stoltenberg (1997) suggested replacing the term "heritability" by "selectability".

Second, there is often a confusion between the contribution of genetic factors to the phenotype and their contribution to the phenotype variability. Heritability says nothing about the causes, the mechanisms at the origin of differences between populations, nor about the etiology of diseases. As Lewontin (1974) reminded us, there is a crucial distinction between the analysis of variance and the analysis of causes. A strong heritability does not mean that the main factors involved in the trait are genetic factors. In a population where there is no environmental variability, the heritability is 100%. Similarly, in a homogeneous social environment, heritability estimates may be high for traits which are mainly due to social environmental factors. This error of interpretation is also prevalent in the literature on GWAS and the discussion around the so-called missing heritability. It is present in the famous paper by Manolio et al (2009) when they list a number of diseases, for which the proportion of heritability currently explained by the loci detected by GWAS is low and conclude that other relevant genetic loci remain to be detected (see discussion by Vieneis & Pearce (2011)).

Third, heritability is often reported as if it were a universal measure for the trait under study. This is wrong as heritability is a local measure in space and time, specific to the studied population. Two groups of individuals, with exactly the same genetic background, will have, for a given trait, a different heritability according to whether they are placed in a context where the environment is constant or variable. Heritability can also vary through time with environmental changes. Differences in heritability can be found depending on the age of the individuals under study. This is well illustrated for Body Mass Index (BMI) with estimates that are systematically larger in children than in adults and also larger when derived from twin studies than from family studies (Elk et al 2012). Note however that in this latter meta-analysis of BMI heritability studies, estimates were found to vary of almost two folds, ranging from 0.47 to 0.90 in twin studies and from 0.24 to 0.81 in family studies. It clearly shows that the measure is not universal and does not have much utility in human populations. It also seriously questioned the missing heritability problem that makes the underlying assumption that heritability should remain the same for a given trait whatever the population context and the sample on which it is measured.

Validity of the assumptions underlying heritability estimates

Heritability estimates rely on strong assumptions that could not be tested and are disputable in human genetics.

A first assumption that is inherent in the polygenic additive model is the existence of many genetic and environmental factors that each have a small contribution. It is assumed that there are no single genetic or environmental factor that makes a major contribution. This is not true for many diseases where major genetic and/or environmental factors have been found. Examples include the contribution of specific HLA heterodimers in celiac disease or other autoimmune disease and diet and physical activity in obesity and type-2 diabetes (for a review on the limit of these assumptions in the context of diabetes, see Génin and Clerget-Darpoux 2015b). For a trait such as IQ, even if Herrnstein & Murray (1994) suggested a limited malleability by schooling, it is now well recognized that school attendance plays an important role. Different studies have shown that education has a direct role on IQ and that it is not reverse causation due, for example, to the fact that people with higher IQ tend to have higher school attendance (see for example the study by Brinch & Galloway (2012) where compulsory schooling in Norway in the 1960s was found to have an effect on IQ scores of early adult men.

A second assumption is the absence of interaction between genetic and environmental factors. It would mean that genetic variance could be estimated without any knowledge on the environment. Yet, contemporary biology has demonstrated that traits are the product of interactions between genetic and non-genetic factors at every point of the development (Moore and Shenk 2017). Genes are part of a "developmental system" (Gottlieb 2001). Moreover, epigenetic phenomena - imprinted genes, methylation, etc. - cannot be ignored. As outlined by Burt (2015) : "the conceptual (biological) model on which heritability studies depend—that of identifiable separate effects of genes vs. the environment on phenotype variance—is unsound".

Another assumption is that of random environment. For most human behavioral traits, this hypothesis is of course not valid (Vetta et Courgeau 2003; Courgeau 2017). Parents pass on alleles to their children with whom they also share environmental factors that can be involved in studied traits, leading to some "co-transmission" of genetic and environmental factors. This is for example the case of educational level for cognitive traits or dietary habits for traits linked to BMI. As shown by Cavalli-Sforza and Feldman (1973), ignoring the co-transmission of genetic and environmental factors could lead to strong bias in heritability estimates. Vertical cultural transmission has a profound effect on correlations between relatives and this effect can be misinterpreted as being due to genetic variation.

The additivity assumption is also not relevant both at the level of the alleles within a genotype but also between genes. Indeed, for many traits and in particular diseases, there exist dominance effects as well as epistasis. The effect of a genotype on phenotype often

depends on the genetic background and on the genotypes at other loci (Carlborg and Haley 2004; Mackay and Moore 2014).

Another underlying assumption is random mating and Hardy-Weinberg equilibrium that is not true, especially for cognitive and cultural traits where homogamy is often the rule (Courgeau, 2017).

Besides these assumptions that are inherent to the model underlying heritability and thus to all the methods to estimate heritability in twin studies, it is further assumed that environment is similarly shared between monozygotic and dizygotic twins. This equal sharing of environment is probably the most debated hypothesis. Since the 1960s, empirical evidence has accumulated that monozygotic twins live in more similar social environments than dizygotic twins. For example, they are more likely to be treated the same by their parents, have the same friends, be in the same class, spend time together, be more attached to each other through their whole life, etc. (Joseph 2013; Burt and Simons 2014). Furthermore, the prenatal (intrauterine) environment of monozygotic and dizygotic twins is different: the prenatal environments of MZ twins (who often share the same placenta) are more similar than those of DZ twins (who never share the same placenta). Most advocates of twin studies recognize that the environments of MZ twins are more similar than those of DZ twins. However, they suggest that, for the model to remain valid, it is only necessary that the environmental factors directly related to the trait under study are the same in MZ and DZ twins ("trait-relevant equal environment assumption"). In doing so they divert potential criticism on the very strong hypothesis of equal environment.

Finally, we see that none of the hypotheses inherent in heritability estimates are verified in humans. More fundamentally, if heritability is used routinely and usefully for plant and animal breeding (to predict the effectiveness of this selection), it is in the context of experimental devices that allow the environment to be controlled, which is impossible in nature and in the case of humans.

From heritability to risk prediction: polygenic risk scores

Besides the questions raised on the usefulness of heritability estimates in human genetics, they continue to be reported in many articles published in major journals and they are sometimes even required by reviewers and/or editors. Even worst, in 2007, a new impetus was given to heritability estimates. It was proposed to extract the genetic variability from the phenotypic variability with, at the end of the day, a phenotypic prediction.

Wray et al (2007) promoted a new predictive tool for clinicians: the polygenic risk score (PRS). They proposed the use of SNP associations to estimate the genetic variance of multifactorial diseases. The underlying idea was that each SNP association reflects a

genetic risk variability in its neighborhood. Assuming an intrinsic value of the disease heritability under study, they claimed that the best predictor of the disease is obtained when all genetic variations, i.e., the whole heritability, is captured and pooled. This triggered a long search for "missing heritability" (Manolio et al 2009), together with an endless extension of GWAS sample sizes. Different methods were developed to better estimate heritability from genome-wide association studies (see Speed, Holmes, Balding, 2020 for a review): Lasso (Tibshirani, 1996), Ridge regression (Meuwissen et al 2001), Bayesian mixed -model (Loh et al, 2015, Moser et al, 2015), and MegaPRS, a summary statistic that allows the user to specify the heritability model (Zhang et al. 2020). The number of papers praising the benefits of using PRS for different complex diseases has grown exponentially during the last decade (1,684 results on Pubmed when searching for "Polygenic Risk Score(s)" on April 28, 2021). Software applications computing individual PRS for numerous diseases and intended to help clinical decision, were also developed.

However, all the limits given for heritability estimation also apply to PRS estimation. The validity of PRS estimations depends on the validity of the polygenic additive liability under which they are computed. Let us just recall that adopting this simple model implies that the genetic variance can be extracted from the phenotypic variance without any knowledge on the environment. The shift from genetics to genomics and from family studies to population studies has led to a shift in causal inference. Once again, association does not mean causation. Indeed, associations may reflect the indirect effect of associated traits or environmental factors. For example, SNPs associated to breast cancer may include SNPs associated to BMI, some of them reflecting environmental factors and socio-cultural stratification. As underlined by Janssens (2019), PRS are not independent of observed clinical factors, they contain indirect information on clinical, familial and environmental factors. Associations can also be due to the effect of ubiquitous genes, such as regulator or transcriptor genes having trans effect on causal genes (Boyle et al 2017).

Besides, like heritability, PRS are not universal measures. Even in a population considered as genetically homogeneous, such as the UK biobank, a simple change on a variable such as age, sex, socio-economic status may impact PRS (Mostafavi et al 2020; Abdellaoui et al 2021).

Nevertheless, the most important misconception comes, here again, from the confusion between *gene effect on the trait* and *effect of gene polymorphism on the trait variance*. Many consider that the polygenic risk score of an individual for a given disease represents the genetic part of his liability. In fact, the PRS curve measures the *relative risk* of each genotype compared to the genotype without any risk allele (often extremely rare and unobserved when the number of SNPs including in the PRS is large) whereas the liability curve is meant to measure *absolute risk*. Some researchers have claimed that the risks of high PRS individuals are equivalent to the one of rare disease mutation carriers (Khera et al 2018). This statement is wrong because the absolute risks are very different

in the two situations. Even for high PRS value, the absolute risk may be very low. This makes PRS a very poor predictor test in terms of specificity and sensitivity, as shown and well-illustrated on coronary artery disease by Wald and Old (2019).

In a recent paper, Wray et al (2019) pointed out the similarity between the Estimated Breeding Value (EBV) in livestock and PRS in human. In their comparison, they considered that both PRS and EBV are estimates of the additive genetic value of a trait. They explained why genetic variance is easier to estimate in a livestock than in humans and called for an increase in GWAS sample sizes to maximize the accuracy of PRS. However, they did not raise the most important and critical points:
- the validity of the assumptions made for human diseases or traits in PRS estimation.
- the simplistic interpretation of SNP associations as reflecting the effect of genetic factors
- the different potential uses of EBV and PRS estimates. While EBV offers reliable classification to predict a global improvement of a trait in the next generation, PRS classification could not be used to predict an individual disease risk.

Conclusion

Acknowledging the limitations of the notion of heritability, many authors have pointed out the impasse it constitutes for human genetics. Lewontin suggested as early as 1974 "to stop the endless search for better methods of estimating useless quantities" (Lewontin 1974), while Jacquard emphasized in 1978 that "the complexity of the mathematics used to answer is not enough to give meaning to an absurd question [...], devoid of any meaning" (Jacquard 1978). The centrality of heritability in human genetics research more than forty years later could thus be a manifestation of the "Garbage In, Garbage Out" syndrome (Génin and Clerget-Darpoux 2015a).

There are growing voices calling for revising and moving beyond the polygenic additive model (Nelson et al 2013; Génin and Clerget-Darpoux 2015b). The "omnigenic model" of Boyle et al (2017) stands as a proposal in this direction. The authors build on the observations that, in genome-wide association studies, statistical associations between genetic variants and disease identify a large number of genes scattered throughout the genome, including many genes with no obvious link to disease. This is in contrast to the expectation that causal variants would be clustered in major disease-related pathways. Boyle et al suggest that gene regulatory networks are so interconnected that all genes are likely to influence the functions of the core disease genes. Thus, a distinction is made between regulatory genes and core genes. But above all, according to the omnigenic model, most of the heritability is explained by the effect of genes located outside the central pathways.

Confusion between measuring a genetic effect on the trait or on its variance is the basis of erroneous interpretations of heritability and PRS. In many articles, PRS is confused with genetic liability and average PRS with average liability. This resulted in the development of a very deterministic view of human traits with false and dangerous predictions of our medical and social future.

In animal or plant populations, where crossbreeding and environment can be controlled, information on genetic variance is central to improving a trait from one generation to the next. Fortunately, human populations are not subject to these same constraints and the objectives of geneticists are totally different. Regarding diseases of complex aetiology, geneticists seek to identify the responsible factors and to understand the complex and heterogeneous interactions between these factors. There is a huge gap between observing associations in a population and understanding the role of genes in the disease development process (Bourgain et al 2007). Sticking to a very simplistic model for all diseases will not allow to reach this goal.

Declarations

Nothing to declare

References

Abdellaoui A, Verweij KJH, Nivard MG (2021) Geographic Confounding in Genome-Wide Association Studies. bioRxiv 2021.03.18.435971. https://doi.org/10.1101/2021.03.18.435971

Baier T, Winkle ZV (2020) Does Parental Separation Lower Genetic Influences on Children's School Performance? Journal of Marriage and Family. https://doi.org/10.1111/jomf.12730

Bell AE (1977) Heritability in retrospect. Journal of Heredity 68:297–300

Benjamin DJ, Cesarini D, Chabris CF, et al (2012) The Promises and Pitfalls of Genoeconomics. Annual Review of Economics 4:627–662. https://doi.org/10.1146/annurev-economics-080511-110939

Binet A, Simon T (1916) New methods for the diagnosis of the intellectual level of subnormals (L'Année Psych., 1905, pp. 191-244). In: Binet A, Simon T, The development of intelligence in children (The Binet-Simon Scale), Williams & Wilkins Co, pp 37–90. https://doi.org/10.1037/11069-002

Bourgain C, Génin E, Cox N, Clerget-Darpoux F (2007) Are genome-wide association studies all that we need to dissect the genetic component of complex human diseases? European Journal of Human Genetics 15:260-3. doi: 10.1038/sj.ejhg.5201753.

Boyle EA, Li YI, Pritchard JK (2017) An Expanded View of Complex Traits: From Polygenic to Omnigenic. Cell 169:1177–1186. https://doi.org/10.1016/j.cell.2017.05.038

Brinch NC, Galloway TA (2012). Schooling in adolescence raises IQ scores. Proc Natl Acad Sci U S A. 109(2): 425–430. doi: 10.1073/pnas.1106077109.

Burt CH (2015) Heritability Studies: Methodological Flaws, Invalidated Dogmas, and Changing Paradigms. Advances in Medical Sociology 16. https://doi.org/10.1108/S1057-629020150000016002

Burt CH, Simons RL (2014) Pulling back the curtain on heritability studies: Biosocial criminology in the postgenomic era. Criminology 52:223–262

Cardno AG, Gottesman II (2000) Twin studies of schizophrenia: From bow-and-arrow concordances to Star Wars Mx and functional genomics. American Journal of Medical Genetics 97:12–17. https://doi.org/10.1002/(SICI)1096-8628(200021)97:1<12::AID-AJMG3>3.0.CO;2-U

Carlborg Ö, Haley CS (2004) Epistasis: too often neglected in complex trait studies? Nature Reviews Genetics 5:618–625. https://doi.org/10.1038/nrg1407

Carter CO (1961) The inheritance of congenital pyloric stenosis. British Medical Bulletin 17:251–253. https://doi.org/10.1093/oxfordjournals.bmb.a069918

Cavalli-Sforza LL, Feldman MW (1973) Cultural versus biological inheritance: phenotypic transmission from parents to children. (A theory of the effect of parental phenotypes on children's phenotypes). Am J Hum Genet 25:618–637

Christian JC, Kang KW, Norton JJ (1974) Choice of an estimate of genetic variance from twin data. Am J Hum Genet 26:154–161

Courgeau D (2017) La génétique du comportement peut-elle améliorer la démographie ? Revue d'études des Populations 2:17

Couzin J (2002) HapMap launched with pledges of $100 million - ProQuest. Science 298:941–942

Elks CE, Den Hoed M, Zhao JH, et al (2012) Variability in the Heritability of Body Mass Index: A Systematic Review and Meta-Regression. Front Endocrinol 3:. https://doi.org/10.3389/fendo.2012.00029

Falconer DS (1965) The inheritance of liability to certain diseases, estimated from the incidence among relatives. Annals of Human Genetics 29:51–76. https://doi.org/10.1111/j.1469-1809.1965.tb00500.x

Falconer DS (1960) Introduction to quantitative genetics. Pearson Education India

Feldman MW, Lewontin RC (1975) The Heritability Hang-up. Science 190:1163–1168

Fisher RA (1918) The Correlation between Relatives on the Supposition of Mendelian Inheritance. Earth and Environmental Science Transactions of The Royal Society of Edinburgh 52:399–433. https://doi.org/10.1017/S0080456800012163

Galton F (1876) The History of Twins, as a Criterion of the Relative Powers of Nature and Nurture. The Journal of the Anthropological Institute of Great Britain and Ireland 5:391–406. https://doi.org/10.2307/2840900

Galton F (1877) Typical Laws of Heredity. Nature 15:492–495. https://doi.org/10.1038/015492a0

Génin E (2020) Missing heritability of complex diseases: case solved? Hum Genet 139:103–113. https://doi.org/10.1007/s00439-019-02034-4

Génin E, Clerget-Darpoux F (2015a) Revisiting the Polygenic Additive Liability Model through the Example of Diabetes Mellitus. HHE 80:171–177. https://doi.org/10.1159/000447683

Génin E, Clerget-Darpoux F (2015b) The missing heritability paradigm: a dramatic resurgence of the GIGO syndrome in genetics. Human heredity 79:1

Gottlieb G (2001) Genetics and Development. In: Smelser NJ, Baltes PB (eds) International Encyclopedia of the Social and Behavioral Sciences, Genetics, Behavior, and Society. Elsevier, New-York, pp 6121–6127.

Herrnstein RJ, Murray CA (1994) The bell curve: Intelligence and class structure in American life. Free Press, New York, NY, US

Holzinger KJ (1929) The relative effect of nature and nurture influences on twin differences. Journal of Educational Psychology 20:241–248. https://doi.org/10.1037/h0072484

Jacquard A (1978) L'inné et l'acquis: l'homme à la merci de l'homme. Journal de la société française de statistique 119:234–251

Janssens ACJW (2019) Validity of polygenic risk scores: are we measuring what we think we are? Human Molecular Genetics 28:R143–R150. https://doi.org/10.1093/hmg/ddz205

Jensen A (1969) How Much Can We Boost IQ and Scholastic Achievement. Harvard Educational Review 39:1–123. https://doi.org/10.17763/haer.39.1.l3u15956627424k7

Johannsen W (1911) The Genotype Conception of Heredity. The American Naturalist 45:129–159. https://doi.org/10.1086/279202

Joseph J (2013) The Use of the Classical Twin Method in the Social and Behavioral Sciences: The Fallacy Continues. The Journal of Mind and Behavior 34:1–39

Kempthorne O (1978) A Biometrics Invited Paper: Logical, Epistemological and Statistical Aspects of Nature-Nurture Data Interpretation. Biometrics 34:1–23. https://doi.org/10.2307/2529584

Khera AV, Chaffin M, Aragam KG, et al (2018) Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nature Genetics 50:1219–1224. https://doi.org/10.1038/s41588-018-0183-z

Larregue J (2016) Sociologie d'une spécialité scientifique. Les désaccords entre les chercheurs 'pro-génétique'et 'pro-environnement'dans la criminologie biosociale états-unienne. Champ pénal/Penal field 13:

Larregue J (2017) La criminologie biosociale à l'aune de la théorie du champ. Ressources et stratégies d'un courant dominé de la criminologie états-unienne. Déviance et société 41:167–201

Larregue J (2018a) «C'est génétique»: ce que les twin studies font dire aux sciences sociales. Sociologie

Larregue J (2018b) «Une bombe dans la discipline»: l'émergence du mouvement génopolitique en science politique. Social Science Information 57:159–195

Lee SH, Wray NR, Goddard ME, Visscher PM (2011) Estimating Missing Heritability for Disease from Genome-wide Association Studies. The American Journal of Human Genetics 88:294–305. https://doi.org/10.1016/j.ajhg.2011.02.002

Lewontin RC (1974) Annotation: the analysis of variance and the analysis of causes. American journal of human genetics 26:400–411

Lewontin RC, Rose S, Kamin LJ (1984) Not in our genes. Pantheon Books, New York

Loh P-R, Tucker G, Bulik-Sullivan BK, et al (2015) Efficient Bayesian mixed-model analysis increases association power in large cohorts. Nature Genetics 47:284–290. https://doi.org/10.1038/ng.3190

Lush JL (1937) Animal breeding plans. Ames. Collegiate Press, Inc

Lush JL (1943) Animal breeding plans., 2nd edn. Iowa State College Press

Mackay TF, Moore JH (2014) Why epistasis is important for tackling complex human disease genetics. Genome Med 6:42. https://doi.org/10.1186/gm561

Maher B (2008) Personal genomes: The case of the missing heritability. Nature 456:18–21. https://doi.org/10.1038/456018a

Manolio TA, Collins FS, Cox NJ, et al (2009) Finding the missing heritability of complex diseases. Nature 461:747–753

McPherson JD, Marra M, Hillier LD, et al (2001) A physical map of the human genome. Nature 409:934–941. https://doi.org/10.1038/35057157

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. Genetics 157:1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Moore DS, Shenk D (2017) The heritability fallacy. Wiley Interdisciplinary Reviews: Cognitive Science 8:e1400

Moser G, Lee SH, Hayes BJ, et al (2015) Simultaneous Discovery, Estimation and Prediction Analysis of Complex Traits Using a Bayesian Mixture Model. PLOS Genetics 11:e1004969. https://doi.org/10.1371/journal.pgen.1004969

Mostafavi H, Harpak A, Agarwal I, et al (2020) Variable prediction accuracy of polygenic scores within an ancestry group. eLife 9:e48376. https://doi.org/10.7554/eLife.48376

Nelson RM, Pettersson ME, Carlborg Ö (2013) A century after Fisher: time for a new paradigm in quantitative genetics. Trends in Genetics 29:669–676

Newman HH, Freeman FN, Holzinger KJ (1937) Twins: a study of heredity and environment. Univ. Chicago Press, Oxford, England

Ollivier L (2008) Jay Lush - reflections on the past. In: Lohmann Breeders. https://lohmann-breeders.com/fr/jay-lush-reflections-on-the-past/. Accessed 19 Apr 2021

Panofsky A (2014) Misbehaving science: Controversy and the development of behavior genetics. University of Chicago Press

Panofsky AL (2011) Field Analysis and Interdisciplinary Science: Scientific Capital Exchange in Behavior Genetics. Minerva 49:295. https://doi.org/10.1007/s11024-011-9175-1

Pearson K (1898) Mathematical contributions to the theory of evolution, On the law of ancestral heredity. Proceedings of the Royal Society of London 386–412. https://doi.org/10.1098/rspl.1897.0128

Ritland K (1996) A Marker-Based Method for Inferences About Quantitative Inheritance in Natural Populations. Evolution 50:1062–1073. https://doi.org/10.1111/j.1558-5646.1996.tb02347.x

Ritland K (2000) Marker-inferred relatedness as a tool for detecting heritability in nature. Molecular Ecology 9:1195–1204. https://doi.org/10.1046/j.1365-294x.2000.00971.x

Robette N (2021) L'hérédité manquante. Promesses et impasses de la sociogénomique. Population. In revision

Siemens DZ (1924) Zwillingspathologie: IhreBedeutung; ihre Methodik, ihre bisherigen Ergebnisse [Twin pathology: Its meaning; its method; results sofar]. Springer Verlag, Berlin, Germany

Speed D, Holmes J, Balding DJ (2020) Evaluating and improving heritability models using summary statistics. Nature Genetics 52:458–462. https://doi.org/10.1038/s41588-020-0600-y

Stoltenberg SF (1997) Coming to terms with heritability. Genetica 99:89–96

Sullivan PF, Kendler KS, Neale MC (2003) Schizophrenia as a Complex Trait: Evidence From a Meta-analysis of Twin Studies. Archives of General Psychiatry 60:1187–1192. https://doi.org/10.1001/archpsyc.60.12.1187

Tenesa A, Haley CS (2013) The heritability of human disease: estimation, uses and abuses. Nature Reviews Genetics 14:139–149. https://doi.org/10.1038/nrg3377

Tibshirani R (1996) Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological) 58:267–288

Tropf FC, Barban N, Mills MC, et al (2015) Genetic influence on age at first birth of female twins born in the UK, 1919–68. Population studies 69:129–145

van Dongen J, Slagboom PE, Draisma HHM, et al (2012) The continuing value of twin studies in the omics era. Nature Reviews Genetics 13:640–653. https://doi.org/10.1038/nrg3243

Vetta A, Courgeau D (2003) Comportements démographiques et génétique du comportement. Population Vol. 58:457–488

Vineis P, E.Pearce N (2011) Genome-wide association studies may be misinterpreted: genes versus heritability. Carcinogenesis 32:1295–1298. https://doi.org/10.1093/carcin/bgr087

Visscher PM (2008) Sizing up human height variation. Nature Genetics 40:489–490. https://doi.org/10.1038/ng0508-489

Visscher PM, Medland SE, Ferreira MAR, et al (2006) Assumption-Free Estimation of Heritability from Genome-Wide Identity-by-Descent Sharing between Full Siblings. PLOS Genetics 2:e41. https://doi.org/10.1371/journal.pgen.0020041

Wald NJ, Old R (2019) The illusion of polygenic disease risk prediction. Genetics in Medicine 21:1705–1707. https://doi.org/10.1038/s41436-018-0418-5

Willemsen G, Ward KJ, Bell CG, et al (2015) The Concordance and Heritability of Type 2 Diabetes in 34,166 Twin Pairs From International Twin Registers: The Discordant Twin (DISCOTWIN) Consortium. Twin Research and Human Genetics 18:762–771. https://doi.org/10.1017/thg.2015.83

Wray NR, Goddard ME, Visscher PM (2007) Prediction of individual genetic risk to disease from genome-wide association studies. Genome Res 17:1520–1528. https://doi.org/10.1101/gr.6665407

Wray NR, Kemper KE, Hayes BJ, et al (2019) Complex Trait Prediction from Genome Data: Contrasting EBV in Livestock to PRS in Humans: Genomic Prediction. Genetics 211:1131–1141. https://doi.org/10.1534/genetics.119.301859

Yang J, Benyamin B, McEvoy BP, et al (2010) Common SNPs explain a large proportion of the heritability for human height. Nature Genetics 42:565–569. https://doi.org/10.1038/ng.608

Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. The American Journal of Human Genetics 88:76–82

Yang J, Zeng J, Goddard ME, et al (2017) Concepts, estimation and interpretation of SNP-based heritability. Nature Genetics 49:1304–1310. https://doi.org/10.1038/ng.3941

Zhang Q, Privé F, Vilhjálmsson B, Speed D (2020) Improved genetic prediction of complex traits from individual-level data or summary statistics. bioRxiv 2020.08.24.265280. https://doi.org/10.1101/2020.08.24.265280