**Appendix A**

In our paper, we adopted the Least Squares (LS) paradigm in Visscher (2004) to derive the algebraic association among statistical power to detect A, the genetic relatedness of same-sex twins (HSS) and the number of pairs of twins (n).

If we characterize the 1st kin and the 2nd kin in a kin pair as two random blocks of random effects, n pairs of kin can be treated as a one-way random effects design. Based on the ANOVA table of one-way random effects design proposed by Shrout and Fleiss (1979), we can easily derive the between-pair (B) and within-pair (W) observed mean squares (MS) table for either n pairs of SS twins or OS twins

|  |  |  |  |
| --- | --- | --- | --- |
|  | df | MS | E(MS) |
| Between pairs | n-1 | B | $$2σ\_{b}^{2}+σ\_{w}^{2}$$ |
| Within pair | n | W | $$σ\_{w}^{2}$$ |

When decomposing the variance into heritability (*h2*) and common environmental impact (*c2*) based on the structure of the ACE model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | $$σ\_{b}^{2}$$ | $$σ\_{w}^{2}$$ | E(B) | E(W) |
| SS pairs | $$H\_{SS}a^{2}+c^{2}$$ | $$1-H\_{SS}a^{2}-c^{2}$$ | $2\left(H\_{SS}a^{2}+c^{2}\right)+\left(1-H\_{SS}a^{2}-c^{2}\right)$  | $$\left(1-H\_{SS}a^{2}-c^{2}\right)$$ |
| OS pairs | $$\frac{1}{2}a^{2}+c^{2}$$ | $$1-\frac{1}{2}a^{2}-c^{2}$$ | $2\left(\frac{1}{2}a^{2}+c^{2}\right)+\left(1-\frac{1}{2}a^{2}-c^{2}\right)$  | $$\left(1-\frac{1}{2}a^{2}-c^{2}\right)$$ |

Where *HSS* is the expected genetic relatedness of SS twins.

The intraclass correlation is calculated based on $\hat{t}=\frac{σ\_{b}^{2}}{σ\_{b}^{2}+σ\_{w}^{2}}$. Combined with the ANOVA table above, for both SS twins and OS twins we can derive $\hat{t}=\frac{B-W}{B+W}$. Visscher [(1998)](https://www.zotero.org/google-docs/?fZ22lN)applied a first-order Taylor-series expansion on the expected mean squared to derive the equations to calculate variance of the intraclass correlations

$$var\left(\hat{t}\right)≈\left(1-t\right)^{2}\left(1+t\right)^{2}=\frac{\left(1-t^{2}\right)^{2}}{n}$$

In our circumstance using SS twins and OS twins, the heritability is estimated via

$$\hat{a}^{2}=\frac{1}{H\_{SS}-.5}\left(\hat{t}\_{SS}-\hat{t}\_{OS}\right)$$

And the variance of the heritability estimate is

$$var\left(\hat{a}^{2}\right)=\left(\frac{1}{H\_{SS}-.5}\right)^{2}\left[var\left(\hat{t}\_{SS}\right)+var\left(\hat{t}\_{OS}\right)\right]=\left(\frac{1}{H\_{SS}-.5}\right)^{2}\left[\frac{\left(1-t\_{SS}^{2}\right)^{2}+\left(1-t\_{OS}^{2}\right)^{2}}{n}\right]$$

In terms of heritability and common environment impact, the equation can also be expressed as

$$var\left(\hat{a}^{2}\right)=\frac{\left(1-\left(H\_{SS}a^{2}+c^{2}\right)^{2}\right)^{2} +\left(1-\left(0.5a^{2}+c^{2}\right)^{2}\right)^{2}}{n\left(H\_{SS}-.5\right)^{2}}$$

When sample sizes are large, the test statistic of a normal test $λ=\frac{a^{2}}{SE\left(\hat{a}^{2}\right)}$ is approximately equal to the noncentrality parameter (NCP) of a chi-square test statistic (Visscher, 2004). Therefore, we assume the alternative distribution of heritability is $T∼N\left(λ,1\right)$, a type-I error rate of α and a type-II error rate of β. In a one-tail test,

$$Power=1-β=Prob\left(x>Z\_{1-α}-λ\right)$$

where x is a random variable following a standard *N*(0,1). We can transform the equation into a format of $λ=Z\_{1-β}+Z\_{1-α}$, where λ can be expressed in terms of HSS and n with $λ^{2}=\frac{a^{4}}{var\left(\hat{a}^{2}\right)}$. Hence, we can illustrate the association among statistical power to detect A, the genetic relatedness of same-sex twins (HSS) and the number of pairs of twins (n)

$$\left(Z\_{1-α}+Z\_{1-β}\right)^{2}=n\frac{\left(a^{2}\right)^{2}\left(H\_{SS}-0.5\right)^{2}}{\left(1-\left(H\_{SS}a^{2}+c^{2}\right)^{2}\right)^{2} +\left(1-\left(0.5 a^{2}+c^{2}\right)^{2}\right)^{2} }$$

which is the equation 3 in our paper.

**Appendix B**

Additional simulation results for variance structure A = 1.0; C = 1.0; E = 1.0 and A = .5; C = 2.0; E = .5 are provided in the supplementary figures. To obtain a rough estimate of power, researchers could match the variance structure, the sample size, and the relatedness difference with the setting of their empirical studies. All figures in the results section and here should not be treated as more than a preliminary reference for any study decisions like model selection or data processing. A more accurate power can be estimated by completing their own customized simulations. More information for customized simulations in R can be found in the vignette of the ACEsimFit package (<https://cran.r-project.org/web/packages/ACEsimFit/vignettes/ACEsimFit.html>).



Supplementary Figure S1-1. Power of the ACE model to detect A under the simulated variance of A = 1.0; C = 1.0; E = 1.0, as a function of sample size per twin group and H of SS twins. Power in each cell was calculated based on the average non-centrality parameter of 1000 simulations under the corresponding condition. Darker cell colors denote lower power. ‘Sample size’ indicates the number of kin pairs in each kin group.



Supplementary Figure S1-2. Power of the ACE model to detect A under the simulated variance of A = .5; C = 2.0; E = .5, as a function of sample size per twin group and H of SS twins. Power in each cell was calculated based on the average non-centrality parameter of 1000 simulations under the corresponding condition. Darker cell colors denote lower power. ‘Sample size’ indicates the number of kin pairs in each kin group.



Supplementary Figure S2-1. Proportion of the fitting results of 1000 simulated datasets that the ACE model has the lowest AIC compared to the AE and CE model under the simulated variance of A = 2.4; C = .3; E = .3, as a function of sample size per twin group and H of SS twins. Darker cell colors denote lower power. ‘Sample size’ indicates the number of kin pairs in each kin group.



Supplementary Figure S2-2. Depicts the proportion of fitting results from 1000 simulated datasets where the ACE model had the lowest AIC compared to the AE and CE models, when the simulated variance is set to A = 1.0; C = 1.0; E = 1.0, as a function of sample size per twin group and H of SS twins. Darker cell colors denote lower power. ‘Sample size’ indicates the number of kin pairs in each kin group.



Supplementary Figure S2-3. Depicts the proportion of fitting results from 1000 simulated datasets where the ACE model had the lowest AIC compared to the AE and CE models, when the simulated variance is set to A = .5; C = 2.0; E = .5, as a function of sample size per twin group and H of SS twins. Darker cell colors denote lower power. ‘Sample size’ indicates the number of kin pairs in each kin group.



Supplementary Figure S3-1. Illustrates the proportion of fitting results from 1000 simulated datasets with at least one negative estimate for A, C, or E variance components, when variance is set to A = 2.4; C = .3; E = .3, as a function of sample size per twin group and H of SS twins. Darker cell colors indicate higher prevalence of negative estimates. ‘Sample size’ indicates the number of kin pairs in each kin group.



Supplementary Figure S3-2. Illustrates the proportion of fitting results from 1000 simulated datasets with at least one negative estimate for A, C, or E variance components, when variance is set to A = 1.0; C = 1.0; E = 1.0. as a function of sample size per twin group and H of SS twins. Darker cell colors denote a higher proportion of models with negative estimates. ‘Sample size’ indicates the number of kin pairs in each kin group.



Supplementary Figure S3-3. Illustrates the proportion of fitting results from 1000 simulated datasets with at least one negative estimate for A, C, or E variance components, when variance is set to A = .5; C = 2.0; E = .5, as a function of sample size per twin group and H of SS twins. Darker cell colors indicate higher prevalence of negative estimates. ‘Sample size’ indicates the number of kin pairs in each kin group.

****** Supplementary Figure S4-1.Depicts the proportion of fitting results from 1000 simulated datasets where the ACE model had the lowest AIC compared to the AE and CE models, when the simulated variance is set to A = 1.5; C = .6; E = .9 (50%, 20%, 30%) and the sex-limited scalar of rc = .95 included, as a function of sample size per twin group and H of SS twins. Darker cell colors denote lower power. ‘Sample size’ indicates the number of kin pairs in each kin group.

******

Supplementary Figure S4-2.Illustrates the proportion of fitting results from 1000 simulated datasets with at least one negative estimate for A, C, or E variance components, when variance is set to A=1.5; C=.6; E=.9 (50%, 20%, 30%) and the sex-limited scalar of rc = .95 included, as a function of sample size per twin group and H of SS twins. Darker cell colors indicate higher prevalence of negative estimates. ‘Sample size’ indicates the number of kin pairs in each kin group.

****

Supplementary Figure S5-1. Average estimates of ‘A’ obtained from 1000 models, each fit to simulate data with variance combination A = 1.0, C = 1.0, E = 1.0 (33.3%, 33.3%, 33.3%). Darker cell colors denote larger deviations from the population parameter A = 1.0. ‘Sample size’ indicates the number of kin pairs in each kin group.

****

Supplementary Figure S5-2. Average estimates of ‘A’ obtained from 1000 models, each fit to simulate data with variance combination A = 2.4, C = .3, E = .3 (80%, 10%, 10%). Darker cell colors denote larger deviations from the population parameter A = 2.4. ‘Sample size’ indicates the number of kin pairs in each kin group.