

Appraisal

Research Note: Twin studies and their value for physiotherapy research

Twin studies have been designed and implemented since the late 19th century. Although relatively new and underused in physiotherapy research, they have made a scientific impact on understanding the possible extent of genetic causes behind human traits.^{1–4} The initial applications of twin studies to health and disease were mainly aimed at elucidating the influences of nature versus nurture. More recently, a variety of modern and simple, yet sophisticated, twin research designs have been developed and applied to unravel causes of and cures for diseases and to make significant contributions to the understanding of how psychological and biological traits are formed. Put simply, the availability of data from a family member, in this context a sibling co-twin who shares some or all genetic material and other familial factors with the relevant individual, offers immense possibilities for understanding genetic and other familial influences, controlling for familial confounding, by applying statistical and modelling approaches that are not possible with a non-twin design. Essentially, twin studies can double or amplify the research horizon and provide novel scientific insights into the health of twins and non-twins.

This Research Note offers an overview of multiple twin designs and their rationales and potential for use in health research. It presents twin designs that are useful for disentangling the contributions of genetic and environmental influences on health and disease, with a focus on the latter, recognising the relevance of modifiable environmental factors for physiotherapy research. It discusses the use of twin studies to identify risk factors for diseases commonly managed by physiotherapists, through powerful adjustments for genetic and familial influences on outcomes and exposures of interest. The emerging use of twins in randomised controlled trials as a promising approach for researchers and clinicians interested in a cost-effective and efficient way to test the efficacy of interventions in physiotherapy is also discussed. When discussing twin research designs, their applications and advantages are presented and possible limitations and challenges in data analysis and aspects of study setup, such as statistical power and recruitment, are addressed. It finishes by describing how registries of twins can make twin research possible and accessible to researchers globally, and the impact of international twin registries networks.

Doubling the value of research: what is unique in twin studies?

The greatest advantage of conducting research with twins is the possibility of accessing data from two family members who are matched by sex, genes and family environments, which are three of the most important determinants of health and disease. Twins are categorised by their zygosity, which captures the degree of genetic similarity between twins within a pair. Twins who both originate from a single zygote are referred to as monozygotic (MZ) twins.⁵ Monozygotic twins are of the same sex, share 100% (or almost 100%) of their genetic variation and are described as identical twin pairs.⁶ Twins who originate from two separate zygotes are referred to as dizygotic (DZ), and are described as non-identical or fraternal twin pairs. Dizygotic

twins share on average 50% of their genetic variation, and are no more genetically similar than two siblings. While approximately 50% of DZ twin pairs are of the same sex, the other half is of the opposite sex, which represents a unique research opportunity to address sex differences that will be described later. Twins, regardless of their zygosity, share a large degree of their early familial environment, starting in the womb. This is an important and sometimes underestimated research concept, which expands beyond the genetic realm. Examples of familial factors (measurable and unmeasurable) that are usually shared by twins early in life include: maternal/uterine environmental factors, parental socio-economic status, diet habits and nutritional factors, or schooling/education. This Research Note further explores the concept of the unique opportunity of controlling for possible confounding by using twin studies, which supports the overall assumption that in twin studies, researchers usually account for factors that are not even predicted or known.

The cornerstone of twin studies is the comparison of similarities and differences within pairs of MZ and DZ twins, as well as the single prospect of matching individuals for a range of known, unknown, measured and unmeasured variables to obtain insights on risk, and treatment estimates for diseases, through well-controlled and powerful data management and modelling. However, the design and statistical implications of the approach is only a partial representation of the potential for twin studies. As will be discussed later, the well-structured organisation and governance framework of twin registries and networks offer an effective (including cost-saving) and efficient (including targeted and agile recruitment) way to conduct research.

The classic twin design

The most traditional twin study uses the classical twin design to estimate the genetic and environmental components of variance of human traits, including those related to disease and health behaviour.² Through advanced statistical approaches, the environmental influences can be partitioned into common environmental influences shared by twins early in life, and non-shared, unique environmental influences. Statistically, this is initially achieved by estimating the correlation or covariance in the trait or condition of interest separately for MZ and DZ twin pairs. A higher correlation in MZ pairs compared with DZ pairs is consistent with genetic effects explaining a component of variation, even though genes are not measured or identified directly. A sub-product of this analysis is the heritability estimate, which reflects the genetic variation as a proportion of the total variance of the trait. It is important to note that heritability should not be used to infer genetic liability or causation of a disease or trait. It is a measure of the proportion of the variation in disease or trait, in a given population at a given time, which is attributed to genetic influences.

This design is recommended as a preliminary step before devoting substantial resources and research focus on a set of risk factors (environmental or genetic) that contribute to diseases (Table 1). Using a classical twin design approach, researchers have claimed that the

Table 1
Twin designs, applications, advantages, and disadvantages.

Design	Research target	Applications/examples	Advantages	Disadvantages
Classical twin design	Genetic and environmental influences on diseases and traits	Genetics can influence up to 63% (95% CI 46 to 77%) of the variance of prevalence of COPD ⁷	All twins, irrespective of exposure or outcome, can be included in analysis	Assumptions need to be met such as: equal shared environment between MZ and DZ twins; MZ twin pairs share 100% of their genetic variation whereas DZ twin pairs share 50% of their genetic variation
Co-twin design	Causes of diseases	Depression does not increase the risk of low back pain. The association between these conditions is confounded ¹⁰	Provides strong control of confounding	The pool of potential study participants may be limited by the requirement of discordant outcomes
Twin RCT	Efficacy of interventions	Calcium supplementation in adolescence has little effect on bone density ¹²	Increases power. Requires smaller sample sizes	Need for both twins in a pair to be concordant for the condition of interest
Epigenetic co-twin design	Effects of genes and their activity on disease and traits	Birth weight is associated with epigenetic differences in growth and metabolism genes ¹⁵	Enables study of gene activity independent of DNA sequence	Does not address shared factors unless the between-pair analysis is also used
Opposite sex DZ twin design	Effect of gender on disease prevalence	Sex is not associated with the prevalence of low back pain (manuscript in preparation)	Controls for 50% of genetics and 100% of early shared environment on the association between sex and condition	Limited availability of study participants reflecting strict inclusion criteria of twins

COPD = chronic obstructive pulmonary disease, RCT = randomised controlled trial, DZ = dizygotic, MZ = monozygotic. Adapted from Scurrah and Hopper (with permission).

proportion of the variance of the prevalence of chronic obstructive pulmonary disease (COPD) that is attributable to genetic influences is 63% (95% CI 46 to 77%).⁷ It should be noted that the remainder of the variance in COPD is still largely dependent on stochastic and environmental – including perhaps modifiable – factors that are managed by physiotherapists, such as lung function and patients' engagement in physical activity. One of the main advantages of the classic twin design is that all types of twins can be included in the analysis, regardless of their zygosity or sex, and their measured outcomes or exposures (Table 1). However, this approach is based on the assumption that effects of the early familial environment are equally shared by MZ and DZ twins, and this might not apply, particularly to DZ opposite-sex pairs.⁸ Ideally, this assumption should be tested as much as possible using, for example, data on length of cohabitation and time living apart,⁹ and not taken for granted. Another relevant assumption of the design is that MZ twin pairs share 100% of their genetic variation, whereas DZ twin pairs share 50% of their genetic variation (Table 1).

The co-twin control study

One of most useful types of twin design for epidemiological research is the co-twin control study. This design is applicable to research questions of causal hypotheses and prognosis of conditions, and due to its clinical nature, it has enormous potential for clinicians and researchers in physiotherapy interested in identifying modifiable factors that can be targeted through preventative strategies. Twin pairs are selected who are discordant for the outcome of interest (outcome discordant) or for the exposure of interest (exposure discordant). Conditional regression models are usually fit to investigate the association between exposures and outcomes. Often a high degree of control for confounding is achieved with this approach, which makes it a very powerful tool. Although the analyses can be conducted with DZ twin pairs, it is usually more efficient with MZ twin pairs. Monozygotic pairs discordant for the outcome or exposure are matched for measured and unmeasured factors, and therefore natural adjustment or control for confounding is perfectly achieved for genetic factors, sex, age, calendar year, and season of birth, as well as for components of the early shared environment (including, but not limited to, maternal/uterine factors, parental socio-economic status, nutritional factors and schooling). A cross-sectional design can be used, but stronger evidence consistent with causation can be obtained with a longitudinal design (eg, MZ twin pairs discordant for an exposure followed up to the first outcome development in a twin). The obvious advantage of this approach is the degree of control for possible confounding. One

disadvantage is the need for MZ pairs to fulfil the discordance criteria (exposure or outcome). This is where the existence of twin registries in Australia and across the world comes into play.

Using the co-twin design it has been found that depression has no significant causal effect on low back pain, as previously thought. The association between these conditions is possibly driven by familial confounding (ie, familial factors that affect both depression and low back pain).¹⁰ Others have confirmed the findings of previous non-twin studies, with less adjustments for confounding, that low birth weight has a causal effect on cerebral palsy – a condition commonly treated by physiotherapists working in paediatrics.³

Twin randomised controlled trials

One of the most promising twin study designs for clinical research in physiotherapy is the twin randomised controlled trial. The design is very flexible, and different randomisation strategies can be used depending on the research question addressed. The most efficient procedure to assess efficacy of treatment using twin trials is within-pair randomisation. With this approach, MZ twins within a pair are randomised to two different interventions. Treatment effects are compared within all pairs across the sample. Because of the high levels of comparability between intervention groups and control for confounding, statistical power can be substantially increased compared with trials of unrelated singletons.¹¹ An interesting twin randomised trial found that calcium supplementation has little effect on bone density during adolescence.¹² Notice that the twin pairs can be swapped and the trial repeated using a cross-over design to ensure that twins within a pair are treated fairly (provided that the usual requirement of no period effects is satisfied). Another example is an ongoing trial that is testing a sleep intervention to improve symptoms in patients with low back pain.¹³ Additional types of twin studies and their application are listed in Table 1.

Twin registries, networks, and the Australian Centre of Research Excellence in Twin Research

Twin registries play a crucial role in supporting researchers to conduct research with twins. They facilitate the implementation of research projects through the identification and recruitment of twin participants through population-based and volunteer-based methods. Countries such as Australia (> 70 000 twins in the registry), USA, Brazil, Denmark, China, South Korea, Finland, UK, Pakistan, Netherlands and Norway have well established twin registries, with

some storing biological samples and twin self-reported data.¹⁴ The International Network of Twin Registries is an organisation of more than 30 registries across the world that was formed to enable the harmonisation and cataloguing of twin data for new studies. The International Society for Twin Studies is another important scientific organisation that supports twin research. More recently, a group of leaders in twin research has been funded by the Australian National Health and Medical Research Council to establish the first Centre of Research Excellence in Twin Research, with the main aim of fostering collaboration, implementation of twin research, and effective translation of findings from twin research to all Australians. Collectively, these organisations represent rich resources that support researchers, including in physiotherapy, to conduct high-quality twin studies that could lead to breakthroughs in the field.

Competing interests: Nil.

Sources of support: Nil.

Acknowledgements: PHF is funded by a NHMRC Career Development Fellowship and a Sydney University SOAR Fellowship. JLH is funded by a NHMRC Senior Principal Research Fellowship.

Provenance: Invited. Peer reviewed.

Correspondence: Paulo Ferreira, Physiotherapy Department, Faculty of Health Sciences, University of Sydney, Australia. Email: Paulo.Ferreira@sydney.edu.au

Paulo Ferreira^a, Jeffrey Craig^b and John L Hopper^c

^aFaculty of Health Sciences, University of Sydney

^bCentre for Molecular and Medical Research, School of Medicine, Deakin University

^cCentre for Epidemiology & Biostatistics, Melbourne School of Population & Global Health, Faculty of Medicine, Dentistry & Health Sciences, University of Melbourne, Australia

References

- Galton F. *The History of Twins as a Criterion of the Relative Powers of Nature and Nurture*. Trubner and Co; 1876.
- Boomsma D, et al. *Nat Rev Genet*. 2002;3:872–882.
- Hrubec Z, Robinette C. *N Engl J Med*. 1984;310:435–441.
- Martin N, et al. *Nat Genet*. 1997;17:387–392.
- Hall J. *Lancet*. 2003;362:735–743.
- Mendel G. *Experiments in plant hybridisation*. Edinburgh: Oliver & Boyd; 1965.
- Ingebrigtsen T, et al. *Respir Med*. 2010;104:1890–1895.
- Feldman M, Lewontin R. *Am Assoc for the Advancement of Science*. 1975:1163.
- Hopper J, Mathews J. *Ann Hum Genet*. 1982;46:373–383.
- Pinheiro M, et al. *Pain*. 2015;156:496–503.
- Yelland L, et al. *Stat Med*. 2017;36:1227–1239.
- Nowson C, et al. *Osteoporos Int*. 1997;7:219–225.
- Pinheiro M, et al. *Twin Res Hum Genet*. 2016;19:492–501.
- Hur Y, Craig J. *Twin Res Hum Genet*. 2013;16:1–12.
- Gluckman P. *Nat Rev Endocrinol*. 2012;8:74–76.