

Article

Conducting Clinical Trials in Twin Populations: A Review of Design, Analysis, Recruitment and Ethical Issues for Twin-Only Trials

Lisa N. Yelland^{1,2}, Katrina J. Scurrah³, Paulo Ferreira⁴, Lucas Calais-Ferreira³, Monica Rankin⁵, Jane Denton⁶, Merryl Harvey⁷, Katherine J. Lee^{8,9}, Evie Kendal¹⁰ and Jeffrey M. Craig^{9,11,12}

¹Women and Kids Theme, South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia, ²School of Public Health, The University of Adelaide, Adelaide, South Australia, Australia, ³Twins Research Australia and Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia, ⁴Faculty of Medicine and Health, University of Sydney, Sydney, New South Wales, Australia, ⁵International Council of Multiple Birth Organisations, Melbourne, Victoria, Australia, ⁶Multiple Births Foundation, Queen Charlotte's & Chelsea Hospital, London, UK, ⁷Faculty of Health, Education and Life Sciences, Birmingham City University, Birmingham, UK, ⁸Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute, Melbourne, Victoria, Australia, ⁹Department of Paediatrics, The University of Melbourne, Melbourne, Victoria, Australia, ¹⁰School of Health Sciences and Biostatistics, Swinburne University of Technology, Melbourne, Victoria, Australia, ¹¹IMPACT – the Institute for Mental and Physical Health and Clinical Translation, School of Medicine, Deakin University, Geelong, Victoria, Australia and ¹²Murdoch Children's Research Institute, Melbourne, Victoria, Australia

Abstract

Although twins often participate in medical research, few clinical trials are conducted entirely in twin populations. The purpose of this review is to demonstrate the substantial benefits and address the key challenges of conducting clinical trials in twin populations, or 'twin-only trials'. We consider the unique design, analysis, recruitment and ethical issues that arise in such trials. In particular, we describe the different approaches available for randomizing twin pairs, highlight the similarity or correlation that exists between outcomes of twins, and discuss the impact of this correlation on sample size calculations and statistical analysis methods for estimating treatment effects. We also consider the role of both monozygotic and dizygotic twins for studying variation in outcomes, the factors that may affect recruitment of twins, and the ethics of conducting trials entirely in twin populations. The advantages and disadvantages of conducting twin-only trials are also discussed. Finally, we recommend that twin-only trials should be considered more often.

Keywords: Analysis; clinical trials; design; ethics; recruitment; twins

(Received 20 October 2021; accepted 25 November 2021; First Published online 25 January 2022)

Twins have participated in medical research for more than a century (Boomsma et al., 2002; Craig et al., 2020; Hrubec & Robinette, 1984; Martin et al., 1997). Previous reviews have highlighted the unique role twins play in understanding genetic and environmental contributions to disease and have provided many examples of the valuable contribution twin studies have made to medical research (Craig et al., 2020; Hrubec & Robinette, 1984). Studies in twins can be used to address questions of gene–environment interactions (Buil et al., 2015) and cause versus association (Sjölander et al., 2012) and have been used in modern specialist areas of medical science such as epigenetics (Bell & Saffery, 2012), stem cells (Hibaoui et al., 2014) and microbiome research (Smith et al., 2013). However, clinical trials conducted entirely in twin populations, or 'twin-only trials', have been limited to date. A recent review on the participation of twins in clinical trials found that among 186,027 clinical trials registered in

ClinicalTrials.gov, only six trials restricted participation to twins (Sumathipala et al., 2018). This finding highlights the potential for conducting more clinical trials in this unique population.

The purpose of the current narrative review is to demonstrate the substantial benefits and address the key challenges of conducting twin-only trials. We consider the design, analysis, recruitment and ethical issues that arise in such trials, including how to randomize twins, the impact of twins on sample size calculations and statistical analysis methods for estimating treatment effects, the utility of monozygotic (MZ) and dizygotic (DZ) twins for studying variation in outcomes, factors affecting recruitment, and the unique ethical considerations. We also discuss the advantages and disadvantages of conducting twin-only trials. For simplicity, we focus on phase III, parallel group randomized trials with an intervention and a control group, although many of the issues discussed also relate to more complex designs.

Author for correspondence: Jeffrey Craig, Email: jeffrey.craig@deakin.edu.au

Cite this article: Yelland LN, Scurrah KJ, Ferreira P, Calais-Ferreira L, Rankin M, Denton J, Harvey M, Lee KJ, Kendal E, and Craig JM. (2021) Conducting Clinical Trials in Twin Populations: A Review of Design, Analysis, Recruitment and Ethical Issues for Twin-Only Trials. *Twin Research and Human Genetics* 24: 359–364, <https://doi.org/10.1017/thg.2021.52>

© The Author(s), 2022. Published by Cambridge University Press on behalf of International Society for Twin Studies. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Design Issues

Randomization of Twins

When trials are conducted in twin populations, participants from a twin pair can be randomized to the intervention or control group

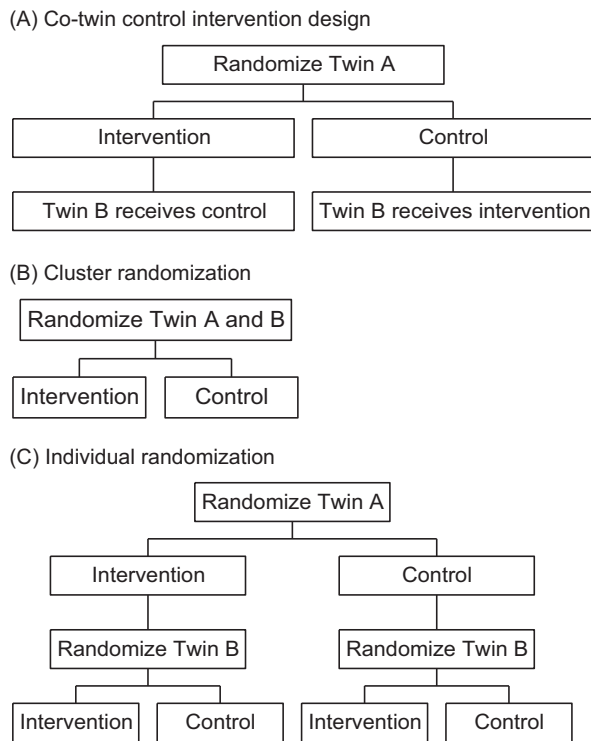


Fig. 1. Different methods of randomization for twins in twin-only trials

in one of three ways: using the co-twin control intervention design, cluster randomization or individual randomization (see Figure 1).

The 'co-twin control intervention design' (Plomin & Haworth, 2010) involves assigning one twin from each pair to the intervention group and the other twin to the control group. It is an example of the classical matched pairs design and has long been recognized as the most powerful design for twin-only trials and an efficient alternative to studies in unrelated individuals (Christian & Kang, 1972) because treatments can be compared within twin pairs. It provides near-perfect control (for MZ twins, who share almost all their genetic material) or partial control (for DZ twins, who share 50% of their genetic variation on average) for key covariates, including those that are not measured. This removes some of the noise from the treatment effect estimate and hence this design requires the smallest sample size (Yelland et al., 2017). The reduced sample size increases the feasibility of recruiting from the relatively small twin population and is most useful in settings where the trial is very expensive, or few twins meet the inclusion–exclusion criteria. While the co-twin control intervention design is clearly the ideal design for twin-only trials for the reasons mentioned above and has been chosen specifically for its efficiency (e.g., Pinheiro et al., 2016), it is also the least preferred approach among potential participants and their caregivers, primarily due to the potential psychological trauma associated with twins in different treatment groups achieving different outcomes (Bernardo et al., 2015). Recruitment may therefore prove difficult for trials using this design and alternative randomization methods may be needed.

Cluster randomization, where both twins in a pair are assigned to the same treatment group, is useful in settings where there is a high risk of treatment group contamination (Donner & Klar, 2000). For example, if twins are assigned to different treatment groups in an unblinded trial, the twin assigned to the intervention group may discuss details or share some part of the intervention with the twin assigned to the control group. Treatment group

contamination can also occur in blinded trials involving active and placebo treatments that are identical in appearance and taken in the home environment, since treatments could get mixed up for twins assigned to different groups who are living together. Both parents of twins and adult twins have a strong preference for cluster randomization due to the increased chance that both twins will achieve similar outcomes (Bernardo et al., 2015) and hence recruitment may be easier for trials that utilize this method of randomization. In some trials, cluster randomization is the only option, as both twins will necessarily receive the same treatment due to the nature of the intervention (e.g., a parental tool designed to increase children's acceptance of vegetables; Fildes et al., 2014).

Individual randomization, where each twin is randomized independently, is useful in settings where there is a need to balance the risk of treatment group contamination and the size of the trial. In such a design, approximately half of all twin pairs will be assigned to the same treatment group and hence treatment group contamination is not an issue among these pairs. The remaining twin pairs will be assigned to different treatment groups and hence treatments can be compared within these pairs. This design is more powerful than the cluster-randomized design but not as powerful as the co-twin control intervention design (Yelland et al., 2017).

In practice, the best way to randomize twins depends on the individual trial, and the main advantages and disadvantages of each approach are summarized in Table 1. Among 50 twin-only trials identified in a recent review, only 6% used individual randomization, while randomizing twins to different treatment groups (66%) and cluster randomization (26%) were more common (Sumathipala et al., 2018).

Sample Size Calculations for Estimating Treatment Effects

At the trial design stage, sample size calculations are typically performed to estimate the number of participants needed to detect a particular treatment effect based on a specified level of power and significance, as well as a set of assumptions about the data (Kirby et al., 2002). Standard sample size calculations assume that observations are independent, but observations collected in a twin-only trial are expected to be similar or correlated within twin pairs. Ignoring this correlation in sample size calculations can result in a trial that is underpowered (for a trial using cluster randomization) or overpowered (for a trial using the co-twin control intervention design; Yelland et al., 2017), both of which are a poor use of resources and may therefore be considered unethical (Carlin & Doyle, 2002).

Sample size calculations for twin-only trials can be performed using a two-step process. First, the sample size is calculated using standard methods that assume the outcomes of all trial participants are independent (Kirby et al., 2002). Second, this sample size is multiplied by a quantity known as the design effect, which measures the degree of inflation required to account for dependence in the data. The design effect depends on trial-specific factors, including how twins will be randomized, how the data will be analyzed and the expected correlation between outcomes of twins. Equations for the design effect have been published elsewhere (Yelland et al., 2017), and sample size calculators that account for the correlation between outcomes of twins are freely available (Yelland et al., 2018).

Including MZ and/or DZ Twins

An important decision to make when designing a twin-only trial is which type(s) of twins to recruit. MZ pairs are more closely matched than DZ pairs. In a co-twin control intervention design,

Table 1. Advantages and disadvantages of different methods of randomization for twins in twin-only trials

Method of randomization	Advantages	Disadvantages
Co-twin control intervention design	<ul style="list-style-type: none"> • Most powerful method requiring the smallest sample size • Ensures balanced treatment groups on many covariates 	<ul style="list-style-type: none"> • High risk of treatment group contamination • Least preferred method by parents of twins and adult twins
Cluster randomization	<ul style="list-style-type: none"> • Low risk of treatment group contamination • Most preferred method by parents of twins and adult twins 	<ul style="list-style-type: none"> • Least powerful method requiring the largest sample size
Individual randomization	<ul style="list-style-type: none"> • More powerful than cluster randomization • More acceptable to parents of twins and adult twins than the co-twin control intervention design 	<ul style="list-style-type: none"> • Less powerful than the co-twin control intervention design • Less acceptable to parents of twins and adult twins than cluster randomization

each MZ pair may be considered an approximate counterfactual, as this is the closest we can get in real life to both treating and not treating the same individual at the same time. The strong correlation between MZ twins makes this a highly efficient design. Although twins in DZ pairs are not as closely matched as MZ pairs, and hence provide a less ideal real-life counterfactual, the matching may still be relatively close. Including DZ twins will increase the required sample size compared to a trial restricted to MZ twins due to the lower correlation (Yelland et al., 2017) but will help with meeting recruitment targets. Twin-only trials most commonly involve MZ twins only, followed by MZ and DZ twins, with few trials restricted to DZ twins (Sumathipala et al., 2018).

Analysis Issues

Analysis Methods for Estimating Treatment Effects

When analyzing data collected in a clinical trial, the primary aim is typically to estimate the effect of treatment on a set of prespecified outcomes. Standard methods of analysis for comparing these outcomes between treatment groups often rely on an assumption of independence between outcomes of all trial participants, which is violated in twin-only trials. Failure to account for the correlation between outcomes of twins in the analysis can result in confidence intervals for treatment effects that are too narrow or too wide, and false conclusions about the effectiveness of the intervention (Carlin et al., 2005; Yelland et al., 2011).

Treatment effects can be estimated in twin-only trials using a range of analysis approaches, including regression models fitted using the generalized estimating equations method of estimation or mixed-effects models. Both approaches are appropriate for analyzing many forms of clustered data, not just data collected on twins, and are discussed in detail elsewhere. Briefly, generalized estimating equations account for the correlation between outcomes of twins implicitly through specification of an assumed pattern of correlation for the data, known as a working correlation structure (Hardin & Hilbe, 2013; Zeger & Liang, 1986). In contrast, mixed-effects models account for the correlation between

outcomes of twins explicitly by including a random effect for each cluster in the analysis model that represents the unique effect of each twin pair on the outcome (Carlin et al., 2005). Simpler analysis methods, such as the paired *t* test, may also be applicable in some settings, but do not provide the same flexibility to control for or explore the effects of other covariates on outcomes.

Analysis Methods for Studying Outcome Variation

A major benefit of conducting twin-only trials is that the data can be used to study the variation and covariation in outcomes, in addition to the effect of the treatment on outcomes. The estimated variances and covariances can be used to make inferences about random variation in outcomes, the influence of genes on outcomes, and gene–treatment interactions, as illustrated below. We begin by discussing the information that can be gained from the co-twin control intervention design. We then describe how the classic twin model can be applied with the co-twin control intervention design and cluster randomization. We conclude by considering individual randomization and the role of opposite-sex twins.

Individuals in a pair assigned to different treatment groups.

If the co-twin control intervention design is used, the trial can provide information about variation in outcomes between similar individuals when treated and untreated. If residual variation differs between treated and untreated individuals after adjusting for mean differences due to treatment effects, this may provide useful information about the utility and effects of the treatment. For example, in an early co-twin control intervention design involving 20 pairs of MZ girls who were randomized to receive a 6-month daily calcium and vitamin D supplement or a placebo (Greene & Naughton, 2011), the standard deviation for some measurements was larger in the supplement group versus the placebo (e.g., 52.3 mm² vs. 39.1 mm² for trabecular area). If significantly different, this would be consistent with variation in the size of the treatment effect between individuals (although this was not formally tested).

The classic twin model for the co-twin control intervention design.

The classic twin model (Fisher, 1951; Hopper & Mathews, 1982) can be applied in twin-only trials, including both MZ and DZ twins, to make inferences about the causes of variability in an outcome. If there is evidence that the correlation between MZ twins is greater than the correlation between DZ twins, this is consistent with additive genetic effects ($A > 0$) and genetic effects contributing to variation in outcomes under the equal environments assumption (Hopper, 2000; Scurreh & Hopper, 2019). If the MZ and DZ correlations are similar, this is consistent with either $A = 0$ or unequal environmental covariance for MZ and DZ pairs, although there is no way to formally test which of these explanations is more appropriate. This type of analysis was used in a recent co-twin control intervention design involving 44 twin pairs (34 MZ and 10 DZ pairs) who were randomly assigned to an 8-week low-fat or high-fat diet (Costanzo et al., 2018). The trial investigated the genetic and environmental factors influencing sensitivity to fatty acid taste at baseline. Baseline correlations for MZ and DZ pairs were similar ($r_{MZ} = .33$, $r_{DZ} = .29$, p value for difference = .41), suggesting that environment rather than genetic factors is the primary influencer of fat taste sensitivity.

The classic twin model for cluster-randomized trials. When cluster randomization is used, the classic twin model can be used to study variation in outcomes after adjustment for treatment and

measured covariate effects, as well as variation in baseline measures. This approach was used in a recent cluster-randomized crossover trial (the STRUETH trial) to assess whether genetic effects influence variation in response to exercise, and whether the contribution of genetic effects depends on the type of exercise (Marsh et al., 2020a, 2020b; Thomas et al., 2021). Twin pairs (30 MZ and 12 DZ pairs) were randomized together to either resistance or endurance training, trained together for 3 months, underwent a washout period, then crossed over to the other training regime for 3 months. For resting heart rate in response to endurance exercise, there was low correlation between outcomes for MZ pairs ($r_{MZ} = .12$), suggesting that within-subject variability (i.e., variation that is not due to shared twin factors, such as genetic or environmental factors) may be high and within-subject repeatability may be low (Marsh et al., 2020b). If there is substantial between-subject variation for two individuals with identical genes and very similar environments, there is likely to be substantial variation in outcomes for the same individual experiencing the same treatment at different times. This finding is strengthened if DZ twins are included in the trial and the correlation between DZ twins is lower than the correlation between MZ twins, as seen in the STRUETH trial ($r_{DZ} < .01$ for resting heart rate in response to endurance exercise). The other plausible explanation for such a finding is that a covariate that frequently differs between twins in a pair and that is associated with the outcome has not been adjusted for. In contrast, if both MZ and DZ correlations are high, this suggests that within-subject repeatability is high and within-subject variability is low. If the MZ correlation is higher than the DZ correlation, as it was for baseline V02Max ($r_{MZ} = .92$ and $r_{DZ} = .78$; Thomas et al., 2021), this is consistent with genetic influences on variation in outcome if equal environmental correlations are assumed for MZ and DZ twins. When the outcome is a change from baseline, if the MZ correlation is higher than the DZ correlation and the treatment can be considered an environmental (i.e., nongenetic) factor, this suggests a gene–environment interaction.

Twins randomized independently. Under this design, all of the above analysis approaches are theoretically possible, but in practice, power is only likely to be sufficient for large trials (see Hopper, 2000, for discussion on the number of pairs needed to detect shared environmental effects). Few twin-only trials randomize twins independently (Sumathipala et al., 2018) and we are not aware of any examples of this design being used to study outcome variation.

Role of opposite-sex twins. Inclusion of opposite-sex twin pairs in a trial enables assessment of whether the correlation between outcomes differs by sex. If the correlation for same-sex DZ twins is higher than the correlation for opposite-sex DZ twins, this suggests that an interaction between sex and unmeasured genetic or environmental effects may be present. If the correlation for female DZ pairs is different from the correlation for male DZ pairs, this may suggest that different environmental factors affect the outcomes in females and males. However, most twin-only trials have either excluded opposite-sex twins (e.g., Marsh et al., 2020b) or had too few opposite-sex pairs to enable these types of analyses (e.g., Costanzo et al., 2018).

Recruitment Issues

One challenge that arises when recruiting for twin-only trials is identifying potential participants, as twins are a relatively small

subgroup of the population. This may be overcome through the involvement of individual twin registries such as Twins Research Australia (Murphy et al., 2019), the International Network of Twin Registries (Buchwald et al., 2014) or multiple birth associations. A recent review of 50 twin-only trials found that 16% used a twin registry to support recruitment (Sumathipala et al., 2018), highlighting the utility of this recruitment strategy. Since both twins are required to participate in a twin-only trial, it is important to minimize the use of exclusion criteria that may render one twin ineligible.

Another recruitment challenge in twin-only trials relates to obtaining consent. Research suggests that consent of a twin is substantially more likely if their co-twin consents (Ullemar et al., 2015) and that twins or their caregivers are more likely to consent if both twins will receive the same treatment, at least in the neonatal setting (Bernardo et al., 2015). Consent is especially complex when the twins are minors and their caregivers are the decision-makers due to concerns over the potential outcomes of treatment and the vulnerability of the children. Further research is needed to explore the full range of factors that influence the involvement of twins from all age groups in clinical trials.

Ethical Issues

As with all research involving human participants, the ethical principles of informed consent and justice must guide the recruitment of twins into clinical trials (Beauchamp & Childress, 2019). However, there are additional issues unique to twins that researchers should be mindful of when planning a twin-only trial. First, researchers have an obligation to ensure that individuals are not unduly influenced in their decision to participate in a trial, and that there are no potentially coercive elements to the recruitment procedure. In the case of twin-only trials, if one twin wishes to enroll then this might exert undue pressure on the other twin to join, since the refusal of one twin necessarily excludes the other from participation. Second, twins or their caregivers may feel an unusual degree of pressure to enroll in a trial, given that many benefits of involving twins in research are likely known to the public.

Twin-only trials may raise justice concerns in recruitment. It is generally considered unethical for the burden of research to be borne by one population, while the benefits are conferred elsewhere. A classic example is the use of research participants in low- and middle-income countries to test drugs that are then sold to citizens in highly developed countries (Zumla & Costello, 2002). For research that is primarily aimed at benefitting twins, there is no issue in exclusively recruiting twins. However, for research that is intended to benefit the general population, it is necessary to ensure that twin populations are not being unfairly burdened. In some cases, this might entail some form of compensation for assuming the risks of research on behalf of the wider population, which may reduce the likelihood that twins feel their special similarity is being used as a commodity.

The ethical principles of beneficence and nonmaleficence aim to ensure that clinical trials are designed to yield the best possible benefit to society while avoiding harm to participants. Given the methodological strengths of twin-only trials, these studies clearly satisfy the first of these principles. The need to avoid harm is the same as in all clinical trials, but with additional considerations related to how the twins will be randomized. While all methods of randomization may be deemed ethical under an assumption of clinical equipoise, the psychological implications of assigning twins to different treatment groups should be taken into

consideration, as some participants may experience psychological trauma if twins receive different treatments and achieve different outcomes (Bernardo et al., 2015). Choosing a method of randomization that is more acceptable to participants may be the more ethical approach (see Figure 1 and Table 1).

Advantages and Disadvantages of Twin-Only Trials

The most obvious reason for conducting a twin-only trial is that twins are the target population for the treatment of interest. For example, a new intervention designed to improve outcomes for infants affected by twin-to-twin transfusion syndrome would necessarily be conducted in twins. However, there are many known or perceived advantages of conducting twin-only trials when the target population extends beyond twins. First, enrolling twins may provide some economic advantage if it is faster to recruit and collect household-level characteristics for a set of twins than two unrelated individuals. Second, twins are generally representative of the broader population (Andrew et al., 2001), and hence treatment effects observed in twins may be extrapolated to other individuals. Third, compliance may be better in twins than singletons if twins encourage each other to follow their treatment schedule or complete outcome assessments, though conversely, noncompliance in one twin could increase the risk of noncompliance in the other twin. Fourth, in trials using the co-twin control intervention design, twins provide a natural control for many factors that influence outcomes and hence a smaller sample size is required to achieve the same power as a trial involving unrelated individuals (Yelland et al., 2017). Finally, twin-only trials can provide useful information about variation in outcomes in addition to treatment group effects. The value of this information needs to be weighed against the increase in sample size required to conduct a twin-only trial if twins will be cluster randomized and the intervention is not specifically targeted at twins.

There are also some known or perceived disadvantages of conducting twin-only trials. First, withdrawals may be a larger problem when studying twins, since the withdrawal of one twin could lead to the withdrawal of the other twin. Second, treatment group contamination may occur within twin pairs assigned to different treatment groups, which can lead to underestimation of the treatment effect and hence an increased risk of missing an effective intervention (Torgerson, 2001). Third, sample size calculations and analysis methods are more complex for trials conducted in twins rather than singletons due to the correlation between outcomes of twins (Carlin et al., 2005; Yelland et al., 2018). Finally, identification and recruitment of twins may be more difficult and costly than singletons due to the low prevalence of twins in most disease groups and the potential for adult twins to live in different geographic regions.

Further research is needed to explore these advantages and disadvantages and to identify additional advantages and disadvantages of conducting twin-only trials. The benefits and challenges of conducting a twin-only trial should be carefully considered in the context of the individual trial when deciding whether recruitment should be restricted to twins.

Conclusions

Given the potential advantages of conducting clinical trials entirely in twin populations, we recommend that twin-only trials be considered to complement and contribute valuable additional information to trials involving singletons. We have outlined the design, analysis, recruitment and ethical issues that should be

considered when conducting twin-only trials and discussed the advantages and disadvantages of undertaking such trials. In conclusion, conducting clinical trials entirely in twin populations can add important insights into the action of specific treatments and interventions for the benefit of all of society and should be considered more often.

Acknowledgments. The authors thank Professor John Hopper for his contribution to useful discussions relating to this review.

Financial Support. This work was supported by the Australian National Health and Medical Research Council (L.Y., Early Career Fellowship APP1052388), (K.S. and L.C., Centre of Research Excellence APP1079102), (P.F., Career Development Fellowship APP1144311) and (K.L., Career Development Fellowship APP1053609).

Conflict of Interest. None.

References

- Andrew, T., Hart, D. J., Snieder, H., de Lange, M., Spector, T. D., & MacGregor, A. J. (2001). Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. *Twin Research and Human Genetics*, 4, 464–477.
- Beauchamp, T. L., & Childress, J. F. (2019). *Principles of biomedical ethics* (8th ed.). Oxford University Press.
- Bell, J. T., & Saffery, R. (2012). The value of twins in epigenetic epidemiology. *International Journal of Epidemiology*, 41, 140–150.
- Bernardo, J., Nowacki, A., Martin, R., Fanaroff, J. M., & Hibbs, A. M. (2015). Multiples and parents of multiples prefer same arm randomization of siblings in neonatal trials. *Journal of Perinatology*, 35, 208–213.
- Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews Genetics*, 3, 872–882.
- Buchwald, D., Kaprio, J., Hopper, J. L., Sung, J., Goldberg, J., Fortier, I., Busjahn, A., Sumathipala, A., Cozen, W., Mack, T., Craig, J. M., & Harris, J. R. (2014). International network of twin registries (INTR): Building a platform for international collaboration. *Twin Research and Human Genetics*, 17, 574–577.
- Buil, A., Brown, A. A., Lappalainen, T., Vinuela, A., Davies, M. N., Zheng, H. F., Richards, J. B., Glass, D., Small, K. S., Durbin, R., Spector, T. D., & Dermitzakis, E. T. (2015). Gene-gene and gene-environment interactions detected by transcriptome sequence analysis in twins. *Nature Genetics*, 47, 88–91.
- Carlin, J. B., & Doyle, L. W. (2002). Statistics for clinicians: 7: Sample size. *Journal of Paediatrics and Child Health*, 38, 300–304.
- Carlin, J. B., Gurrin, L. C., Sterne, J. A., Morley, R., & Dwyer, T. (2005). Regression models for twin studies: a critical review. *International Journal of Epidemiology*, 34, 1089–1099.
- Christian, J. C., & Kang, K. W. (1972). Efficiency of human monozygotic twins in studies of blood lipids. *Metabolism*, 21, 691–699.
- Costanzo, A., Nowson, C., Orellana, L., Bolhuis, D., Duesing, K., & Keast, R. (2018). Effect of dietary fat intake and genetics on fat taste sensitivity: A co-twin randomized controlled trial. *American Journal of Clinical Nutrition*, 107, 683–694.
- Craig, J. M., Calais-Ferreira, L., Umstad, M. P., & Buchwald, D. (2020). The value of twins for health and medical research: A third of a century of progress. *Twin Research and Human Genetics*, 23, 8–15.
- Donner, A., & Klar, N. (2000). *Design and analysis of cluster randomization trials in health research*. Arnold.
- Fildes, A., van Jaarsveld, C. H. M., Wardle, J., & Cooke, L. (2014). Parent-administered exposure to increase children's vegetable acceptance: A randomized controlled trial. *Journal of the Academy of Nutrition and Dietetics*, 114, 881–888.
- Fisher, R. A. (1951). Limits to intensive production in animals. *British Agricultural Bulletin*, 4, 217–218.
- Greene, D. A., & Naughton, G. A. (2011). Calcium and vitamin-D supplementation on bone structural properties in peripubertal female identical twins: A randomised controlled trial. *Osteoporosis International*, 22, 489–498.

- Hardin, J. W., & Hilbe, J. M. (2013). *Generalized estimating equations* (2nd ed.). Chapman & Hall/CRC.
- Hibaoui, Y., Grad, I., Letourneau, A., Sailani, M. R., Dahoun, S., Santoni, F. A., Gimelli, S., Guipponi, M., Pelte, M. F., Bena, F., Antonarakis, S. E., & Feki, A. (2014). Modelling and rescuing neurodevelopmental defect of Down syndrome using induced pluripotent stem cells from monozygotic twins discordant for trisomy 21. *EMBO Molecular Medicine*, 6, 259–277.
- Hopper, J. L. (2000). Why ‘common environmental effects’ are so uncommon in the literature. In T. D. Spector, H. Snieder, & A. J. MacGregor (Eds.), *Advances in twin and sib-pair analysis* (pp. 151–165). Oxford University Press.
- Hopper, J. L., & Mathews, J. D. (1982). Extensions to multivariate normal models for pedigree analysis. *Annals of Human Genetics*, 46, 373–383.
- Hrubec, Z., & Robinette, C. D. (1984). The study of human twins in medical research. *New England Journal of Medicine*, 310, 435–441.
- Kirby, A., Gebiski, V., & Keech, A. C. (2002). Determining the sample size in a clinical trial. *Medical Journal of Australia*, 177, 256–257.
- Marsh, C. E., Thomas, H. J., Naylor, L. H., Scurrah, K. J., & Green, D. J. (2020a). Exploring human trainability: Design and rationale of Studies of Twin Responses to Understand Exercise as a Therapy (STRUETH) study. *Contemporary Clinical Trials Communications*, 19, 100584.
- Marsh, C. E., Thomas, H. J., Naylor, L. H., Scurrah, K. J., & Green, D. J. (2020b). Fitness and strength responses to distinct exercise modes in twins: Studies of Twin Responses to Understand Exercise as a Therapy (STRUETH) study. *Journal of Physiology*, 598, 3845–3858.
- Martin, N., Boomsma, D., & Machin, G. (1997). A twin-pronged attack on complex traits. *Nature Genetics*, 17, 387–392.
- Murphy, K., Lam, J., Cutler, T., Tyler, J., Calais-Ferreira, L., Li, S., Little, C., Ferreira, P., Craig, J. M., Scurrah, K. J., & Hopper, J. L. (2019). Twins Research Australia: A new paradigm for driving twin research. *Twin Research and Human Genetics*, 22, 438–445.
- Pinheiro, M. B., Ho, K. K., Ferreira, M. L., Refshauge, K. M., Grunstein, R., Hopper, J. L., Maher, C. G., Koes, B. W., Ordoñana, J. R., & Ferreira, P. H. (2016). Efficacy of a sleep quality intervention in people with low back pain: protocol for a feasibility randomized co-twin controlled trial. *Twin Research in Human Genetics*, 19, 492–501.
- Plomin, R., & Haworth, C. M. A. (2010). Genetics and intervention research [research article]. *Perspectives on Psychological Science*, 5, 557–563.
- Scurrah, K. J., & Hopper, J. L. (2019). *Twin research: Designs and analytic approaches*. Twins Research Australia. <https://www.twins.org.au/research/tools-and-resources/125-conversation-in-twin-research/377-twin-research-designs-and-analytic-approaches>.
- Sjölander, A., Frisell, T., & Öberg, S. (2012). Causal interpretation of between-within models for twin research. *Epidemiologic Methods*, 1, 217–237.
- Smith, M. I., Yatsunenko, T., Manary, M. J., Trehan, I., Mkakosya, R., Cheng, J., Kau, A. L., Rich, S. S., Concannon, P., Mychaleckyj, J. C., Liu, J., Houpt, E., Li, J. V., Holmes, E., Nicholson, J., Knights, D., Ursell, L. K., Knight, R., & Gordon, J. I. (2013). Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. *Science*, 339, 548–554.
- Sumathipala, A., Yelland, L., Green, D., Shepherd, T., Jayaweera, K., Ferreira, P., & Craig, J. M. (2018). Twins as participants in randomized controlled trials: A review of published literature. *Twin Research and Human Genetics*, 21, 51–56.
- Thomas, H. J., Marsh, C. E., Maslen, B. A., Scurrah, K. J., Naylor, L. H., & Green, D. J. (2021). Studies of Twin Responses to Understand Exercise Therapy (STRUETH): Body composition. *Medicine & Science in Sports Exercise*, 53, 58–67.
- Torgerson, D. J. (2001). Contamination in trials: is cluster randomisation the answer? *British Medical Journal*, 322, 355.
- Ullemar, V., Lundholm, C., Ortqvist, A. K., Gumpert, C. H., Anckarsater, H., Lundstrom, S., & Almqvist, C. (2015). Predictors of adolescents’ consent to use health records for research and results from data collection in a Swedish twin cohort. *Twin Research and Human Genetics*, 18, 256–265.
- Yelland, L. N., Salter, A. B., Ryan, P., & Makrides, M. (2011). Analysis of binary outcomes from randomised trials including multiple births: when should clustering be taken into account? *Paediatric and Perinatal Epidemiology*, 25, 283–297.
- Yelland, L. N., Sullivan, T. R., Collins, C. T., Price, D. J., McPhee, A. J., & Lee, K. J. (2018). Accounting for twin births in sample size calculations for randomised trials. *Paediatric and Perinatal Epidemiology*, 32, 380–387.
- Yelland, L. N., Sullivan, T. R., Price, D. J., & Lee, K. J. (2017). Sample size calculations for randomised trials including both independent and paired data. *Statistics in Medicine*, 36, 1227–1239.
- Zeger, S. L., & Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, 42, 121–130.
- Zumla, A., & Costello, A. (2002). Ethics of healthcare research in developing countries. *Journal of the Royal Society of Medicine*, 95, 275–276.