









Maximizing the value of twin studies in health and behaviour

Received: 20 December 2022

Accepted: 19 April 2023

Published online: 15 May 2023

 Check for updates

Fiona A. Hagenbeek ^{1,2,10} ✉, Jana S. Hirzinger ^{1,3,10}, Sophie Breunig ^{1,4,5,10},
Susanne Bruins ^{1,2}, Dmitry V. Kuznetsov ^{1,6}, Kirsten Schut ^{1,7},
Veronika V. Odintsova ^{1,8,9} & Dorret I. Boomsma ^{1,2,8} ✉

In the classical twin design, researchers compare trait resemblance in cohorts of identical and non-identical twins to understand how genetic and environmental factors correlate with resemblance in behaviour and other phenotypes. The twin design is also a valuable tool for studying causality, intergenerational transmission, and gene–environment correlation and interaction. Here we review recent developments in twin studies, recent results from twin studies of new phenotypes and recent insights into twinning. We ask whether the results of existing twin studies are representative of the general population and of global diversity, and we conclude that stronger efforts to increase representativeness are needed. We provide an updated overview of twin concordance and discordance for major diseases and mental disorders, which conveys a crucial message: genetic influences are not as deterministic as many believe. This has important implications for public understanding of genetic risk prediction tools, as the accuracy of genetic predictions can never exceed identical twin concordance rates.

Behavioural scientists are often interested in understanding how human traits, such as intelligence, educational attainment, depression and anxiety, are influenced by inherited and environmental factors. Since long before the advent of human genomics, researchers have relied on twin studies to answer these questions. The classical twin design compares trait resemblances in monozygotic (MZ) twins to those in dizygotic (DZ) twins. Because MZ twins are genetically identical and DZ twins share 50% of their genes on average, any additional similarity between identical twins should be related to genes, provided that the twins share the same environments^{1,2}.

Researchers can apply the classical twin design to estimate the influence of genes for any trait they can measure. They can also extend the classical twin design to model causal relations between exposures

and traits, to describe differences within discordant twin pairs, to estimate genetic and cultural components of inheritance, and to account for measured genetic variables and environmental exposures.

In this Review, we examine the use of twin studies as researchers become interested in new phenotypes, including those based on omics technologies. We reflect on methods for achieving progress in fields that have traditionally valued twin studies of individual differences, and we explore which fields offer new opportunities for twin research and how insights from other disciplines can inform twin studies. We begin by introducing the classical twin design. We then summarize concordance and discordance in twin pairs for major psychiatric and somatic disorders, arguing that no other design is as informative for studying penetrance and genetic prediction of disease risk. Next, we

¹Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands. ²Amsterdam Public Health Research Institute, Amsterdam, the Netherlands. ³Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Amsterdam Neuroscience, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands. ⁴Department of Psychology & Neuroscience, University of Colorado Boulder, Boulder, CO, USA. ⁵Institute for Behavioral Genetics, University of Colorado Boulder, Boulder, CO, USA. ⁶Faculty of Sociology, Bielefeld University, Bielefeld, Germany. ⁷Nightingale Health Plc, Helsinki, Finland. ⁸Amsterdam Reproduction & Development (AR&D) Research Institute, Amsterdam, the Netherlands. ⁹Department of Psychiatry, University Medical Center of Groningen, University of Groningen, Groningen, the Netherlands. ¹⁰These authors contributed equally: Fiona A. Hagenbeek, Jana S. Hirzinger, Sophie Breunig. ✉ e-mail: f.a.hagenbeek@vu.nl; di.boomsma@vu.nl

BOX 1

Glossary

Amnion. A membrane that contains the amniotic fluid and provides a protective environment for an embryo by closely covering it. Forms the amniotic sac together with the chorion.

Assortative mating. Individuals with similar phenotypic characteristics, such as height or intelligence, are more likely to mate with each other than would be expected by chance. This impacts heritability estimation in the classical twin design.

Chimerism. A phenomenon in which an individual carries some genetic material originating from another individual (for example, from their co-twin or mother).

Chorion. The outermost fetal membrane around the embryo. Forms the amniotic sac together with the amnion.

Classical twin design. The comparison of the resemblance of MZ and DZ twins for a phenotype or phenotypes to estimate the contribution of genetic and non-genetic (environmental) sources of (co)variance.

Concordance, pairwise. Defined as $C/(C+D)$, where C is the number of concordant pairs (that is, where both twins of a pair are affected) and D is the number of discordant pairs (that is, where only one twin of a pair is affected).

Concordance, proband-wise. Defined as $2C/(2C+D)$, in which C is the number of concordant pairs and D is the number of discordant pairs. Proband-wise concordance is preferred over pairwise concordance as it reduces the impact of non-genetic factors that can affect both twins in a pair.

Cultural transmission. A type of phenotypic transmission from parent to offspring that is independent of genetic transmission and gives rise to r_{GE} .

De novo mutation. A mutation that is not inherited from either parent and that can occur during the formation of the egg or sperm (germ cells) or during early embryonic development.

Direction of causation (DOC) modelling. Modelling based on cross-twin cross-trait correlations to deduce causal relationships between two variables with different inheritance patterns.

Discordance. When two twins within a pair do not exhibit the same trait—for example, when one twin develops a disease, whereas the other does not.

Environmental correlation. The extent to which the same environmental factors contribute to variation in two phenotypes. Can be estimated via multivariate twin models.

Epigenome. All chemical modifications that occur to the DNA molecule and the proteins associated with it that can affect gene expression without changing the underlying DNA sequence.

Gene–environment correlation (r_{GE}). Correlation between genotype and environment, which occurs when genotypes are not randomly distributed over environments. Multiple mechanisms (passive, reactive and active) can lead to r_{GE} .

Gene–environment interaction ($G \times E$). Present when genetic effects on a phenotype depend on the individual's environment or when environmental effects on a phenotype depend on an individual's genotype.

Genetic correlation. The extent to which the same genetic factors contribute to variation in two phenotypes. Can be estimated via multivariate twin models or be based on results from GWASs.

Genetic relationship matrix (GRM). A genetic covariance matrix between persons calculated from genotype data of the individuals.

Genome-based restricted maximum likelihood. A statistical method for variance component estimation in genetics that quantifies the

total narrow-sense contribution to a trait's heritability of a particular subset of genetic variants (for example, SNPs).

Genome-wide association study (GWAS). A regression-based approach to identify genomic variants (for example, SNPs) that are statistically associated with a disease or trait.

Heritability. The proportion of phenotypic variance that can be explained by differences in genetic make-up. In twin studies, we can differentiate between variance due to all genetic effects (broad-sense heritability) and due to additive genetic effects (narrow-sense heritability). Heritability derived from measured SNPs is referred to as SNP heritability.

Heteroscedasticity. A statistical phenomenon in a regression model where the variance of the residual (error) term is not the same for different values of the independent variable.

Locus. The chromosomal location of a genetic variant or gene.

Liability. A continuous variable that underlies human diseases or dichotomous traits. Variation of liability can be both genetic and environmental. Such diseases or traits are sometimes referred to as threshold characters, because a threshold divides the population into affected and unaffected persons.

Maximum likelihood estimation. A statistical approach to estimate the parameters (for example, heritability) of complex models on the basis of observed data by finding the values for parameters that maximize the likelihood function—that is, a function that measures how likely the observed data are to have been generated by the model.

Mendelian randomization (MR). An analytical approach that uses random segregation of genetic variants strongly associated with an exposure as an instrument to investigate the putative causal effect of the exposure on an outcome, allowing confounders and reverse causation to be reduced.

Metabolome. The entire range of small molecules (that is, lipids and sugars) in a cell, tissue or organism that are involved in cellular metabolism.

MR–DOC. A combination of DOC modelling and MR to test causal relationships while accounting for confounding factors.

Multilevel twin model. A reparameterized form of the classical twin model where variance is decomposed into levels: (1) within-pair variance, (2) between-family variance and (3) the variance of a higher-level clustering variable.

Multivariate twin models. Models in which the outcome consists of more than one phenotype per person to estimate the relative contributions of genetic and non-genetic (environmental) sources of covariance.

Network parameters. Descriptions of the connectivity of brain structures or psychological concepts as determined by (statistical) models.

New phenotype. A phenotype newly present in a population due to technological and scientific advancements or global phenomena (such as COVID-19).

Omics. Biological sciences that aim to study an organism's entire collection of molecules.

Penetrance. The probability that a disorder will occur if an individual is a carrier of a particular genotype.

Phenotype. Any measured or observed trait.

Pleiotropy. When a single gene (or a set of genes) affects multiple phenotypes. It is one of the mechanisms leading to genetic correlations among traits.

(continued from previous page)

Polygenic score (PGS). A value that characterizes an individual's genetic value for a disease or trait, which reflects prediction relative to the general population. It estimates the combined effects of multiple DNA variants in determining the probability of an individual having specific phenotypes or traits.

Proteome. The entire range of proteins, the vital molecules directly involved in cellular function, in a cell, tissue or organism.

Single nucleotide polymorphisms (SNPs). Substitutions of a single nucleotide at known genomic locations.

Singleton, twin, triplet, quadruplet, quintuplet (and so on). An individual from a single pregnancy that resulted in one, two, three, four or five (or more) offspring.

Statistical power. The probability of detecting an effect when it is there—in other words, the likelihood of correctly rejecting the null hypothesis.

Structural equation model. A statistical method to estimate the relationship between unobserved (latent) and observed variables.

Tetrachoric correlation. A special case of the polychoric correlation that represents a correlation between two, unobserved, continuous variables ('liabilities') from two observed variables that are assessed on an ordinal scale. Tetrachoric refers to the case when both observed variables are dichotomous.

Transcriptome. All RNA transcripts in a cell from protein-coding to non-coding RNA. It refers to the set of RNA molecules such as mRNA, tRNA, rRNA and others.

Transmitted/non-transmitted genotype design. A design in which transmission from parents to offspring is analysed via two types of genetic information: (1) genetic variants transmitted to offspring (this constitutes the offspring's genotype) and (2) genetic variants that are not passed down from parents to offspring (the non-transmitted genotype). The non-transmitted genotype can influence an offspring's phenotype; this is referred to as indirect genetic effects, genetic nurture or cultural inheritance.

Twinning. The process in which two or more offspring result from a single pregnancy.

X-inactivation. The process by which one of the copies of the X chromosome is inactivated (silenced).

Zygosity. Twins are either MZ (sometimes called 'identical') or DZ ('fraternal'). MZ twins develop from a single fertilized egg and, except for somatic mutations, share the entire DNA sequence. DZ twins develop from two separate eggs that are fertilized by two separate sperm. They share approximately 50% of their genetic material, like non-twin siblings. Triplets can be MZ, DZ or trizygotic.

turn to assumptions and advances in research methods and the study of twinning as a phenotype, which is important for the mothers of twins and for female fertility and may inform developmental biology. We conclude by discussing existing knowledge gaps in twin studies.

Classic twin design and twin concordance and discordance

Classic twin design

The classical twin design compares the resemblance of MZ and DZ twins for univariate or multivariate traits to estimate heritability and genetic correlations among phenotypes (a glossary of key terms is provided in Box 1). It decomposes phenotypic variance and covariance among traits into genetic and non-genetic components, on the basis of the biometrical model $P = G + E$, where P , G and E represent univariate or multivariate phenotypes, individual genetic values and environmental deviations, respectively³. P is measured, whereas G and E need not be observed, but their influence is estimated by comparing MZ and DZ twins, as illustrated in Fig. 1. The influence of the genotype can be summarized as the heritability of a trait: the proportion of variance due to genetic factors.

Twin resemblance can be quantified in concordance or correlations. Concordance and discordance refer to the degree of similarity or difference between twins in terms of their phenotypes. Concordance is typically applied to dichotomous phenotypes, such as the existence of a disorder, while correlations are used to summarize twin resemblance for continuous traits. Similar concordance rates in MZ and DZ twin pairs suggest a larger role of environmental factors, while a higher concordance in MZ than DZ twin pairs suggests stronger genetic effects. For continuous traits, we double the difference between the MZ and DZ correlations to obtain a first estimate of heritability—that is, the proportion of variance explained by genetic factors. However, the relation between twin concordance and trait heritability is more complex because it depends on the prevalence of a disorder in the population.

Twin concordance and discordance

Figure 2a presents an overview of estimates of twin proband-wise concordance rates (or case-wise with full ascertainment)⁴ in complex diseases, either directly obtained from publications by computing

proband-wise concordance from the published data or obtained by contacting the authors. To estimate heritability, researchers cannot only look at the differences in concordance rates between MZ and DZ twins; their resemblance must be obtained on the underlying liability scale by tetrachoric correlations^{5,6}. Figure 2b (based on Smith⁷) shows the relation between concordance, prevalence of a disorder in the population and heritability on the liability scale⁷. The figure specifies this relation for prevalences ranging from 0.01% for rare traits to 80% for common ones. It is especially striking that for traits with a low prevalence and a high heritability, most MZ twins will be discordant.

These results provide valuable information for research and society. While genetic effects are evident from the higher concordance in MZ twins than in DZ pairs, concordance in MZ twins does not approach 100%, even for highly heritable traits such as schizophrenia⁸. Explaining why one person becomes affected and their twin with an identical genotype is protected involves identifying environmental factors, epigenetic mechanisms and DNA mutations that occurred in one twin but not in the co-twin.

The extent to which the concordance rates in MZ twins are lower than 100% conveys how predictive the genotype will be for individual outcomes. An individual's genetic liability to a disorder can be seen as providing a baseline, with other factors influencing whether that liability ends up passing the threshold for disease penetrance. Thus, twin concordance rates powerfully illustrate that even a strong genetic predisposition does not have to result in the development of a disease.

The genome is not deterministic. Processes such as epigenetics (including X-inactivation, in utero exposure and post-natal environment) may lead to changes in DNA expression within MZ pairs. Epigenetic and DNA expression differences were found in MZ twins discordant for inflammation-related diseases (such as lupus⁹) and for psychiatric and neurodegenerative disorders (including schizophrenia¹⁰ and Alzheimer's disease¹¹). Differences in discordant MZ twin pairs have also been observed for proteomic^{12,13} and metabolomic profiles^{14,15} and in multi-omics studies^{16,17}.

For biomarkers, including omics profiles, it remains challenging to establish whether twin discordance is a cause or consequence of the disease. Here MZ twin pairs are valuable to examine causality¹⁸, especially when combined with the longitudinal data in twin registries.

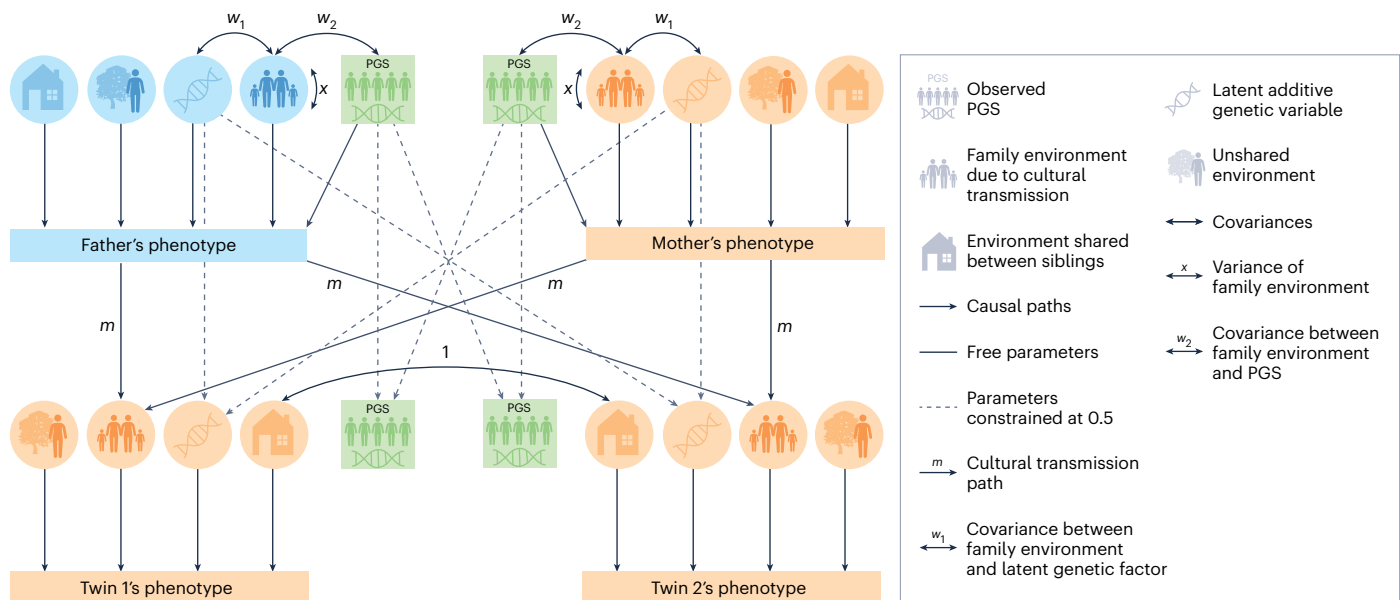


Fig. 1 Path diagram of a twins-and-parents design plus PGSs to disentangle genetic and cultural transmission. The circles denote latent (that is, unmeasured) variables, and the squares denote measured variables. To ease the presentation, variances for latent variables that are constrained at 1 are not drawn. The model as depicted is for DZ twins.

In one example addressing the causes versus consequences of smoking behaviour¹⁹, longitudinal phenotype data were combined with gene expression profiles, and 132 differentially expressed genes were identified in current, former and never smokers. Nearly all genes (125) had reversible effects on gene expression levels. In 56 MZ pairs discordant for current smoking, only six genes were differentially expressed, but the effects for 75% of the genes in the discordant MZ twin pairs were in the same direction as in the total population.

For diseases that are more common at an older age, one needs to consider that the disease can still manifest at later follow-up. Careful consideration of the age composition of the sample is therefore required, as illustrated by studies investigating twin concordance rates for Parkinson's disease^{20,21}. When twins were studied at an average age of 74 years, researchers found similar concordance rates for late-onset Parkinson's disease (>50 years of age) in MZ and DZ pairs. When they later searched the US National Death Index for all twin pairs with Parkinson's disease in at least one twin, they found higher concordance among MZ than DZ twins at follow-up, and heritability was now estimated at 20% for late-onset Parkinson's disease.

In an era of genetic testing services and polygenic scores (PGSs), the upper limit of genomic predictions must be defined and acknowledged. Researchers and clinicians may currently see risk prediction tools as having an upper limit set by trait heritability, but in fact predictions can go only as far as the MZ twin concordance rates.

Assumptions of the classical twin design

In its most basic application, the classical twin design assumes random mating in the parents of twins, and that environmental factors that influence a trait are similar for MZ and DZ twins so that any differences in the similarity of the twins' traits can be attributed to genetic factors. This so-called equal environment assumption has been tested numerous times and holds for most phenotypes²². Researchers can test assumptions regarding equality of means and variances in MZ and DZ twins and the absence of interactions or correlations between genes and environment²³ with appropriate data and methods, and deviations from these assumptions may provide new insights, as we discuss below.

Another assumption entails that twins are representative of the population to which heritability and other estimates apply. Twins and

singletons are similar in most respects^{24,25}, even though twins are born earlier on average and grow up with a sibling of the same age. Twins and non-twins have similar biomarker profiles and diseases²⁶, cognitive functions²⁷, health behaviours²⁸, personality²⁹ and psychopathology³⁰. Despite early hypotheses that twins from DZ pairs are more often chimeric than singletons, a large study in twin-family pedigrees found this not to be the case³¹. Although twins are born earlier on average and have lower birth weight³² and body mass index (BMI)^{33,34}, genetic correlations of twin and singleton birth weights are not significantly different from each other³⁵, meaning that the same genes influence birth weight in twins and singletons. Consequently, data from twins can be combined with other population-based results or included in population-based genome-wide association studies (GWASs) to increase the sample size. As David Lykken stated in his 1982 presidential address to the Society for Psychophysiological Research, "Twins are probably more representative of the general population than any other group ... This representativeness is even more true of the families of twins"³⁶.

DNA sequence differences within MZ twin pairs

Post-zygotic (that is, after fertilization) mutations occurring in one twin of an identical pair can lead to genetic and phenotypic differences between them. However, if such mutations occur early in embryonic development—that is, before a split occurs—both twins will have the mutation. In a DNA sequencing project of 381 Icelandic twin pairs and two triplet sets, 39 pairs differed by over 100 mutations, while 38 pairs did not differ at all³⁷. The median number of post-zygotic mutations differing within twin pairs was 14, with a higher estimate of 48 for high-coverage pairs. In 78 parent-offspring trios from the same population, 4,933 de novo mutations were reported—that is, an average rate of 1.2×10^{-8} per nucleotide per generation³⁸. A review also put 1.18×10^{-8} per position as the best estimate of the average human germline mutation rate, corresponding to 74 novel single nucleotide polymorphisms (SNPs) and approximately three novel structural variants³⁹ per genome per generation. De novo mutations are enriched in coding and regulatory regions of the genome⁴⁰. Thus, it is not surprising that de novo mutations account for a substantial component of some rare genetic syndromes and that mutations in one twin but not the other can lead to discordance between them. For highly polygenic traits, individual



Fig. 2 | Concordance rates in twins. a, Proband-wise MZ and DZ twin concordance rates. References for the studies and the non-rounded concordance rates are provided in Supplementary Table 2. **b**, Expected concordance rates in MZ twins based on Smith⁷ detailing the relation between twin concordance,

disorder prevalence and disorder heritability on the liability scale. The expected MZ twin concordance, prevalence and heritability for the highlighted disorders are provided in Supplementary Table 3.

SNP effects are very small. If only a few SNPs are discordant per pair, the effects of these are negligible. While it is likely that de novo mutations do not impact twin-based heritability estimates of polygenic traits, we should be cautious when investigating rare genetic disorders that involve only a small number of genes or disorders that involve de novo mutations particularly.

DNA sequence differences between MZ twins can also inform on the timing of developmental mutations. Germ cells are specified around the third week after fertilization, a process referred to as primordial germ cell specification (PGCS). If post-zygotic de novo mutations are present in both the soma and germline of a twin, we can classify them as pre-PGCS. Both twins can share pre-PGCS mutations, or these can be present in just one twin, indicating whether mutations occurred before or after the twinning split. When post-zygotic mutations are present in the soma or the germline of both a proband twin and their offspring, we can classify them as post-PGCS mutations.

Methodological developments in twin studies

The classical twin design can be extended to explore gene–environment interactions and correlations. Gene–environment interaction ($G \times E$) can be conceptualized as moderation, where genes moderate the environmental effects or where the environment moderates or controls genetic effects. Gene–environment correlation (r_{GE}) describes a correlation between G and E in the model $P = G + E$ and is sometimes viewed as genetic control over exposure to different environments. There is ample theoretical work on gene–environment interactions but relatively little empirical work. In contrast, several developments have stimulated the study of r_{GE} . These involve adding measured G or E

to the design, extensions to longitudinal data and the addition of family members, such as the parents or the offspring of twins.

Gene–environment correlation

Detecting r_{GE} in the twin design can be achieved by including measured genetic or environmental variables. For instance, if the phenotype is depression, marital status may be a relevant measured environment, and a PGS for depression a relevant genetic variable. If genetic factors influence depression and marital status, this gives rise to r_{GE} . Obviously, this approach requires a prior expectation about which environmental variable to assess. By including a measured genetic variable—for example, a PGS for depression—we can estimate the correlation between the shared family environment and all genetic factors that influence depression⁴¹. This approach is exploratory, as it can detect r_{GE} but does not require prior hypotheses concerning the environmental influences involved in the r_{GE} .

An alternative approach to studying r_{GE} in the classical twin design is by analysing longitudinal data⁴². Given repeated measures, r_{GE} can be modelled by estimating the regression of the latent environmental variable at time t (E_t) on the phenotype at time $t - 1$ (P_{t-1}). This regression implies r_{GE} by the following chain of regression relationships: the genotype at $t - 1$ influences the phenotype at $t - 1$, which influences E_t . So, P_{t-1} mediates the relationship between G_{t-1} and E_t , thus giving rise to r_{GE} (Fig. 3). This influence of the phenotype on the environment may arise if children actively seek out or create environmental circumstances, which match their genotype. Given the own role of the child, this is called active r_{GE} or ‘niche-picking’. This may apply to intelligence and musicality, for example. Alternatively, the behaviour of the child may

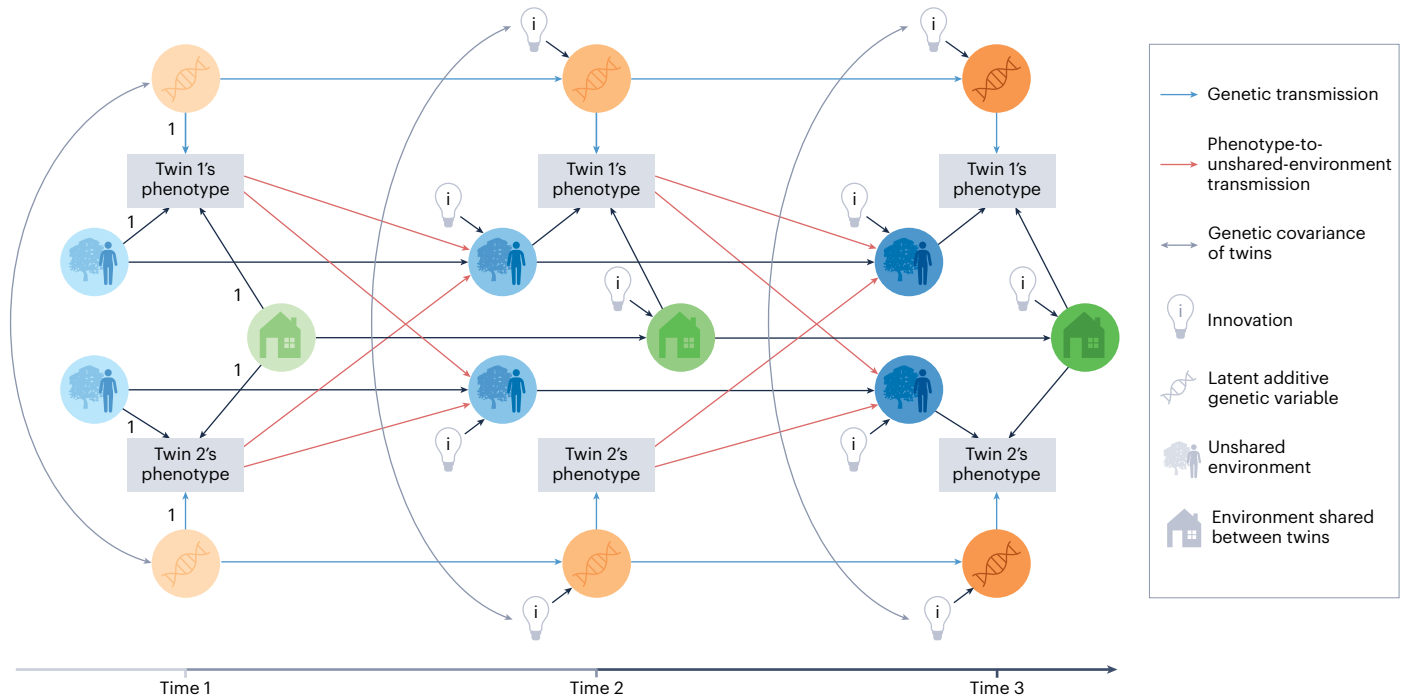


Fig. 3 | Path diagram for a time-dependent process. The model is shown for a pair of twins, whose longitudinal phenotypes are influenced by their genotype and the shared and unshared environment. The latent genotype and the shared and unshared environment can be influenced by earlier times (for example,

stable genetic influences) and by innovations (for example, new expressed genes). The red arrows show phenotype-to-environment transmission, which leads to GE correlation.

elicit responses from persons in the environment, called evocative r_{GE} . Examples are highly outgoing children, who encourage social interaction, or aggressive children, who may discourage social interaction and elicit corrective parenting.

Simultaneous environmental and genetic transmission across generations also leads to correlations between genotype and environment in offspring, as their genotype depends on the parental genotypes (genetic transmission), and their environment (partly) depends on the parental genotypes, as mediated by the parental phenotype (cultural transmission). This form of r_{GE} is called passive r_{GE} , as the twin offspring are the (passive) recipients of parental genes and an environment (partly) created by parents. The environment that parents create depends on the genotype they transmit to their offspring and on the part that is not transmitted. The parent–twin design can incorporate the effects of the non-transmitted genotypes on offspring outcomes⁴³ (depicted in Fig. 1).

In the children-of-twins design, which involves MZ and DZ twin pairs as parents, researchers compare the resemblance between parents and their own children to the resemblance between offspring and their parent’s co-twin^{44–46}. In both the children-of-twins design and the parents-of-twins design, researchers estimate the effect that remains after accounting for direct genetic transmission, which induces a correlation between genotype and the shared environment.

A new way of modelling parental transmitted and non-transmitted genotypes is structural equation modelling combined with PGSs (SEM–PGS), which disentangles parental genetic and environmental effects on offspring traits while controlling for assortative mating, thereby providing unbiased estimates of genetic transmission⁴⁷. Researchers can use SEM–PGS with trio data but also with twin pairs⁴⁸. This way, they can study parental effects by combining both GWAS and twin-based designs.

The outcomes of intergenerational twin studies can help to inform intervention approaches⁴⁹ and avoid parent-blaming⁵⁰. An example can be seen in cases in which a child is struggling in school, and the

teacher blames its parents for not providing a suitable home environment or not being involved in their child’s education. In a parent–twin study of offspring years of education, there was a direct effect of parental genes that were inherited by the twin offspring⁵¹, making it difficult to hold the parents responsible for passing on their genes. The indirect (genetic nurturing) effect of parental genes that were not transmitted to offspring weakened after controlling for parental IQ or socio-economic status, implying that the genetic nurturing effect on education reflects family socio-economic status rather than any deliberate parenting behaviour⁵¹. Genetic nurture effects are often interpreted as parenting effects but may also reflect the broader family environment.

Gene–environment interaction

Researchers can investigate $G \times E$ by comparing heritability estimates across different environments. One recent example is before and during exposure to the COVID-19 pandemic. A comparison of heritability before and during the first lockdown in the Netherlands (gene-by-crisis interaction) showed a slight increase in heritability estimates for optimism and meaning of life during the first months of the pandemic and lower-than-unity genetic correlations across time (0.75 and 0.63), implying $G \times E$, with both quantitative and qualitative differences in genetic influences⁵².

Purcell suggested a broad method to measure changes in heritability on the basis of traits that could potentially have a heritable factor⁵³. Some caution is needed if such traits, or moderators, are themselves correlated between twins, but several solutions exist⁵⁴. Molenaar et al. developed an alternative approach that does not require a measured moderator to estimate $G \times E$ ⁵⁵. Their method exploits the fact that $G \times E$ will introduce heteroscedasticity^{56,57}. By modelling environmental variance conditional on the genetic variance, they test for departures from bivariate normality in twin data. This method can explicitly estimate $G \times E$ without a priori hypotheses about specific interacting genotypes or environments.

Interaction among family members

Individuals can contribute to the environment of their family members, including twins and their co-twins^{44,58}, and these effects can be either competitive or cooperative⁵⁹. Under competition, the higher the expression of a trait in one sibling, the less it will be exhibited in the other sibling. Under cooperation, a trait expressed more strongly in one sibling will also be exhibited more in the other sibling. The presence of sibling interaction may be optimally detected in twin designs. Here, violation of the equal variance assumption in the two zygosity groups implies either cooperation or competition: inflated phenotypic variances and increased twin correlations imply cooperation, while deflated variances and decreased DZ correlations relative to MZ correlations suggest competition. Both inflation and deflation act more on the variance of MZ twins than that of DZ twins, thereby making the twin design a powerful tool to detect and distinguish competitive and collaborative interactions.

Multilevel twin models

When working with data from twins, natural clusters occur that are created by the fact that we include individuals from the same family unit. These family units are further nested in higher-level clusters such as geographic region, neighbourhood or school. If both twins in a pair share higher-level clustering variables, their effects will be reflected in the estimate for the common environment in the classic twin design. In multilevel twin models, variance is decomposed into within-pair variance on the first level (that is, representing differences between twins within a pair), between-family variance on the second level (that is, differences between pairs or families) and the variance of a higher-level clustering variable on the third level (for example, differences between geographic regions)⁶⁰. For instance, multilevel twin modelling can clarify the relationship between geographic regional clustering and ancestry, as seen in regional clustering for height in seven-year-old Dutch twins, where, after adjusting for ancestry, the variation in height was no longer explained by region⁶¹. Hence, regional clustering effects may mask genetic ancestry effects, which can be disentangled in multilevel twin models.

Causal modelling in twin designs

Randomized controlled trials are considered the gold standard for establishing causality, but these are often not feasible or ethical in health and behavioural research. In such cases, researchers can turn to genetic designs such as Mendelian randomization (MR) and the direction of causation (DOC) twin models to study causality. MR uses an observed genetic variable (for example, SNP or PGS) as an instrumental variable that is assumed to correlate with the exposure variable, but not with confounding variables or the outcome⁶². However, MR analysis with SNPs has low statistical power, and PGSs will often be unfeasible due to genetic pleiotropy. The bivariate DOC twin model was proposed for twin data to test causal relationships between two variables. Depending on whether trait A causes trait B or vice versa (that is, the direction of causation), the model predicts different cross-twin cross-trait correlations. The DOC model works only in situations in which two variables have different modes of inheritance—that is, one trait is influenced by the shared environment, while the other is mainly influenced by genes⁶³. Combining MR with the DOC twin approach in an MR–DOC model yields a causal test that does not require the mode of inheritance to differ for the two variables and relaxes several critical MR assumptions concerning pleiotropy. If PGSs of twins are used as instrumental variables in a bivariate classical twin design, the PGSs can correlate with both predictor and outcome^{64,65}. Combining DNA-based methods with twin-based designs can thus overcome the limitations of individual methods. While no single approach is free of bias, triangulation can help to address the limitations of methods and facilitate replicable and reproducible results.

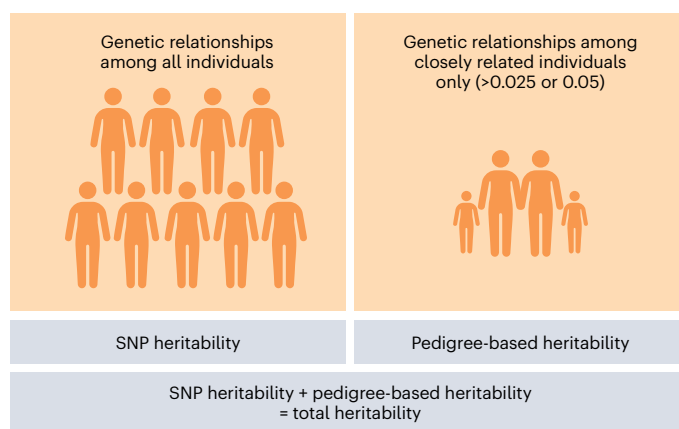


Fig. 4 | The threshold GRM approach. Two GRMs (orange boxes) are the basis to obtain two heritability estimates: SNP heritability and pedigree-based heritability. The two estimates sum to the total heritability. In the GRM giving rise to the pedigree-based heritability, only relations among third or fourth cousins (who share 2.5% of their genetic material—that is, 0.025) or among second cousins (0.05) or more closely related family members are retained.

The added value of twins when decomposing heritability

The classical twin design can be combined with genome-based restricted maximum likelihood methods to simultaneously estimate SNP heritability and pedigree-based heritability^{66,67}. SNP heritability refers to the contribution of measured SNPs to trait heritability. Pedigree-based heritability captures heritability on the basis of the known relationships among not-too-distantly related family members. Here, closely related family members are typically defined as cousins two or three times removed or closer. This approach is referred to as threshold genetic relationship matrix (GRM) and relies on two GRMs: one GRM including all genetic relationships, and one where the genetic relationship among distantly related pairs of individuals is set to zero to capture pedigree-associated variation. The total heritability is obtained by summing the estimates as obtained from both GRMs (Fig. 4). By including additional GRMs that comprise specific SNPs, researchers can extend the threshold GRM approach to capture the contributions of functional or biologically relevant elements⁶⁸. For example, a four-GRM approach was used to obtain heritability estimates for known metabolite loci and distinguish between metabolite classes⁶⁹. Besides the GRMs in the threshold approach, from which all metabolite loci were removed, this study included a third GRM to capture the metabolite loci of a specific metabolite class and a fourth GRM to include all metabolite loci for all other metabolite classes.

Advances in health and behavioural phenotypes

Twinning

DZ and MZ twinning have different biological mechanisms. DZ twins result from two separate fertilizations and have different placentas and fetal membranes. MZ twins, in contrast, come from a single sperm and egg and separate within two weeks of fertilization. MZ twins may have shared or separate amnions and chorions⁷⁰. DZ twinning has a genetic component^{71–73}, but estimates of heritability for being a DZ twin mother are remarkably scarce. Duffy and Martin estimated the heritability of twinning in historical pedigrees to lie between 8% and 20%⁷⁴. They could not distinguish between MZ and DZ twins, potentially underestimating the true heritability, but their estimate was consistent across time (8–19 generations) and ancestries (West Africa, Europe and Canada).

The first GWAS for being a mother of natural DZ twins identified two genes, which replicated in the deCODE Icelandic database⁷⁵—namely, follicle-stimulating hormone beta subunit (*FSHB*) and SMAD family member 3 (*SMAD3*)—with risk alleles increasing the frequency of twin births in Iceland by 18% and 9%, respectively. Whereas FSH genes

are strong candidates for DZ twinning, the finding for *SMAD3*, which is expressed in the human ovaries, was new. SNPs associated with DZ twinning have an impact on several reproductive traits, including increased fertility (having more children), decreased risk of polycystic ovary syndrome, earlier onset of menstruation, earlier natural menopause and giving birth to the first child at a younger age.

By contrast, the genetic aetiology for MZ twinning remains unclear, but van Dongen et al. discovered a strong epigenetic signature for MZ twinning comprising 834 differentially methylated positions in adult somatic tissues⁷⁶. The loci were enriched for putative metastable epi-alleles, which are epigenetic marks or modifications that occur during early development. This lifelong signature opens up new avenues to investigate the vanishing MZ twin syndrome (that is, the disappearance of an embryo during early pregnancy⁷⁷) and congenital disorders that have an overrepresentation of MZ twins⁷⁸.

Phenotypes of increasing interest

Twin studies are of interest not only to psychology, biology and medicine but also to fields such as political science, sociology and economics^{79,80}. In economics, twin studies have demonstrated genetic effects in classic economic paradigms such as the trust game, which manipulates participants' willingness to invest money and to reciprocate others' trust⁸¹. In political science, twin studies have shown that genetic contributions to political ideology are largely independent of the chosen measurement, time or population⁸². Inside and outside of these fields, topics previously studied become relevant again, and new ones arise. For example, during the global COVID-19 pandemic, a study in twins estimated a heritability of 31% for COVID-19 on the basis of self-reported symptoms⁸³.

The outcomes of twin studies can benefit interventions. As genetic influences can change over a lifetime, the longitudinal nature of twin registers can inform policy decisions aimed at a specific age group⁸⁴. For example, as obesity often starts in childhood⁸⁵ and twin studies show that the heritability of BMI increases throughout childhood⁸⁶, this suggests that public health interventions might be most beneficial earlier in childhood rather than later in life. We may see this process as a recursive loop⁸⁷ where policies can inspire new questions for researchers, and results from research can lead to new policies. Society and its political, economic and social agents influence funding decisions, thus making twin studies relevant to a wide audience. This includes economists looking at the costs and benefits of policy decisions, sociologists studying the effects of policy decisions in society and political scientists studying the entire process of policy development.

Scientists develop interests in phenotypes that become accessible because of our technological advancement. Digital fingerprinting, the unique trace of a person's activity on the internet^{88,89}, is one example of such a phenotype. Language used on Facebook can predict depression symptoms⁹⁰, and more complete digital fingerprints may serve as predictors for other health and behavioural outcomes. However, researchers do not know whether digital fingerprints in MZ twins are as similar as their physical fingerprints. One study found that genetic effects can fully explain the familial resemblance in the frequency of internet use and that genetic and environmental factors account for different aspects of use⁹¹. It seems plausible that variance in digital fingerprints will also have a heritable component. If MZ twins have highly similar digital fingerprints (more so than DZ twins or siblings), this also creates options for the application of the classical twin design to nationwide register data, as these registers contain information on the entire population of twins but often lack information on zygosity for same-sex pairs.

An approach to handle unknown zygosity was suggested by combining twin and sibling data in register data⁹². Male–female pairs are DZ, but same-sex pairs are MZ or DZ. As there are equal numbers of opposite- and same-sex DZ twin pairs, we know the proportion of DZ pairs. As same-sex DZ twins and same-sex siblings share the same

genetic similarity, we can utilize the ratios of concordant and discordant pairs of the latter as weighting factors to determine the number of concordant and discordant pairs for same-sex DZ and MZ twin pairs. This approach has been used to estimate the heritability for attention deficit hyperactivity disorder in German health insurance data (0.77 for females and 0.88 for males) as inferred from ICD-10 diagnoses and drug prescriptions. These estimates closely resemble those obtained from classical twin studies. Thus, with only sex and diagnostic status of twin and sibling pairs, researchers could determine the contribution of genetics.

Neural and symptom networks

Twin studies can shed light on the genetic influence on the development, prognosis and comorbidity of neuropsychological and psychopathological disorders by analysing the network parameters of these disorders^{93,94}. Researchers investigate the genetics of nodes and edges in graph-theory-based networks or in network analyses of psychopathological phenotypes. Network parameters, such as the efficiency of the structural brain network (derived from magnetic resonance imaging data), have strong genetic correlations with intelligence^{95,96} and schizophrenia⁹⁷. In networks of structural and functional brain connectivity, the heritability of network parameters ranges from 0.05 to 0.74 depending on participants' age and the analytical method used (Supplementary Table 1). In networks for psychopathological disorders, central nodes (nodes with a relatively large number of connections to other nodes in the network) generally have higher heritability⁹⁸.

Applications of network analyses face numerous challenges, including a lack of criteria for identifying the completeness of networks, and the heterogeneity and replicability of networks^{99,100}. Network configuration can vary for different patients and across time, challenging the generalizability and reliability of network approaches⁹⁸. Twin studies can help to resolve such issues by considering twin resemblance for network parameters, where high correlations in MZ twins imply high reliability³⁶.

Omicstrait

Recent technological advancements in measuring genetic variation, transcriptomes, epigenomes, proteomes and metabolomes have enabled researchers to study heritability in twin cohorts through classical twin designs or a combination of classical twin designs and genome-based restricted maximum likelihood. Heritability estimates for gene expression from twin studies that characterize RNA transcripts in a cell by microarrays or sequencing¹⁰¹ differ by tissue. For example, adipose tissue has a higher average heritability (26%) than lymphoblastoid cell lines (21%) and skin tissue (16%)¹⁰². The average heritability of peripheral blood gene expression is between 10% and 20%, with mean heritability estimates significantly higher in RNA sequencing studies than those from microarray-expression data^{68,103}.

Twin-based heritability estimates for DNA methylation vary by sex and age. DNA methylation, the most commonly measured epigenetic mechanism in epidemiological studies, involves the addition of a methyl group to the C5 position of cytosines in the DNA, which decreases the genome accessibility for transcription, thereby regulating gene expression¹⁰⁴. On the basis of adult twin data, the heritability of blood DNA methylation across the genome was 19% on average, with common SNPs explaining on average 7% of the variance—that is, 37% of the total twin-based heritability¹⁰⁵. Heritability for DNA methylation decreased with age, driven by an increase in environmental variance. A combination of multiple twin datasets across the lifespan (0–92 years) demonstrated an increase in familial correlations in MZ and DZ twins from birth to adolescence and a decrease after young adulthood¹⁰⁶. The rate of change in familial correlations was similar in MZ and DZ pairs, which suggests that age affected the influence of environmental rather than genetic effects on DNA methylation across the lifespan.

Twin studies of proteomics and metabolomics generally report higher heritability estimates than those found for transcriptome and epigenome variables. One study in female twins estimated that genes accounted for 13.6% of the variance in plasma protein levels¹⁰⁷. Post-translational modifications of proteins, particularly glycosylation (attachment of a glycan to a protein), are also heritable, with many *N*-glycans showing high heritability (>50%) in plasma^{108,109}. Metabolites comprise a diverse set of small molecules (<1.5 kDa) involved in cellular metabolism, including amino acids, sugars and lipids¹¹⁰. Twin-based heritability estimates of blood metabolite levels cluster around 50% but can differ among metabolite classes^{111,112}. Similar heritability estimates have been observed for urinary metabolites^{113,114}.

To fully understand biological processes, we need to collect and analyse data from multiple omics domains together¹¹⁵. Simultaneous multi-omics modelling can complement single-omics analyses by accounting for the relationships among omics domains¹¹⁶. We anticipate that the fusion of twin designs with data from multiple omics domains will enhance the categorization or differentiation of diseases and the forecasting of biomarkers.

Concluding remarks

In today's globalized and increasingly interconnected research world, collaboration and cooperation are key. Over 60 twin registers are in place worldwide, across multiple continents in 26 countries¹¹⁷, with many collaborating in the Collaborative Project of Development of Anthropometrical Measures in Twins (CODATwins)¹¹⁸, which amasses and shares data on height, BMI and size at birth from 54 twin projects from 24 countries. The results for height and BMI are remarkable, as the differences in heritability estimates were minor across twin cohorts from different cultural–geographic regions and in individuals from different birth cohorts or for whom height and BMI were obtained at different ages.

Existing registries cannot always answer novel research questions, and the replication of findings in understudied twin populations warrants the development of new twin registries¹¹⁹. Particularly, African, Arab, Hispanic and other non-European populations are still underrepresented in twin research. Also, some age groups are still underrepresented in twin studies. With the lengthening of average life expectancies and the rapid rise of aging populations, understanding the causes and effects of aging and age-related declines in health is vital¹²⁰. While twin studies can make valuable contributions to these research topics, few current registries include twins of advanced age. Eventually, longitudinal registries may aid here by following their participants into advanced age, and establishing geriatric cohorts will accelerate research into aging. Another solution, which also tackles the attrition often observed in longitudinal studies, involves linking existing twin registers to electronic health records and other national registers.

Periconceptual and prenatal variables remain understudied¹²¹. In a large meta-analysis of heritability estimates in over 14 million twin pairs for over 17,000 traits¹²², less than 5% of the studies published between 1958 and 2012 investigated early life traits or childhood-onset disorders. The low heritability estimates and substantial shared environmental effects reported for several of these traits might make them especially interesting to study^{123,124}. Researchers should ideally follow mothers of twins and twins from the pre-pregnancy period. This would seem difficult, as predicting a twin pregnancy is hardly possible, but many fertility clinics hold medical records on the conception of twins and have the potential to create such research resources.

For all phenotypes, we need to consider sample sizes when applying twin methods. Is larger always better, or can 'better' also involve phenotypes with improved reliability or designs that profit from unique events? An example would be the National Aeronautics and Space Administration twin study that enabled the investigation of the effect of long-duration spaceflight on the human body in one pair

of identical twins¹²⁵. While astronaut Mark Kelly stayed on Earth, his identical twin Scott spent a year in space. Spaceflight affected some bodily functions even six months after returning to Earth, including gene expression changes. Other changes, such as body weight and metabolite levels, were less persistent and returned to the pre-flight levels after a shorter time.

Twin studies have helped in understanding the influence of the environment by analysing traits that tend to be labelled as 'environmental', such as social environment, leisure-time activities and life events, indicating that with an average heritability of 49%, these 'environmental' variables are partially under genetic control¹²⁶. The classical twin design is still one of the most powerful to distinguish between unique and shared environmental influences. While GWAS approaches are increasingly powerful to estimate (besides heritability) the bivariate genetic correlations between traits, in contrast to twin studies, they do not inform on the phenotypic or environmental correlations between phenotypes and do not account for the total heritability of the traits. Moreover, traits that show evidence of shared environment are predicted to show evidence of genetic nurture in designs with non-transmitted PGSs¹²⁷. The results from twin studies can thus generate new hypotheses about genetic nurture.

Twin registers often have multigenerational genotype data and a wide variety of phenotypes and will continue to be essential contributors of data to GWASs, within-family designs, causality modelling, inter-generational transmission and longitudinal studies. Twin researchers mostly collect twin data, though interest in twin research by other disciplines is beneficial, as every discipline has its way of measuring constructs¹²⁸. Diverse disciplines should collaborate to ensure clarity about the meaning of results, such as when a trait is reported to have high heritability¹²⁹. Despite rapid technological advances enabling increasingly large-scale omics investigations for complex human traits and the renewed interest in family-based methods, twin designs have contributed to the major discoveries in behavioural genomics and will continue to do so¹³⁰.

Code availability

The code to reproduce Fig. 2b is available on GitHub: <https://github.com/KuznetsovDima/NATHUMBEHAV-22123354/>.

References

- Boomsma, D., Busjahn, A. & Peltonen, L. Classical twin studies and beyond. *Nat. Rev. Genet.* **3**, 872–882 (2002).
- Craig, J. M., Calais-Ferreira, L., Umstad, M. P. & Buchwald, D. The value of twins for health and medical research: a third of a century of progress. *Twin Res. Hum. Genet.* **23**, 8–15 (2020).
- Falconer, D. S. & Mackay, T. F. *Introduction to Quantitative Genetics* (Prentice Hall, 1996).
- McGue, M. When assessing twin concordance, use the probandwise not the pairwise rate. *Schizophr. Bull.* **18**, 171–176 (1992).
- Visscher, P. M. & Wray, N. R. Concepts and misconceptions about the polygenic additive model applied to disease. *Hum. Hered.* **80**, 165–170 (2015).
- Falconer, D. S. The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann. Hum. Genet.* **29**, 51–76 (1965).
- Smith, C. Heritability of liability and concordance in monozygous twins. *Ann. Hum. Genet.* **34**, 85–91 (1970).
- Owen, M. J., Sawa, A. & Mortensen, P. B. Schizophrenia. *Lancet* **388**, 86–97 (2016).
- Marion, M. C. et al. Nucleic acid-sensing and interferon-inducible pathways show differential methylation in MZ twins discordant for lupus and overexpression in independent lupus samples: implications for pathogenic mechanism and drug targeting. *Genes* **12**, 1898 (2021).

10. Castellani, C. A. et al. DNA methylation differences in monozygotic twin pairs discordant for schizophrenia identifies psychosis related genes and networks. *BMC Med. Genomics* **8**, 17 (2015).
11. Konki, M. et al. Peripheral blood DNA methylation differences in twin pairs discordant for Alzheimer's disease. *Clin. Epigenet.* **11**, 130 (2019).
12. Kazuno, A. et al. Proteomic analysis of lymphoblastoid cells derived from monozygotic twins discordant for bipolar disorder: a preliminary study. *PLoS ONE* **8**, e53855 (2013).
13. O'Hanlon, T. P. et al. Plasma proteomic profiles from disease-discordant monozygotic twins suggest that molecular pathways are shared in multiple systemic autoimmune diseases. *Arthritis Res. Ther.* **13**, R181 (2011).
14. Muniandy, M. et al. Plasma metabolites reveal distinct profiles associating with different metabolic risk factors in monozygotic twin pairs. *Int J. Obes.* **43**, 487–502 (2019).
15. Tsang, T. M., Huang, J. T.-J., Holmes, E. & Bahn, S. Metabolic profiling of plasma from discordant schizophrenia twins: correlation between lipid signals and global functioning in female schizophrenia patients. *J. Proteome Res.* **5**, 756–760 (2006).
16. Bondia-Pons, I. et al. Metabolome and fecal microbiota in monozygotic twin pairs discordant for weight: a Big Mac challenge. *FASEB J.* **28**, 4169–4179 (2014).
17. Zhu, Y. et al. Genome-wide profiling of DNA methylome and transcriptome in peripheral blood monocytes for major depression: a monozygotic discordant twin study. *Transl. Psychiatry* **9**, 215 (2019).
18. Vitaro, F., Brendgen, M. & Arseneault, L. The discordant MZ-twin method: one step closer to the holy grail of causality. *Int. J. Behav. Dev.* **33**, 376–382 (2009).
19. Vink, J. M. et al. Differential gene expression patterns between smokers and non-smokers: cause or consequence? *Addict. Biol.* **22**, 550–560 (2017).
20. Goldman, S. M. et al. Concordance for Parkinson's disease in twins: a 20-year update. *Ann. Neurol.* **85**, 600–605 (2019).
21. Tanner, C. M. et al. Parkinson disease in twins: an etiologic study. *JAMA* **281**, 341–346 (1999).
22. Plomin, R., DeFries, J. C., Knopik, V. S. & Neiderhiser, J. M. *Behavioral Genetics* (Worth, 2012).
23. Willoughby, E. A., Polderman, T. J. C. & Boutwell, B. B. Behavioural genetics methods. *Nat. Rev. Methods Prim.* **3**, 11 (2023).
24. Christensen, K. & McGue, M. in *Twin Research for Everyone* (eds Tarnoki, A. et al.) 439–456 (Academic Press, 2022).
25. Willemsen, G., Odintsova, V., de Geus, E. & Boomsma, D. I. in *Twin and Higher-Order Pregnancies* (eds Khalil, A. et al.) 51–71 (Springer International, 2021).
26. de Geus, E. J. C., Posthuma, D., IJzerman, R. G. & Boomsma, D. I. Comparing blood pressure of twins and their singleton siblings: being a twin does not affect adult blood pressure. *Twin Res.* **4**, 385–391 (2001).
27. Christensen, K. et al. Comparison of academic performance of twins and singletons in adolescence: follow-up study. *Br. Med. J.* **333**, 1095 (2006).
28. Barnes, J. C. & Boutwell, B. B. A demonstration of the generalizability of twin-based research on antisocial behavior. *Behav. Genet.* **43**, 120–131 (2013).
29. Johnson, W., Krueger, R. F., Bouchard, T. J. & McGue, M. The personalities of twins: just ordinary folks. *Twin Res. Hum. Genet.* **5**, 125–131 (2002).
30. Kendler, K. S., Ohlsson, H., Lichtenstein, P., Sundquist, J. & Sundquist, K. The genetic epidemiology of treated major depression in Sweden. *Am. J. Psychiatry* **175**, 1137–1144 (2018).
31. Johnson, B. N. et al. Male microchimerism in females: a quantitative study of twin pedigrees to investigate mechanisms. *Hum. Reprod.* **36**, 2529–2537 (2021).
32. Mook-Kanamori, D. O. et al. Heritability estimates of body size in fetal life and early childhood. *PLoS ONE* **7**, e39901 (2012).
33. Silventoinen, K., Magnusson, P. K. E., Tynelius, P., Kaprio, J. & Rasmussen, F. Heritability of body size and muscle strength in young adulthood: a study of one million Swedish men. *Genet. Epidemiol.* **32**, 341–349 (2008).
34. Estourgie-van Burk, G. F., Bartels, M., Boomsma, D. I. & Deleamarre-van de Waal, H. A. Body size of twins compared with siblings and the general population: from birth to late adolescence. *J. Pediatr.* **156**, 586–591 (2010).
35. Beck, J. J. et al. Genetic meta-analysis of twin birth weight shows high genetic correlation with singleton birth weight. *Hum. Mol. Genet.* **30**, 1894–1905 (2021).
36. Lykken, D. T. Research with twins: the concept of emergence. *Psychophysiology* **19**, 361–372 (1982).
37. Jonsson, H. et al. Differences between germline genomes of monozygotic twins. *Nat. Genet.* **53**, 27–34 (2021).
38. Kong, A. et al. Rate of de novo mutations and the importance of father's age to disease risk. *Nature* **488**, 471–475 (2012).
39. Veltman, J. A. & Brunner, H. G. De novo mutations in human genetic disease. *Nat. Rev. Genet.* **13**, 565–575 (2012).
40. Francioli, L. C. et al. Genome-wide patterns and properties of de novo mutations in humans. *Nat. Genet.* **47**, 822–826 (2015).
41. Dolan, C. V., Huijskens, R. C. A., Minică, C. C., Neale, M. C. & Boomsma, D. I. Incorporating polygenic risk scores in the ACE twin model to estimate A–C covariance. *Behav. Genet.* **51**, 237–249 (2021).
42. Dolan, C. V., de Kort, J. M., van Beijsterveldt, T. C. E. M., Bartels, M. & Boomsma, D. I. GE covariance through phenotype to environment transmission: an assessment in longitudinal twin data and application to childhood anxiety. *Behav. Genet.* **44**, 240–253 (2014).
43. Wang, B. et al. Robust genetic nurture effects on education: a systematic review and meta-analysis based on 38,654 families across 8 cohorts. *Am. J. Hum. Genet.* **108**, 1780–1791 (2021).
44. D'Onofrio, B. M. et al. The role of the children of twins design in elucidating causal relations between parent characteristics and child outcomes. *J. Child Psychol. Psychiatry* **44**, 1130–1144 (2003).
45. McAdams, T. A. et al. Revisiting the children-of-twins design: improving existing models for the exploration of intergenerational associations. *Behav. Genet.* **48**, 397–412 (2018).
46. McAdams, T. A. et al. Accounting for genetic and environmental confounds in associations between parent and child characteristics: a systematic review of children-of-twins studies. *Psychol. Bull.* **140**, 1138–1173 (2014).
47. Balbona, J. V., Kim, Y. & Keller, M. C. Estimation of parental effects using polygenic scores. *Behav. Genet.* **51**, 264–278 (2021).
48. Balbona, J. V., Kim, Y. & Keller, M. C. The estimation of environmental and genetic parental influences. *Dev. Psychopathol.* **34**, 1876–1886 (2022).
49. Conley, D. et al. Is the effect of parental education on offspring biased or moderated by genotype? *Sociol. Sci.* **2**, 82–105 (2015).
50. Hart, S. A., Little, C. & van Bergen, E. Nurture might be nature: cautionary tales and proposed solutions. *NPJ Sci. Learn.* **6**, 2 (2021).
51. Willoughby, E. A., McGue, M., Iacono, W. G., Rustichini, A. & Lee, J. J. The role of parental genotype in predicting offspring years of education: evidence for genetic nurture. *Mol. Psychiatry* **26**, 3896–3904 (2021).
52. de Vries, L. P. et al. Gene-by-crisis interaction for optimism and meaning in life: the effects of the COVID-19 pandemic. *Behav. Genet.* **52**, 13–25 (2022).
53. Purcell, S. Variance components models for gene–environment interaction in twin analysis. *Twin Res.* **5**, 554–571 (2002).
54. van der Sluis, S., Posthuma, D. & Dolan, C. V. A note on false positives and power in $G \times E$ modelling of twin data. *Behav. Genet.* **42**, 170–186 (2012).

55. Molenaar, D., van der Sluis, S., Boomsma, D. I. & Dolan, C. V. Detecting specific genotype by environment interactions using marginal maximum likelihood estimation in the classical twin design. *Behav. Genet* **42**, 483–499 (2012).
56. Jinks, J. L. & Fulker, D. W. Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behavior. *Psychol. Bull.* **73**, 311–349 (1970).
57. van der Sluis, S., Dolan, C. V., Neale, M. C., Boomsma, D. I. & Posthuma, D. Detecting genotype–environment interaction in monozygotic twin data: comparing the Jinks and Fulker test and a new test based on marginal maximum likelihood estimation. *Twin Res. Hum. Genet.* **9**, 377–392 (2006).
58. Plomin, R., DeFries, J. C. & Loehlin, J. C. Genotype–environment interaction and correlation in the analysis of human behavior. *Psychol. Bull.* **84**, 309–322 (1977).
59. Eaves, L. A model for sibling effects in man. *Heredity* **36**, 205–214 (1976).
60. Hunter, M. D. Multilevel modeling in classical twin and modern molecular behavior genetics. *Behav. Genet.* **51**, 301–318 (2021).
61. Tamimy, Z. et al. Multilevel twin models: geographical region as a third level variable. *Behav. Genet.* **51**, 319–330 (2021).
62. Davey Smith, G. & Hemani, G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum. Mol. Genet.* **23**, R89–R98 (2014).
63. Heath, A. C. et al. Testing hypotheses about direction of causation using cross-sectional family data. *Behav. Genet.* **23**, 29–50 (1993).
64. Castro-de-Araujo, L. F. S. et al. MR-DoC2: bidirectional causal modeling with instrumental variables and data from relatives. *Behav. Genet.* **53**, 63–73 (2022).
65. Minică, C. C., Dolan, C. V., Boomsma, D. I., de Geus, E. & Neale, M. C. Extending causality tests with genetic instruments: an integration of Mendelian randomization with the classical twin design. *Behav. Genet.* **48**, 337–349 (2018).
66. Evans, L. M. et al. Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. *Nat. Genet.* **50**, 737–745 (2018).
67. Zaitlen, N. et al. Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. *PLoS Genet.* **9**, e1003520 (2013).
68. Ouwens, K. G. et al. A characterization of cis- and trans-heritability of RNA-seq-based gene expression. *Eur. J. Hum. Genet.* **28**, 253–263 (2020).
69. Hagenbeek, F. A. et al. Heritability estimates for 361 blood metabolites across 40 genome-wide association studies. *Nat. Commun.* **11**, 39 (2020).
70. Hall, J. G. Twinning. *Lancet* **362**, 735–743 (2003).
71. Bulmer, M. G. *The Biology of Twinning in Man* (Clarendon, 1970).
72. Lewis, C. M., Healey, S. C. & Martin, N. G. Genetic contribution to DZ twinning. *Am. J. Med. Genet.* **61**, 237–246 (1996).
73. Meulemans, W. J. et al. Genetic modelling of dizygotic twinning in pedigrees of spontaneous dizygotic twins. *Am. J. Med. Genet.* **61**, 258–263 (1996).
74. Duffy, D. L. & Martin, N. G. The heritability of twinning in seven large historic pedigrees. *Twin Res. Hum. Genet.* **25**, 63–66 (2022).
75. Mbarek, H. et al. Identification of common genetic variants influencing spontaneous dizygotic twinning and female fertility. *Am. J. Hum. Genet.* **98**, 898–908 (2016).
76. van Dongen, J. et al. Identical twins carry a persistent epigenetic signature of early genome programming. *Nat. Commun.* **12**, 5618 (2021).
77. Levi, S. Ultrasonic assessment of the high rate of human multiple pregnancy in the first trimester. *J. Clin. Ultrasound* **4**, 3–5 (1976).
78. Hall, J. G. The mystery of monozygotic twinning I: what can amyoplasia tell us about monozygotic twinning and the possible role of twin–twin transfusion? *Am. J. Med. Genet. A* **185**, 1816–1821 (2021).
79. Eaves, L. J. & Eysenck, H. J. Genetics and the development of social attitudes. *Nature* **249**, 288–289 (1974).
80. Taubman, P. Earnings, education, genetics, and environment. *J. Hum. Resour.* **11**, 447–461 (1976).
81. Cesarini, D. et al. Heritability of cooperative behavior in the trust game. *Proc. Natl Acad. Sci. USA* **105**, 3721–3726 (2008).
82. Hatemi, P. K. et al. Genetic influences on political ideologies: twin analyses of 19 measures of political ideologies from five democracies and genome-wide findings from three populations. *Behav. Genet.* **44**, 282–294 (2014).
83. Williams, F. M. K. et al. Self-reported symptoms of COVID-19, including symptoms most predictive of SARS-CoV-2 infection, are heritable. *Twin Res. Hum. Genet.* **23**, 316–321 (2020).
84. Baird, P. N. & Hysi, P. Twin registries moving forward and meeting the future: a review. *Twin Res. Hum. Genet.* **22**, 201–209 (2019).
85. Geserick, M. et al. Acceleration of BMI in early childhood and risk of sustained obesity. *N. Engl. J. Med.* **379**, 1303–1312 (2018).
86. Elks, C. et al. Variability in the heritability of body mass index: a systematic review and meta-regression. *Front. Endocrinol.* **3**, 29 (2012).
87. Huston, A. C. From research to policy and back. *Child Dev.* **79**, 1–12 (2008).
88. Bird, S., Segall, I. & Lopatka, M. Replication: why we still can't browse in peace—on the uniqueness and reidentifiability of web browsing histories. in *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*, 489–503 (USENIX Association, 2020).
89. Olejnik, L., Castelluccia, C. & Janc, A. Why Johnny can't browse in peace: on the uniqueness of web browsing history patterns. In *5th Workshop on Hot Topics in Privacy Enhancing Technologies (HotPETs 2012)*, hal-00747841 (2012).
90. Eichstaedt, J. C. et al. Facebook language predicts depression in medical records. *Proc. Natl Acad. Sci. USA* **115**, 11203–11208 (2018).
91. Long, E. C. et al. The genetic and environmental contributions to internet use and associations with psychopathology: a twin study. *Twin Res. Hum. Genet.* **19**, 1–9 (2016).
92. Langner, I., Garbe, E., Banaschewski, T. & Mikolajczyk, R. T. Twin and sibling studies using health insurance data: the example of attention deficit/hyperactivity disorder (ADHD). *PLoS ONE* **8**, e62177 (2013).
93. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**, 186–198 (2009).
94. Borsboom, D. & Cramer, A. O. J. Network analysis: an integrative approach to the structure of psychopathology. *Annu. Rev. Clin. Psychol.* **9**, 91–121 (2013).
95. Koenis, M. M. G. et al. Association between structural brain network efficiency and intelligence increases during adolescence. *Hum. Brain Mapp.* **39**, 822–836 (2018).
96. Koenis, M. M. G. et al. Development of the brain's structural network efficiency in early adolescence: a longitudinal DTI twin study. *Hum. Brain Mapp.* **36**, 4938–4953 (2015).
97. Bohlken, M. M. et al. Structural brain connectivity as a genetic marker for schizophrenia. *JAMA Psychiatry* **73**, 11–19 (2016).
98. Olatunji, B. O., Christian, C., Strachan, E. & Levinson, C. A. Central and peripheral symptoms in network analysis are differentially heritable: a twin study of anxious misery. *J. Affect. Disord.* **274**, 986–994 (2020).
99. Forbes, M. K., Wright, A. G. C., Markon, K. E. & Krueger, R. F. Quantifying the reliability and replicability of psychopathology network characteristics. *Multivar. Behav. Res.* **56**, 224–242 (2021).
100. Fried, E. I. & Cramer, A. O. J. Moving forward: challenges and directions for psychopathological network theory and methodology. *Perspect. Psychol. Sci.* **12**, 999–1020 (2017).

101. Zhang, H. The review of transcriptome sequencing: principles, history and advances. *IOP Conf. Ser. Earth Environ. Sci.* **332**, 042003 (2019).
102. Grundberg, E. et al. Mapping *cis*- and *trans*-regulatory effects across multiple tissues in twins. *Nat. Genet.* **44**, 1084–1089 (2012).
103. Wright, F. A. et al. Heritability and genomics of gene expression in peripheral blood. *Nat. Genet.* **46**, 430–437 (2014).
104. Aristizabal, M. J. et al. Biological embedding of experience: a primer on epigenetics. *Proc. Natl Acad. Sci. USA* **117**, 23261–23269 (2020).
105. van Dongen, J. et al. Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nat. Commun.* **7**, 11115 (2016).
106. Li, S. et al. Early life affects late-life health through determining DNA methylation across the lifespan: a twin study. *eBioMedicine* **77**, 103927 (2022).
107. Liu, Y. et al. Quantitative variability of 342 plasma proteins in a human twin population. *Mol. Syst. Biol.* **11**, 786 (2015).
108. Menni, C. et al. Glycosylation of immunoglobulin G: role of genetic and epigenetic influences. *PLoS ONE* **8**, e82558 (2013).
109. Zaytseva, O. O. et al. Heritability of human plasma *N*-glycome. *J. Proteome Res.* **19**, 85–91 (2020).
110. Yang, Q. et al. Metabolomics biotechnology, applications, and future trends: a systematic review. *RSC Adv.* **9**, 37245–37257 (2019).
111. Pool, R. et al. Genetics and not shared environment explains familial resemblance in adult metabolomics data. *Twin Res. Hum. Genet.* **23**, 145–155 (2020).
112. Shin, S.-Y. et al. An atlas of genetic influences on human blood metabolites. *Nat. Genet.* **46**, 543–550 (2014).
113. Bermingham, K. M. et al. Genetic and environmental contributions to variation in the stable urinary NMR metabolome over time: a classic twin study. *J. Proteome Res.* **20**, 3992–4000 (2021).
114. Hagenbeek, F. A. et al. Heritability of urinary amines, organic acids, and steroid hormones in children. *Metabolites* **12**, 474 (2022).
115. Hasin, Y., Seldin, M. & Lusis, A. Multi-omics approaches to disease. *Genome Biol.* **18**, 83 (2017).
116. Duruflé, H. et al. A powerful framework for an integrative study with heterogeneous omics data: from univariate statistics to multi-block analysis. *Brief. Bioinform.* **22**, bbaa166 (2021).
117. Hur, Y.-M., Odintsova, V. V., Ordoñana, J. R., Silventoinen, K. & Willemssen, G. in *Twin Research for Everyone* (eds Tarnoki, A. et al.) 23–50 (Academic Press, 2022).
118. Silventoinen, K. et al. The CODATwins Project: the current status and recent findings of COLlaborative Project of Development of Anthropometrical Measures in Twins. *Twin Res. Hum. Genet.* **22**, 800–808 (2019).
119. Odintsova, V. V. et al. Establishing a twin register: an invaluable resource for (behavior) genetic, epidemiological, biomarker, and ‘omics’ studies. *Twin Res. Hum. Genet.* **21**, 239–252 (2018).
120. Boutwell, B. B., Narvey, C. S., Helton, J. J. & Piquero, A. R. Why twin studies are important for health span science research: the case of maltreatment of aging adults. *BMC Geriatr.* **22**, 943 (2022).
121. Austerberry, C., Mateen, M., Fearon, P. & Ronald, A. Heritability of psychological traits and developmental milestones in infancy: a systematic review and meta-analysis. *JAMA Netw. Open* **5**, e2227887 (2022).
122. Polderman, T. J. C. et al. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* **47**, 702–709 (2015).
123. Hur, Y.-M. et al. A comparison of twin birthweight data from Australia, the Netherlands, the United States, Japan, and South Korea: are genetic and environmental variations in birthweight similar in Caucasians and East Asians? *Twin Res. Hum. Genet.* **8**, 638–648 (2005).
124. Laursen, M. et al. Genetic influence on prolonged gestation: a population-based Danish twin study. *Am. J. Obstet. Gynecol.* **190**, 489–494 (2004).
125. Garrett-Bakelman, F. E. et al. The NASA Twins Study: a multidimensional analysis of a year-long human spaceflight. *Science* **364**, eaau8650 (2019).
126. Vinkhuyzen, A. A. E., Van Der Sluis, S., De Geus, E. J. C., Boomsma, D. I. & Posthuma, D. Genetic influences on ‘environmental’ factors. *Genes Brain Behav.* **9**, 276–287 (2010).
127. Friedman, N. P., Banich, M. T. & Keller, M. C. Twin studies to GWAS: there and back again. *Trends Cogn. Sci.* **25**, 855–869 (2021).
128. Hatemi, P. K. The intersection of behavioral genetics and political science: introduction to the special issue. *Twin Res. Hum. Genet.* **15**, 1–5 (2012).
129. Cesarini, D. & Visscher, P. M. Genetics and educational attainment. *NPJ Sci. Learn.* **2**, 4 (2017).
130. Plomin, R. The next 10 years of behavioural genomic research. *JCPP Adv.* **2**, e12112 (2022).

Acknowledgements

We thank N. G. Martin, C. V. Dolan, J. Couvée and J. van Dongen for their valuable contributions to the draft versions of this manuscript. The current work is supported by the Royal Netherlands Academy of Science Professor Award (no. PAH/6635) to D.I.B. The funder had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

Conceptualization: D.I.B. Funding acquisition: D.I.B. Project administration: F.A.H. and D.I.B. Supervision: D.I.B. Visualization: F.A.H., S. Bruins, D.V.K. and V.V.O. Writing—original draft: all authors. Writing—review and editing: all authors. All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41562-023-01609-6>.

Correspondence and requests for materials should be addressed to Fiona A. Hagenbeek or Dorret I. Boomsma.

Peer review information *Nature Human Behaviour* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature Limited 2023