

Improving Heritability Estimates with Restricted Maximum Likelihood (ReML)



Thomas E Nichols^{1,2}, Karl Friston³, Jonathan Roiser³, Essi Viding³

1. GlaxoSmithKline Clinical Imaging Centre, London, United Kingdom, thomas.e.nichols@gsk.com

2. Oxford FMRIB Centre, Oxford, United Kingdom

3. University College London, London, United Kingdom

Introduction

Heritability is the proportion of variability in a phenotype that can be explained by genetic sources. Heritability is usually measured with twins studies, where the correlation of monozygotic (MZ, identical) and dizygotic (DZ, fraternal) twins indicates shared variation due to common environmental and/or genetic influences.

The simplest estimate of narrow-sense (additive genetic) heritability h^2 is Falconer's estimate (Falconer, 1996), which is just twice the difference of MZ and DZ correlations.

Another estimate comes from a components-of-variance approach, where a structured covariance model expresses the shared genetic and environmental effects (Neale, 1998). This approach has much greater flexibility and can be more powerful (Christian, 1995). While such models are often represented by a Structural Equation Model (Mx, <http://www.vcu.edu/mx>), SEM is just a mechanism to specify the model, and a standard maximum likelihood or restricted maximum likelihood method (ReML) is actually used to estimate the variance parameters.

In this initial work we use simulations to compare the bias, variance and MSE of the Falconer's and REML estimates of h^2 .

Methods

Falconer's estimate is

$$h^2_F = 2(r_{MZ} - r_{DZ})$$

where r_{MZ} is the sample correlation of identical twin pairs and r_{DZ} is the sample correlation of fraternal twin pairs (h^2_F is truncated at zero if negative).

The covariance model has 3 components: An additive genetic term (A), a common environmental term (C), and an independent error term (E). The variance is the same for all subjects,

$$\text{Var}(MZ) = \text{Var}(DZ) = A + C + E,$$

but the covariance depends on the twin type:

$$\begin{aligned} \text{Cov}(MZ_1, MZ_2) &= A + C \\ \text{Cov}(DZ_1, DZ_2) &= A/2 + C. \end{aligned}$$

Heritability is $h^2_R = A/(A+C+E)$.

Our simulation uses four sets of true parameter settings:

- "Null" No correlation
- "A Only" Only additive genetic effects
- "C Only" No genetic but common environmental effects
- "A & C" Both genetic and common environmental effects

For each setting we compare the bias, standard deviation and mean squared error of

the h^2 estimates (12,000 realizations). The true A, C & E parameters satisfy $A + C + E = 1$, and so A can be interpreted as the true h^2 .

Results

Figure 1 shows the bias in h^2 for the two methods for each of the four types of true parameters settings. For only 20 twins the bias in Falconer's method can exceed 0.5 (i.e. 50%), and is particularly bad in the "Null" case, with an average h^2 of 0.7. The bias of ReML never exceeds 0.3. Both models have the worst bias in the "C Only" case, suggesting that the environment term can often be mistaken for heritability.

Figure 2 shows the standard deviation of the h^2 estimators. The standard deviation of Falconer's estimate is often more than double that of the REML method.

Figure 3 shows the results of using model selection techniques based on choosing the model with the best log evidence. For each true parameter setting considered, four models (Null, A only, C only, A&C) were fit and the one with the best log evidence was selected. Even with 100 subjects, the A&C models were never selected, suggesting that much larger samples are needed for accurate inference on model type. This also mirrors the recommendation of Christian (1995) to simply always fit A&C due to poor performance of model comparisons.

Conclusions

Through a Monte Carlo simulation we have shown that Falconer's method, a technique commonly used in the imaging genetics literature, has extremely poor performance, with bias in small samples exceeding 50% and dramatically higher variability relative to the best-practice method for estimating heritability. This suggests there is an extreme risk of false positives with Falconer's method applied to

small samples.

While custom genetics software is often used to find these better REML estimates, we have used the ReML variance components framework of SPM and will be able to rapidly develop voxel-wise estimates of h^2 .

References

- Falconer, DS & Mackay, TFC (1996), *Introduction to Quantitative Genetics*, 4th Ed, Longmans Green, Harlow, UK.
- Christian, JC (1995), "Comparison of analysis of variance and maximum likelihood based path analysis of twin data: partitioning genetic and environmental sources of covariance", *Genetic Epidemiology*, vol. 12, no. 1, pp. 27-35.
- Neale, MC & Cardon, LR, eds. (1992), *Methodology for Genetic Studies of Twins and Families*, Cluwer, Dordrecht.
- Neale, MC (1998), "Twin studies", *Encyclopedia of Biostatistics*, Wiley, Hoboken, NJ.

Figure 1. Bias of Falconer's and ReML heritability estimates. Falconer's greatly over-estimates heritability for small samples. ReML has negligible bias when A & C are both non-zero.

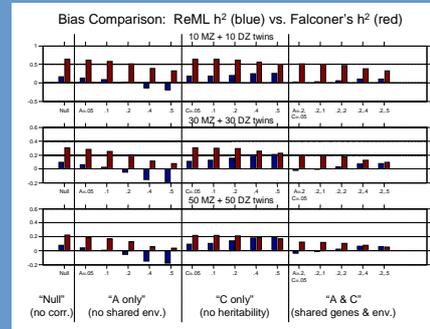


Figure 2. Standard deviation of Falconer's and ReML heritability estimates. With low n Falconer's has 3- to 4-times greater standard deviation than ReML's h^2 estimates.

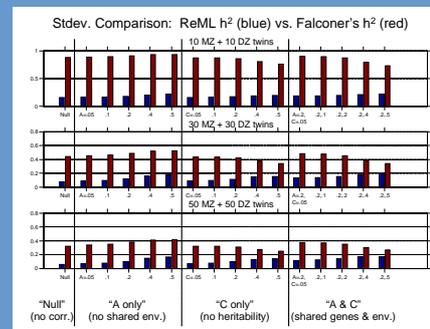


Figure 3. Accuracy of model selection with ReML. Even with 100 subjects the A & C models are almost never correctly selected, suggesting an A&C model should be used by default.

