CAMBRIDGE
UNIVERSITY PRESS

Article

The Co-Twin Control Design: Implementation and Methodological Considerations

Bodine M. A. Gonggrijp[1,2] (iD), Steve G. A. van de Weijer[1], Catrien C. J. H. Bijleveld[1,3], Jenny van Dongen[2,4] (iD) and Dorret I. Boomsma[2,4]

[1]Netherlands Institute for the Study of Crime and Law Enforcement (NSCR), Amsterdam, the Netherlands, [2]Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands, [3]Department of Criminal Law and Criminology, VU University Amsterdam, Amsterdam, the Netherlands and [4]Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

Abstract

Establishing causal relationships in observational studies is an important step in research and policy decision making. The association between an exposure and an outcome can be confounded by multiple factors, often making it hard to draw causal conclusions. The co-twin control design (CTCD) is a powerful approach that allows for the investigation of causal effects while controlling for genetic and shared environmental confounding factors. This article introduces the CTCD and offers an overview of analysis methods for binary and continuous outcome and exposure variables. Tools for data simulation are provided, along with practical guidance and accompanying scripts for implementing the CTCD in R, SPSS, and Stata. While the CTCD offers valuable insights into causal inference, it depends on several assumptions that are important when interpreting CTCD results. By presenting a broad overview of the CTCD, this article aims to equip researchers with actionable recommendations and a comprehensive understanding of the design's strengths and limitations.

Keywords: Discordant twin design; causal inference; genetic confounding; twin research

(Received 12 July 2023; accepted 23 July 2023; First Published online 1 September 2023)

Establishing and quantifying the causal effect of an exposure on an outcome trait is a major goal in many fields of research. Causal inference plays a pivotal role inside and outside science. It advances our understanding of the complex relationships between exposures and outcomes, informs public health interventions, and shapes policy decisions. Exposures and outcomes can both refer to a wide array of variables, including external environmental impact, comprehensive lifetime exposure records, individual traits, and disorders or diseases. In observational studies, unraveling causal relationships is challenging. The association between an exposure and an outcome can be confounded by multiple factors, leading to spurious associations and potentially incorrect causal inference. Confounders, if identified a priori and are observable (measurable), can be included in the analysis of causal relations. However, it is hard to rule out unobserved, possibly latent, confounders. Genetic confounding, for instance, arises when genes influence the exposure and the outcome independent of a causal relationship (horizontal pleiotropy; Solovieff et al., 2013). If the genes involved in the confounding are unknown, which is often the case in studies of genetically complex traits, the confounding is latent. Failing to account for confounding variables can result in erroneous causal

inference (Sjölander & Greenland, 2013). This highlights the need for robust methods that can help account for confounding factors in observational studies.

Family-based designs, including the co-twin control design (CTCD) or discordant twin-pair design, were developed to investigate causal relations in the presence of latent confounding. The CTCD utilizes data of discordant monozygotic (MZ) twin pairs to examine the association between an exposure and an outcome variable (Gesell, 1942). One of the earliest experiments to implement a CTCD was conducted in 1927 when a pair of identical twins were taught to climb stairs at different time intervals. Twin A underwent a 6-week training period at 46 weeks old and managed to climb the stairs in 26 seconds. Twin B, on the other hand, climbed the stairs in 45 seconds without any training. Subsequently, Twin B received training for 2 weeks and improved her time to 10 seconds, despite having a much shorter training duration compared to her co-twin. By the time they reached 56 weeks and 3 years, their performance on the staircase was similar. This study's findings established an association between learning and maturity (Gesell & Thompson, 1943).

In observational data, the CTCD is applied to discordant MZ twin data when one twin has been exposed to a condition or variable of interest and the other has not. For example, one twin is a smoker and the other is not, or one twin was the victim of a crime and the other was not. Discordance can also be for disorders and diseases such as MS (multiple sclerosis), cancer or psychiatric conditions. Discordancy is not limited to binary exposures. MZ twins may differ with respect to continuous exposures such as birth

Corresponding author: Bodine Gonggrijp; Email: b.m.a.gonggrijp@vu.nl
Cite this article: Gonggrijp BMA, van de Weijer SGA, Bijleveld CCJH, van Dongen J, Boomsma DI. (2023) The Co-Twin Control Design: Implementation and Methodological Considerations. *Twin Research and Human Genetics* **26**: 249–256, https://doi.org/10.1017/thg.2023.35

Table 1. Overview of articles focusing on the CTCD and their additional value

Author	Year	Title	Focus
Christian & Kang	1972	Efficiency of human monozygotic twins in studies of blood lipids	Extensive comparison of the paired and unpaired design, including the efficiency of both the designs, and estimation of the experimental error for studies involving twins versus unrelated subjects.
Martin et al.	1982	Co-twin control studies: Vitamin C and the common cold	Provides an introduction of the CTCD and discussion of the advantages as well as an application of the CTCD to study the effect of vitamin C on cold symptoms.
Hu et al.	1998	Modelling ordinal responses from co-twin control studies	Ordinal data in a CTCD by investigating the applicability of the random-effects and GEE approaches.
Lichtenstein et al.	2002	The Swedish Twin Registry: A unique resource for clinical, epidemiological and genetic studies	Step-by-step description of the discordant twin design.
Goldberg & Fisher	2005	Co-twin control methods	The basic analytic methods that can be applied to CTCD with a (dichotomous) discordant environmental risk factor and a continuous or dichotomous outcome variable are given.
Madsen & Osler	2009	Commentary: Strengths and limitations of the discordant twin-pair design in social epidemiology. Where do we go from here?	Discussion of the strengths and limitations of the CTCD.
Vitaro et al.	2009	The discordant MZ-twin method: One step closer to the holy grail of causality	Presents two analytical strategies based on the discordant MZ twin method; the difference score strategy and mixed strategy.
McGue et al.	2010	Causal inference and observational research: The utility of twins	Methods of causal inference are discussed, with a focus on the discordant twin design.
van Dongen et al.	2012	The continuing value of twin studies in the omics era	Review that considers the continuing value of twin studies in the current era of molecular genetic studies.
Sjölander et al.	2012	Analysis of 1:1 matched cohort studies and twin studies, with binary exposures and binary outcomes	An overview of methods for matched cohort and twin studies with binary exposures and outcomes.
D'Onofrio	2013	Critical need for family-based, quasi-experimental designs in integrating genetic and social science research	Discusses major advantages, limitations, and assumptions of family-based quasi-experimental designs for examining environmental risks.
Røysamb & Tambs	2016	The beauty, logic and limitations of twin studies	Central theoretical foundations of the classic and extended twin designs are discussed, including the CTCD.
Sahu & Prasuna	2016	Twin studies: A unique epidemiological tool	A general overview of twin studies is given and the steps of the co-twin control analyses with recommended software packages.
Segal	2019	Co-twin control studies: Natural events, experimental interventions and rare happenings/twin research	Issues of the CTCD are discussed with reference to earlier studies.
McAdams et al.	2021	Twins and causal inference: Leveraging nature's experiment	Review discussing how monozygotic and dizygotic twin pairs can be used to strengthen causal inference.
van Dijk et al.	2022	Using twins to assess what might have been: The co-twin control design	Description of the CTCD, including statistical framework, value and limitation, and example code for SAS and R.

weight (e.g., Groen-Blokhuis et al., 2011) or body mass index (e.g., van Dongen et al., 2015). The major advantage of the CTCD is the inherent matching on genetic and shared environmental confounders. Specifically, MZ twins are genetically identical, or nearly identical, and have shared many environmental influences including intrauterine exposures. In addition, MZ twins are perfectly matched for age and sex. The inclusion of discordant dizygotic (DZ) twins who are likewise matched for age and sex (in the case of same-sex DZ twins), and shared environment also is informative, as are designs with full siblings. However, DZ twins and siblings are not perfectly matched for genetic factors, as they share on average 50% of their alleles, and so are not genetically identical. Thus, the CTCD applied to MZ twins corrects for genetic and shared environmental confounding, and confounders directly associated with age and sex. There are other family-based designs that can be applied, including a sibling design, but only the analysis of data from MZ twins, who are nearly genetically identical, controls for genetic confounding (D'Onofrio, 2013). It is worth

noting that in rare cases MZ twins inherit very similar, but not entirely identical, genotypes. In such cases, discordant MZ twins can be valuable in identifying specific genetic mutations (post zygotic) that may be causing the differences. For instance, a study conducted by Kondo et al. (2002) successfully identified the interferon regulatory factor 6 gene (IRF6) as the locus responsible for the development of Van der Woude syndrome (VWS) in a pair of MZ twins discordant for VWS.

The CTCD has generated substantial interest and has led to multiple methodological and review articles. Table 1 contains a brief description of these articles and their scope. The aim of the current article is to provide an overview of the CTCD with a focus on its application to data from large twin registries and summarize the methodological considerations. We discuss the analysis of binary and continuous exposures and outcome variables. We add to the existing literature summarized in Table 1 by presenting simulation analyses that illustrate various scenarios encountered in the CTCD, such as different sources of confounding factors

(e.g., no confounding, genetic confounding, shared environmental confounding). Additionally, we offer a collection of scripts that can be readily utilized with the popular statistical software packages SPSS, R, and Stata.

Co-Twin Control Analyses Design

The CTCD, as applied to MZ and DZ twin pairs, typically involves three steps, as outlined by Lichtenstein et al. (2002). The first step focuses on a total sample. We assume that the sample is representative of the general population, which is likely to be the case in population-based twin registries. The second step focuses on DZ twins, while the third step concentrates on MZ twins. We outline these steps in more detail below.

Step 1: Test for the Association at the Population Level Between Exposure and Outcome

In this first step, the exposed individuals are compared to all nonexposed individuals (i.e., controls) to examine the association between the exposure and the outcome. Thus, the relationship between exposure and outcome is assessed in all participants in the sample, without considering twin status. Cases can come from nontwins (if included in the sample), discordant twin pairs, or pairs that are concordant for exposure. When analyzing data from twin registries, the total sample includes data from related individuals; thus it is necessary to correct for familial clustering, since the assumption of independent observations is otherwise violated.

Subsequent steps focus exclusively on DZ and MZ twin pairs. These steps aim to investigate the presence of familial and genetic confounding in the association between the exposure and outcome variables.

Step 2: Matched Analysis in Same-Sex Dizygotic Twin Pairs

In the second step, a within-pair analysis is conducted in DZ exposure-discordant twin pairs. Often, only same-sex DZ twins are considered. However, opposite-sex twins can be included, and in cases where the research hypothesis specifically relates to sex as a potential cause of the association under investigation, considering opposite-sex twins is an optimal design. For example, Cui et al. (2005) conducted a study on birth defects in a sample of 4768 opposite-sex twins, aiming to examine potential sex differences. The findings revealed that among the opposite-sex twin pairs, males exhibited a 29% higher risk of birth defects compared to their twin sisters.

DZ twins and nontwin siblings share on average 50% of their segregating alleles, whereas MZ twins share 100% of their alleles. Both MZ and DZ twins share certain environmental factors, such as shared family experiences, and early intrauterine exposures, such as maternal smoking during pregnancy. A within-DZ twin pairs analysis controls for shared environmental factors that were not accounted for in the population-level analyses.

Step 3: Matched Co-Twin Analysis in the Monozygotic Twin Pairs

The analysis of step 2 is repeated for the exposure-discordant MZ twin pairs. Given the aim of controlling for confounders, the analyses of exposure-discordant MZ twins is the most powerful, given the matching for sex, age, genetic influences, and shared environmental influences. Note that in studies that target rare diseases or involve large-scale omics, measurements in biological samples, such as epigenetic profiles in blood, often only this third step is conducted. In such cases, researchers actively seek pairs of MZ twins where one twin is affected by the disease, trait or exposure of interest, while the other twin remains unaffected or only the most extreme discordant twin pairs are selected. For example, a study by Dempster et al. (2014) revealed distinct DNA methylation differences in 18 pairs of MZ twins discordant for adolescent depression. Another example is the study by van Dongen et al. (2015) where only twin pairs' BMI were selected ($N = 120$ pairs) who were extremely discordant for BMI. The result supported causal effects of obesity, as the heavier twin had a more unfavorable blood biomarker profile than their leaner co-twin.

Evaluating Results Across the Three Steps

As explained, the CTCD tests the relationship between exposure and outcome in individuals who differ in their exposures in the population sample (step 1), within DZ twin pairs who are discordant for the exposure (step 2), and within discordant MZ twin pairs (step 3). To support a causal hypothesis, we inspect the strength of these relationships. The associations within the population (step 1, see above) may be confounded by genetic or environmental factors. Associations within DZ twin pairs discordant for exposure (step 2) control for shared environmental effects, and associations within MZ twin pairs discordant for exposure (step 3) control for both shared environmental and genetic effects.

Figure 1 illustrates four scenarios resulting from these three steps. In scenario 1A, where exposure causally affects the outcome, we expect to observe associations at the population level and within discordant twin pairs. For example, if victimization of a crime causes depression, the victims and nonvictims in the population will differ in depression status, and a similar difference will be observed within discordant twin pairs. However, this pattern is necessary, but not sufficient, evidence of causality since nonshared environmental factors are not accounted for by the CTCD design.

Scenario 1B represents the case where the association between exposure and outcome is entirely explained by genetic confounding. While an association is expected to be observed at the population sample level, it is expected to be zero within the MZ pairs, as the genetic factors that confound the association are identical between the twins. The association within discordant DZ twin pairs will be intermediate, with the exact effect size depending on the size of the genetic covariance between the exposure and outcome variable and the variance of the exposure variable.

Scenario 1C depicts shared environmental factors as the sole confounder. Associations between exposure and outcome are again expected to be observed at the population level, but absent within both DZ and MZ twin pairs, as the shared environment is equal in both types of pairs. Finally, scenario 1D suggests partial confounding by genetic and environmental factors. Associations would be reduced within DZ twin pairs and further reduced within MZ pairs compared to the population-level effect. The presence of an exposure effect within discordant MZ pairs supports at least a partial causal effect.

Statistical Approaches

Multiple statistical approaches have been developed to estimate the association between exposure and outcome in the three steps outlined above. An overview of commonly applied analyses is presented in Table 2. We note that Table 2 does not provide an

Table 2. Possible statistical models for the population and discordant twin analyses for both binary and continuous variables

			Outcome	
			Binary	Continuous
Population analysis	Exposure	Binary	Logistic regression	Linear regression Two-sample *t* test ANOVA
		Continuous	Logistic regression	Linear regression
Discordant twin analyses	Exposure	Binary	McNemar's test Conditional logistic regression	Fixed effects regression Paired *t* test
		Continuous	Conditional logistic regression	Fixed effects regression

Note: ANOVA, analysis of variance.
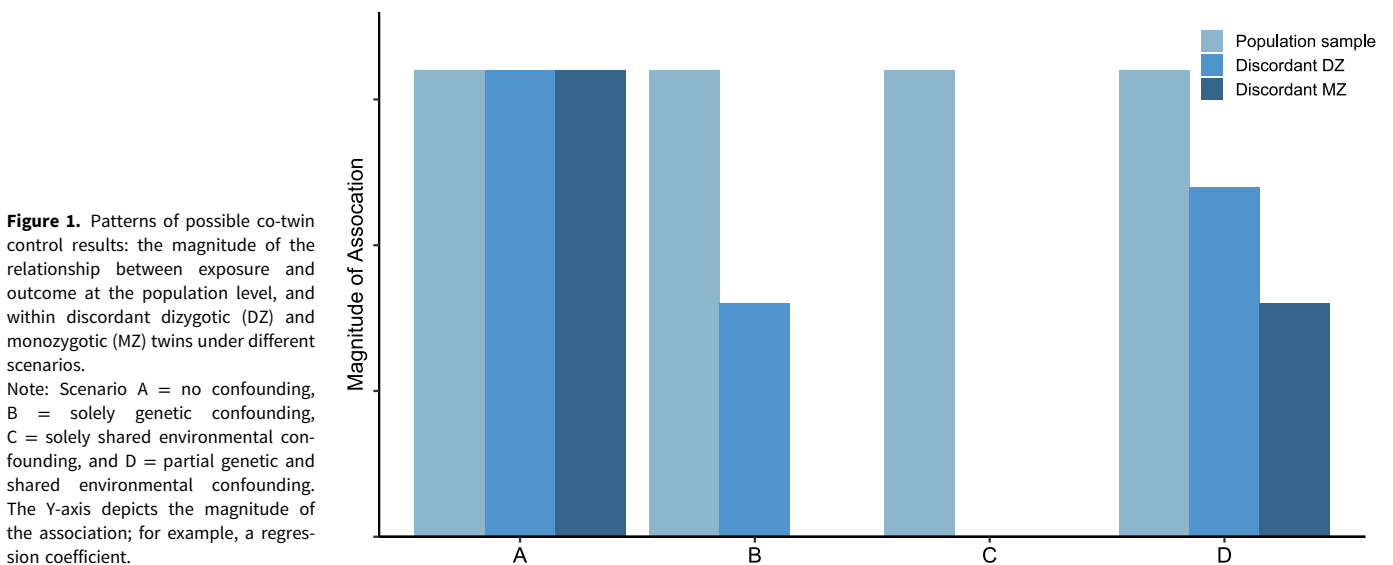
Figure 1. Patterns of possible co-twin control results: the magnitude of the relationship between exposure and outcome at the population level, and within discordant dizygotic (DZ) and monozygotic (MZ) twins under different scenarios.
Note: Scenario A = no confounding, B = solely genetic confounding, C = solely shared environmental confounding, and D = partial genetic and shared environmental confounding. The Y-axis depicts the magnitude of the association; for example, a regression coefficient.

exhaustive overview of all possible analyses but aims to summarize some common approaches.

One point to emphasize is that in twin data analysis, the inclusion of twin pairs in the analysis relies on their discordance in both exposure level and outcome variables. If the outcome or exposure variable is the same within a twin pair, it indicates a lack of within-pair variation. Consequently, such twin pairs do not contribute information. Thus, only twin pairs discordant for the exposure and outcome variable contribute to the estimation of the association between the exposure and outcome (Frisell et al., 2012). While this phenomenon of exclusion is often discussed in relation to binary variables, where many twin pairs may share the same exposure or outcome values, it is not limited to binary data. In the case of continuous variables, twins can be concordant, that is, share the same values (e.g., have the same height). Thus, in both continuous and binary data, twin concordance on the exposure or outcome results in the exclusion of twin pairs, and the true sample size in these analyses is given by the number of discordant twin pairs for the exposure as well as the outcome.

Outcome Variable: Binary

Analyses of binary outcomes often focus on odds ratio (OR) estimates. An OR of 1 indicates no association, while ORs below

1 or above 1 indicate decreased or increased odds of the outcome with increasing exposure, respectively (Moser & Coombs, 2004).

In population sample analyses, logistic regression models adjusted for covariates are often employed to analyze binary outcomes. If there is clustering in the data because of related participants, a correction for clustering is necessary; for example, by mixed models with random effects or a generalized estimating equation approach (GEE). In discordant twin analyses with continuous exposure and binary outcome variables, conditional logistic regression analyses (CLR) are common, accounting for stratification and matching (Graubard & Korn, 2011; Sjölander et al., 2012). For example, Kendler and Gardner (2010) analyzed data from 4910 female and male twins by CLR to investigate the association between dependent stressful life events and major depression (MD). They found a strong association between dependent stressful life events and MD in both female and male twins. The association was lower in DZ twins and lower still in MZ twins, suggesting confounding by genetic factors.

When the exposure and outcome variables are binary, an alternative approach for discordant twin analyses is provided by McNemar's test. This test is based on the 2 x 2 contingency table of matched responses to determine whether there are statistically significant differences in a dichotomous dependent variable between two related groups (Adedokun & Burgess, 2012). For instance, Romanov et al. (2003) employed McNemar's test in a

study of 9947 Finnish adult twins and found that the effect of multiple life events on depression was similar in both MZ and DZ twins discordant for depression, indicating that the relationship between experiencing multiple life events and depression is not due to genetic and shared environmental confounding.

Outcome Variable: Continuous

For continuous outcome variables, there are several modeling options for population analyses. One common approach is to use linear regression models, which may include measured covariates. Correction for clustering may be necessary using mixed models with random effects for clustering factors such as family and zygosity of twins, or a generalized estimating equation (GEE) approach.

In discordant twin analyses, a fixed-effect regression analysis can be conducted. This involves regressing the within-pair difference on the continuous outcome variable on the within-pair difference on the exposure variable. An example of this approach is demonstrated in the study by Middeldorp et al. (2008), which examined the associations of life events with anxious depression and personality in data from 5782 MZ and DZ twins participating in longitudinal survey studies in the Netherlands. The findings did not provide evidence to support a causal hypothesis, as there were no differences in anxious depression, neuroticism, and extraversion scores between exposed and nonexposed unrelated subjects and within discordant MZ and DZ twin pairs.

When both the exposure and outcome variables are continuous, often studies create two groups based on the continuous exposure variables to determine discordance between twins. For example, Groen-Blokhuis et al. (2011) investigated the association between low birth weight and attention problems. They classified twin pairs as discordant for birth weight if the birth weight of the smaller twin was at least 15% lower than the birth weight of the larger twin or if there was a birth weight difference of at least 400 grams. Children with lower birth weights had higher scores on hyperactivity and attention problems compared to children with higher birth weights. Similar findings were observed for unrelated pairs, as well as MZ and DZ twin pairs, providing evidence for a causal relationship between birth weight and attention problems. However, we note that this method of classification results in a loss of information. An alternative approach is to use the original scores of the exposure variable in the analyses, without dichotomizing it.

Statistical Approaches Simulation

A data simulation was conducted using R to demonstrate the application of the CTCD for assessing causality and confounding. Three datasets were generated: one for a population sample, one for a dizygotic (DZ) twin sample, and one for a monozygotic (MZ) twin sample, where the desired samples can be specified by the user. Note that when either the exposure variable or the outcome variable is binary, only the discordant twin pairs with respect to both the exposure and outcome will be included in the analyses, resulting in a reduced sample size for MZ and DZ pairs. The script for the data simulation can be found in Supplementary File 1. The simulation implemented the scenarios depicted in Figure 1 for each group (population, DZ twins, and MZ twins). Normally distributed continuous variables were generated for both the exposure (x) and the outcome (y) variables by exact data simulation. Subsequently, binary variables (dx and dy) were derived from the continuous variables. All continuous values above zero were assigned a value of

1, indicating a positive outcome ('case'), while the remaining values were assigned a value of zero ('control'). The simulation scenarios can be customized to incorporate different strengths of association and genetic and nongenetic variance components of x and y based on specific simulation requirements.

In the current simulation, the variance components for x and y were chosen based on an ACE model — that is, in this model the variation in a phenotype is due to additive genetic effects (A), the common environment (C), and the unique, random, environment (E) — with a relatively small contribution of C. The simulated exposure variable was based on traits such as victimization, with variance components of $a^2 = .45$, $c^2 = 0.22$, and $e^2 = 0.33$ (e.g., Beaver et al., 2009), while the simulated outcome variable represented a trait such as depression, with variance components of $a^2 = 0.40$, $c^2 = 0.10$, and $e^2 = 0.50$ (e.g., Huider et al., 2021; Sullivan et al., 2000).

Twin correlations and cross-twin, cross-trait correlations, either Pearson or tetrachoric correlations, are summarized in Supplementary Table 1. Additionally, the correlation between the difference scores of MZ and DZ twins for two traits is presented in Supplementary Table 1.

Analyses of Simulated Data in R, STATA, and SPSS

Statistical analyses were performed in R, STATA, and SPSS to analyze the simulated data in the CTCD. All analysis scripts can be found in the Github repository (https://github.com/bmagonggrijp/CTCD-Implementation-and-Methodological-Considerations) and in Supplementary File 1. The results from the analyses conducted in R, STATA, and SPSS are consistent, as shown in Supplementary Tables 2, 3, and 4 respectively. We again note that the estimation of associations in the twin analyses relies on discordant twin pairs contributing to the within-pair association's estimation. Consequently, not all twin pairs are included when, for example, a conditional logistic regression is performed, and thus the exact number of contributing twins should be carefully checked and reported in publications. Among the software packages we worked with, STATA and R provided this information in the results. In SPSS, conducting a conditional logistic regression requires specifying that only discordant twins should be selected for analysis to obtain the correct within-pair estimate for fixed effects and conditional logistic regression. The analyses of the simulated data demonstrate that if these steps are performed correctly, the three software packages produce exactly the same regression coefficients, with slightly different confidence intervals in SPSS compared to those in R and STATA.

Figure 2 presents the results of the different simulation analyses for each scenario depicted in Figure 1. The analyses vary depending on the type of variables involved (continuous or binary) and reveal differences in the magnitude of the association within the same scenario. In the top part of Figure 2, scenario A is depicted, where no confounding is present. The simulation analyses of continuous exposure and outcome variables show an expected pattern with consistent association magnitudes across the population sample and both DZ and MZ discordant twin analyses. However, when examining binary variables, there is a slight deviation from the expected pattern, due to the introduction of variability when deriving binary variables from continuous variables during the simulation process.

The second part of Figure 2 illustrates scenario B, where only genetic confounding is present. In each analysis, for both continuous and binary variables, an association is observed in
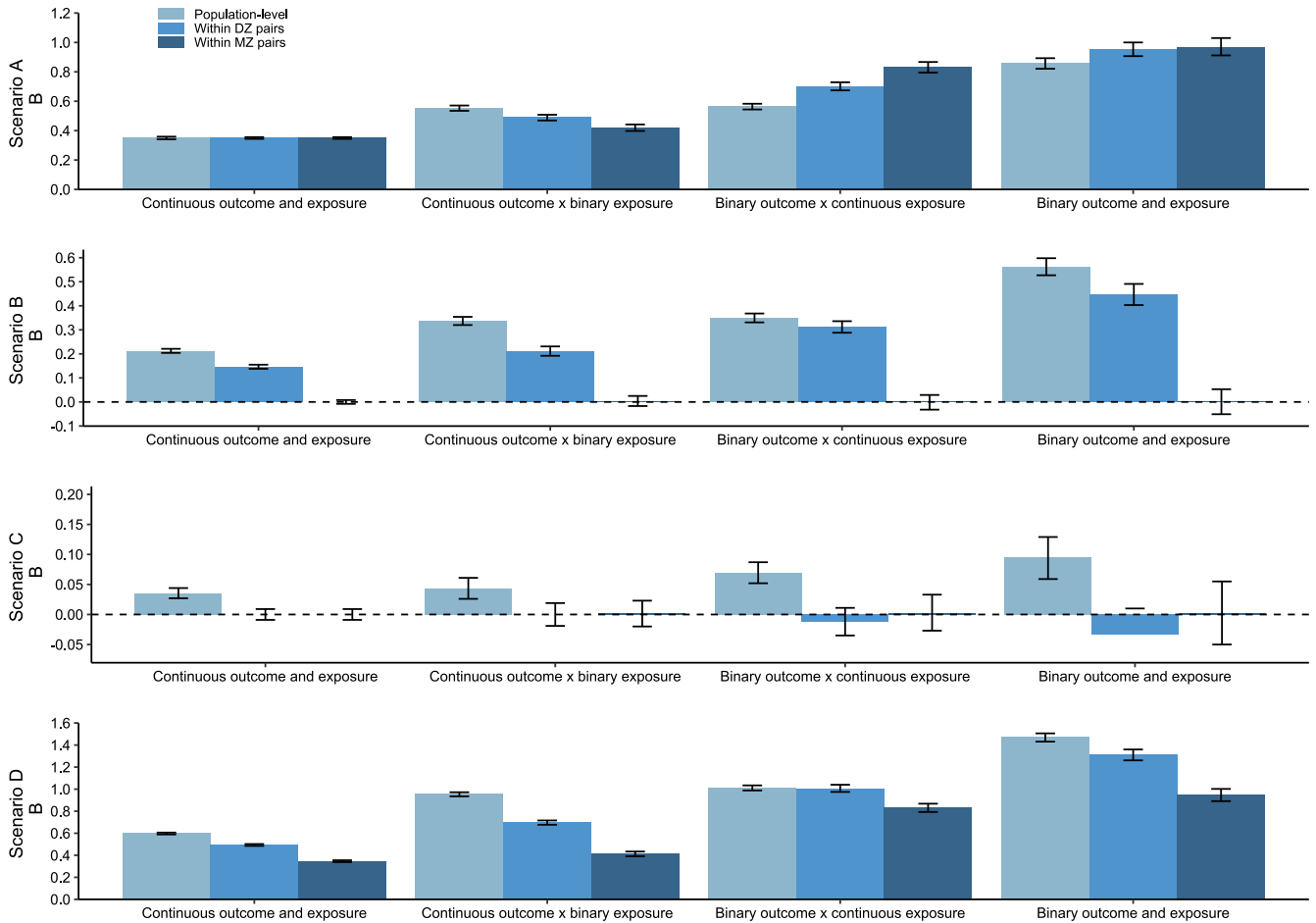
Figure 2. Results from simulation analyses, depicting the relationship between exposure and outcome at the population level, and within discordant dizygotic (DZ) and monozygotic (MZ) twins under different scenarios and types of variables.

Note: B, regression coefficient. Scenarios A = no confounding, B = solely genetic confounding, C = solely shared environmental confounding) and D = Causal effect with genetic and shared environmental confounding.

the population sample, while there is no association in the MZ discordant twins, and the DZ twins show an intermediate pattern. The third part of Figure 2 shows the simulation of scenario C, which involves confounding by shared environmental factors. As expected, only the population sample shows an association, while the association within the twins is either zero or not significantly different from zero. The association in the population sample is relatively small compared to the other scenarios because the simulated exposure and outcome variables both had a relatively low contribution of shared environmental factors (0.22 and 0.10, respectively).

Finally, the fourth part of Figure 2 depicts a scenario where there is partly genetic and shared environmental confounding. Here again, the expected patterns emerge, with reduced association magnitudes within DZ and MZ pairs compared to the effect within the population sample. Thus, under the assumption of no unshared environmental confounding, we draw correct conclusions with respect to the presence or absence of genetic or shared environmental confounding.

Discussion

The co-twin control design (CTCD) is a valuable tool for investigating causal associations while controlling for confounding

by shared environmental and genetic factors. In this article, we reviewed common analysis methods for CTCD with binary or continuous outcome and exposure variables and provided practical guidance and scripts for implementing the design in R, SPSS, and Stata, and illustrating it using simulated data. The CTCD's greatest strength lies in its ability to control for unmeasured genetic and shared environmental factors, providing evidence for a causal association. However, the CTCD cannot completely rule out alternative explanations for observed associations.

One assumption in the CTCD is the absence of confounding by nonshared environmental factors. While MZ twins share genetic and some environmental influences, they also have nonshared experiences and exposures that make them unique. Consequently, when the association of exposure with the outcome within the population sample is similar to the within-DZ and within-MZ pairs association this may reflect true causality, but it can also reflect the effect of the nonshared experiences that led to differences in exposure.

Second, the CTCD cannot distinguish between causation and reverse causation (McGue et al., 2010). Causation refers to the traditional understanding that variable X causes variable Y, while reverse causation means that Y causes X. The CTCD is designed to test for a possible causal association based on prior theory about the cause and the consequence, which could be the case, for

example, when a causal variable was measured before the outcome variable. Analyzing data from longitudinal twin studies can help rule out possible reverse causation. Within twin research, several approaches have been developed to explore the direction of causation, such as the direction of causation model (DCM; Gillespie & Martin, 2005; Heath et al., 1993). The DCM is designed to address the direction of causality when two variables have different genetic architectures, such as an AE versus a CE model.

The CTCD, and other causal designs, may fail to detect a within-pair association due to measurement error. According to classical test score theory, the presence of measurement error in the exposure variable has a more pronounced impact on within-pair associations compared to individual-level associations due to the compounded effect of error inherent in calculating difference scores (Ashenfelter & Krueger, 1994). To illustrate, consider the interplay between σ, an estimate of the proportion of measurement error in the exposure variable, and ρ, the twin correlation. At the individual level, the regression coefficient of the outcome on the exposure is attenuated by σ. However, at the within-pair level, the attenuation is influenced by $\sigma/(1-\rho)$ (Frissell et al., 2012; McGue et al., 2010). Consequently, when measurement error accounts for, for example, 5% of the exposure variable, the individual-level estimate would be attenuated by 5%. However, the within-pair estimate would be attenuated by 25% when the twin correlation on exposure is .8 and by 12.5% when the twin correlation is .6. Measurement error thus leads to a greater attenuation in the DZ and MZ analyses compared to the association at the population level. Moreover, as MZ twin correlations are typically higher than DZ twin correlations, it is expected that the within-pair attenuation will be greater for MZ than for DZ pairs. Frissell et al. (2012) conducted a series of simulations in paired sibling data under a logistic model with binary exposure and outcome, where the association of x and y was causal and not confounded. Their simulations demonstrated measurement error to lead to an attenuation of the unpaired *OR* estimates and to an ever greater attenuation of the within-sibling *OR*. That is, measurement error in the exposure variable can result in the underestimation of causal effects or the failure to observe significant within-pair associations, even when a causal association is present between the exposure and outcome (Duffy & Martin, 1994; Frissell et al., 2012; McGue et al., 2010).

The CTCD may also fail to detect a within-pair association due to a 'spill-over' effect. This effect occurs when the exposure being investigated not only affects the exposed twin but also influences the nonexposed twin within the pair. The likelihood of observing a spill-over effect varies depending on the nature of the exposure. For example, research has shown that certain exposures, such as victimization of sexual abuse within a family, can impact not only the victimized individual but also nonvictimized family members (de Jong, 2022). In contrast, exposures like birth weight are highly unlikely to exert an influence on the co-twin.

In this article we provided an overview of the CTCD with a focus on its application to data from large twin registries. However, this is not always feasible when, for example, studying rare outcomes or exposures. In such cases, researchers may face challenges in obtaining a sufficient sample size of discordant MZ twin pairs. For such rare diseases and disorders, the ascertainment into a research project may be through other routes, such as patient registries.

In conclusion, the CTCD can be a powerful design for investigating causal associations while controlling for genetic and shared environmental confounding factors. This article aids future researchers to employ the CTCD by presenting an overview and discussion of its strengths and limitations. By providing a set of scripts that can be readily used with the popular statistical software packages SPSS, R, and STATA, and by providing tools to aid in data simulation, we aimed to contribute to an understanding and application of the design.

References

Adedokun, O. A., & Burgess, W. D. (2012). Analysis of paired dichotomous data: A gentle introduction to the McNemar test in SPSS. *Journal of MultiDisciplinary Evaluation*, 8, 125–131. https://doi.org/10.56645/jmde.v8i17.336

Ashenfelter, O., & Krueger, A. (1994). Estimates of economic return to schooling from a new sample of twins. *American Economic Review*, 84, 1157–1117.

Beaver, K. M., Boutwell, B. B., Barnes, J. C., & Cooper, J. A. (2009). The biosocial underpinnings to adolescent victimization: Results from a longitudinal sample of twins. *Youth Violence and Juvenile Justice*, 7, 223–238. https://doi.10.1177/1541204009333830

Christian, J. C., & Kang, K. W. (1972). Efficiency of human monozygotic twins in studies of blood lipids. *Metabolism*, 21, 691–699. https://doi.10.1016/0026-0495(72)90118-7

Cui, W., Ma, C. X., Tang, Y., Chang, V., Rao, P. V., Ariet, M., Resnick, M. B., & Roth, J. (2005). Sex differences in birth defects: A study of opposite-sex twins. *Birth Defects Research Part A: Clinical and Molecular Teratology*, 73, 876–880. https://doi.10.1002/bdra.20196

de Jong, R. (2022). *Long-term consequences of child sexual abuse: Adult role fulfillment of child sexual abuse victims* [Unpublished doctoral thesis]. Vrije Universiteit Amsterdam. https://research.vu.nl/en/publications/long-term-consequences-of-child-sexual-abuse-adult-role-fulfillme

Dempster, E. L., Wong, C. C., Lester, K. J., Burrage, J., Gregory, A. M., Mill, J., & Eley, T. C. (2014). Genome-wide methylomic analysis of monozygotic twins discordant for adolescent depression. *Biological Psychiatry*, 76, 977–983. https://doi.10.1016/j.biopsych.2014.04.013

D'Onofrio, B. M., Lahey, B. B., Turkheimer, E., & Lichtenstein, P. (2013). Critical need for family-based, quasi-experimental designs in integrating genetic and social science research. *American Journal of Public Health*, 103, S46–S55. https://doi.10.2105/AJPH.2013.301252

Duffy, D., & Martin, G. (1994). Inferring the direction of causation in cross-sectional twin data: theoretical and empirical considerations. *Genetic Epidemiology*, 11, 483–502. https://doi.10.1002/gepi.1370110606

Frissell, T., Öberg, S., Kuja-Halkola, R., & Sjölander, A. (2012). Sibling comparison designs: bias from non-shared confounders and measurement error. *Epidemiology*, 23, 713–720. https://doi.10.1097/EDE.0b013e31825fa230

Gesell, A. (1942). The method of co-twin control. *Science*, 95, 446–448. https://doi.10.1126/science.95.2470.446

Gesell, A., & Thompson, H. (1943). Learning and maturation in identical infant twins. In R. G. Barker, J. S. Kounin, & H. F. Wright (Eds.), *Child behavior and development: A course of representative studies* (pp. 209–227). McGraw-Hill. https://doi.org/10.1037/10786-013

Gillespie, N. A., & Martin, N. G. (2005). Direction of causation models. *Encyclopedia of Statistics in Behavioral Science*, 1, 496–499.

Goldberg, J., & Fischer, M. (2005). Co-twin control methods. *Encyclopedia of Statistics in Behavioral Science*. https://doi.org/10.1002/0470013192.bsa143

Graubard, B. I., & Korn, E. L. (2011). Conditional logistic regression with survey data. *Statistics in Biopharmaceutical Research*, 3, 398–408. https://doi.10.1198/sbr.2010.10002

Groen-Blokhuis, M. M., Middeldorp, C. M., van Beijsterveldt, C. E., & Boomsma, D. I. (2011). Evidence for a causal association of low birth weight and attention problems. *Journal of the American Academy of Child & Adolescent Psychiatry*, *50,* 1247–1254. https://doi.10.1016/j.jaac.2011.09.007

Heath, A. C., Kessler, R. C., Neale, M. C., Hewitt, J. K., Eaves, L. J., & Kendler, K. S. (1993). Testing hypotheses about direction of causation using cross-sectional family data. *Behavior Genetics*, *23*, 29–50. https://doi.10.1007/BF01067552

Hu, F. B., Goldberg, J., Hedeker, D., & Henderson, W. G. (1998). Modelling ordinal responses from co-twin control studies. *Statistics in Medicine*, *17*, 957–970. https://doi:10.1002/(sici)1097-0258(19980515)17:9<957::aid-sim791>3.0.co;2-o

Huider, F., Milaneschi, Y., van der Zee, M. D., de Geus, E. J. C., Helmer, Q., Penninx, B. W. J. H., & Boomsma, D. I. (2021). Major depressive disorder and lifestyle: Correlated genetic effects in extended twin pedigrees. *Genes*, *12*, 1509. https://doi.10.3390/genes12101509

Kendler, K. S., & Gardner, C. O. (2010). Dependent stressful life events and prior depressive episodes in the prediction of major depression: The problem of causal inference in psychiatric epidemiology. *Archives of General Psychiatry*, *67*, 1120–1127. https://doi.10.1001/archgenpsychiatry.2010.136

Kondo, S., Schutte, B. C., Richardson, R. J., Bjork, B. C., Knight, A. S., Watanabe, Y., Howard, E., de Lima, R. L., Daack-Hirsch, S., Sander, A., McDonald-McGinn, D. M., Zackai, E. H., Lammer, E. J., Aylsworth, A. S., Ardinger, H. H., Lidral, A. C., Pober, B. R., Moreno, L., Arcos-Burgos, M., . . . Murray, J. C. (2002). Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. *Nature Genetics*, *32,* 285–289. https://doi.10.1038/ng985

Lichtenstein, P., de Faire, U., Floderus, B., Svartengren, M., Svedberg, P., & Pedersen, N. L. (2002). The Swedish Twin Registry: A unique resource for clinical, epidemiological and genetic studies. *Journal of Internal Medicine*, *252*, 184–205. https://doi.10.1046/j.1365-2796.2002.01032.x

Madsen, M., & Osler, M. (2009). Commentary: strengths and limitations of the discordant twin-pair design in social epidemiology. Where do we go from here? *International Journal of Epidemiology*, *38,* 1322–1323. https://doi.10.1093/ije/dyp264

Martin, N. G., Carr, A. B., Oakeshott, J. G., & Clark, P. (1982). Co-twin control studies: Vitamin C and the common cold. *Progress in Clinical and Biological Research*, *103*, 365–373. https://doi.10.1017/s0001566000006450

McAdams, T. A., Rijsdijk, F. V., Zavos, H. M., & Pingault, J. B. (2021). Twins and causal inference: Leveraging nature's experiment. *Cold Spring Harbor Perspectives in Medicine*, *11*, a039552. https://doi.10.1101/cshperspect.a039552

McGue, M., Osler, M., & Christensen, K. (2010). Causal inference and observational research: The utility of twins. *Perspectives on Psychological Science*, *5*, 546–556. https://doi.10.1177/1745691610383511

Middeldorp, C. M., Cath, D. C., Beem, A. L., Willemsen, G., & Boomsma, D. I. (2008). Life events, anxious depression and personality: A prospective and genetic study. *Psychological Medicine*, *38,* 1557–1565. https://doi.10.1017/S0033291708002985

Moser, B. K., & Coombs, L. P. (2004). Odds ratios for a continuous outcome variable without dichotomizing. *Statistics in Medicine*, *23*, 1843–1860. https://doi.10.1002/sim.1776

Romanov, K., Varjonen, J., Kaprio, J., & Koskenvuo, M. (2003). Life events and depressiveness ¾ The effect of adjustment for psychosocial factors, somatic health and genetic liability. *Acta Psychiatrica Scandinavica*, *107,* 25–33. https://doi.10.1034/j.1600-0447.2003.01419.x

Røysamb, E., & Tambs, K. (2016). The beauty, logic and limitations of twin studies. *Norsk Epidemiologi*, *26,* 35–46. https://doi.org/10.5324/nje.v26i1-2.2014

Sahu, M., & Prasuna, J. G. (2016). Twin studies: A unique epidemiological tool. *Indian journal of Community Medicine*, *41*, 177–182. https://doi.10.4103/0970-0218.183593

Segal, N. L. (2019). Co-twin control studies: Natural events, experimental interventions and rare happenings/Twin research: Cancer risk in overweight twins; Prognosis after fetal loss of one twin; Twin concordance for Parkinson's disease; Neuroanatomy of musically discordant MZ twins/News articles: Twin birth with two wombs; Twins' prenatal interactions; Switched-at-birth twins; Fetus-in-fetu; Unsolved paternity. *Twin Research and Human Genetics*, *22*, 272–276. https://doi.10.1017/thg.2023.21

Sjölander, A., & Greenland, S. (2013). Ignoring the matching variables in cohort studies ¾ When is it valid and why? *Statistics in Medicine*, *32*, 4696–4708. doi: 10.1002/sim.5879

Sjölander, A., Johansson, A., Lundholm, C., Altman, D., Almqvist, C., & Pawitan, Y. (2012). Analysis of 1:1 matched cohort studies and twin studies, with binary exposures and binary outcomes. *Statistical Science*, *27*, 395–411. https://www.jstor.org/stable/41714772

Solovieff, N., Cotsapas, C., Lee, P., Purcell, S. M., & Smoller, J. W. (2013). Pleiotropy in complex traits: Challenges and strategies. *Nature Reviews Genetics*, *14*, 483–495. https://doi.10.1038/nrg3461

Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression: Review and meta-analysis. *American Journal of Psychiatry*, *157*, 1552–1562. https://doi.10.1176/appi.ajp.157.10.1552

van Dijk, W., Norris, C. U., & Hart, S. A. (2022). Using twins to assess what might have been: The co-twin control design. *Research on Social Work Practice*. Advance online publication. https://doi.org/10.1177/10497315221120604

van Dongen, J., Slagboom, P. E., Draisma, H. H., Martin, N. G., & Boomsma, D. I. (2012). The continuing value of twin studies in the omics era. *Nature Reviews Genetics*, *13*, 640–653. https://doi.org/10.1038/nrg3243

van Dongen, J., Willemsen, G., Heijmans, B. T., Neuteboom, J., Kluft, C., Jansen, R., Penninx, B. W., Slagboom, P. E., de Geus, E. J., & Boomsma, D. I. (2015). Longitudinal weight differences, gene expression and blood biomarkers in BMI-discordant identical twins. *International Journal of Obesity*, *39*, 899–909. https://doi.10.1038/ijo.2015.24

Vitaro, F., Brendgen, M., & Arseneault, L. (2009). The discordant MZ-twin method: One step closer to the holy grail of causality. *International Journal of Behavioral Development*, *33,* 376–382. https://doi.org/10.1177/0165025409340805