Published in final edited form as:

Stat Biosci. 2016 October ; 8(2): 264–283. doi:10.1007/s12561-016-9143-2.

Power and Effective Study Size in Heritability Studies

Jesse D. Raffa and

University of Washington, Department of Statistics, Box 354322, Seattle, WA, 98195-4322,
U.S.A., Phone: 206-685-8145

Elizabeth A. Thompson

University of Washington, Department of Statistics, Box 354322, Seattle, WA, 98195-4322, U.S.A

Jesse D. Raffa: jraffa@mit.edu; Elizabeth A. Thompson: eathomp@u.washington.edu

Abstract

Correlation between study units in quantitative genetics studies often makes it difficult to compare important inferential aspects of studies. Describing the relatedness between study units is critical to capture features of pedigree studies involving heritability, including power and precision of heritability estimates. Blangero et al (2012) showed that in pedigree studies the power to detect heritability is a function of the true heritability and the eigenvalues of the kinship matrix. We extend this to a more general setting which allows statements about expected precision of heritability estimates. Using two different Taylor series approximations, we summarize the relatedness in a study design by one or two parameters. These relatedness summary parameters (RSPs) are functions of the eigenvalues or log-eigenvalues of the kinship matrix. Using the RSPs based on the log-eigenvalues, we accurately approximate the expectation of the likelihood ratio test and expected confidence interval widths. We define an effective sample size of a target study as one which has the equivalent power and precision to a reference design. Using unrelated sibpairs as the reference design provides very accurate assessments of power. RSPs and effective sample sizes provide new tools for comparing studies and communicating information about relatedness in heritability studies.

Keywords

eigenvalues; expected log-likelihood; kinship matrix; log-eigenvalues; pedigree; precision; relatedness summary parameters; variance of the eigenvalues

1 Introduction

The sample size of a study is fundamental quantity in all areas of statistics. It dictates many aspects of a study, including its power and the precision of estimates it generates. When all other aspects of two studies are the same, the interpretation of sample size is straightforward – larger is better. However, this applies only to studies involving independent study units.

Software

The authors have made R code for calculating the relatedness summary parameters (RSPs) from a .ped file and other calculations available on GitHub. Please visit: <https://www.github.com/jraffa/pedRSP> for more information.

The situation can be much more complex in cases where there is dependence among study units [1], for example when studying genetically related individuals, where inherited traits are correlated between individuals who share a common ancestor [10]. Discussing power and precision in this setting involves describing both the sample size and relatedness of the observed individuals. This often makes tasks such as calculating power or comparing study designs very difficult.

One situation where this occurs is when a quantitative genetic trait is studied in pedigrees under an additive genetic model, where the primary inferential focus is heritability (i.e., the proportion of variance of the trait explained by genetic effects) [10]. Then, the correlation between two trait values depends on how related the two individuals are and on the heritability of the trait. When such studies are comprised solely of pairs of individuals of the same type of relationship who are otherwise considered unrelated (e.g., studies of unrelated sibpairs or twins [15]), it is straightforward to understand how altering the sample size will affect inference. However, recent studies [3] of heritability often involve complex multi-generational pedigrees which are more difficult to describe.

Communicating the relatedness between individuals can be done in several ways (Figure 1). Historically, describing relatedness involved representing the pedigree structure graphically (Figure 1a). This format identifies the individual as a founder (i.e., not related to anyone except through their own children), or identifies their parents. An important quantity for estimating heritability is the twice the kinship matrix [12] (Figure 1b). For N individuals, this is an N by N positive semi-definite matrix which describes the expected proportion of the genome shared between two non-inbred individuals via a common ancestor. Even where confidentiality concerns and feasibility would allow publication of a pedigree or full kinship matrix, there is no obvious way to interpret this information statistically. Simpler summaries are needed, both in the context of planning a new study (e.g., what types of pedigrees should one try to recruit) and in interpreting the results in the literature (e.g., given two conflicting results, which study has stronger evidence).

Blangero et al (2012) [3] recently demonstrated the usefulness of performing an eigendecomposition of the kinship matrix in calculating power for detecting heritability. This approach showed that the power to detect heritability is a function of eigenvalues of the kinship matrix and the true heritability (Figure 1c). In contrast to previous methods, which often calculated power through simulation [19], this approach revealed that the effect of the relatedness between individuals on inference can largely be captured by the N eigenvalues of the kinship matrix. However, Blangero's approach [3] requires knowledge of N eigenvalues, and in even moderately sized studies, presenting this information in full, in a manner others can use could be impractical and inefficient.

In genetic studies, there is a need to have simple interpretable quantities, similar to sample size in studies involving independent data, but accommodating the relationships between study subjects. In this paper, we extend the approach of Blangero et al (2012) to the general setting of testing an arbitrary null-hypothesis (Section 2). This is then used to derive a method to compute expected precision of heritability estimates. In Section 3, we approximate this general result using Taylor series approximations, which simplify to be

functions of quantities we call relatedness summary parameters (RSPs). These RSPs are one or two simple functions of the eigenvalues of the kinship matrix (Figure 1d). The RSPs are then used to develop the idea of an effective sample size by relating the power or precision of a study of interest to a reference study design (Section 4). We demonstrate through simulation that these RSPs are an effective way to describe the impact of relatedness in a study on power and precision (Section 5). We conclude with a discussion (Section 6).

2 Estimation of Heritability in Pedigree Studies

Let y_i be an additive polygenic trait for subject i ($i = 1, \dots, N$), and write $\mathbf{y} = [y_1, y_2, \dots, y_N]$ as a vector of trait values for all N individuals in the study. Under the standard, additive polygenic model, $\mathbf{y}$ is distributed as multivariate normal (MVN) as follows:

$$\mathbf{y} \sim MVN(\mathbf{X}'\boldsymbol{\beta}, \boldsymbol{\Sigma}_p), \quad (1)$$

where $\mathbf{X}$ is the covariate design matrix for any relevant fixed effects, $\boldsymbol{\beta}$ is a vector of regression coefficients for the covariates, and $\boldsymbol{\Sigma}_p$ is the phenotypic covariance. Under an additive polygenic model $\boldsymbol{\Sigma}_p = \mathbf{K}\sigma_G^2 + \mathbf{I}_N\sigma_e^2$ where $\mathbf{K}$ is twice the kinship matrix, Φ , a $N \times N$ matrix calculated through the pedigree information, σ_G^2 is the additive genetic variance, $\mathbf{I}_N$ is the $N \times N$ identity matrix, and σ_e^2 is the individual residual variance. Under this model, the heritability, h^2 of this trait is $\sigma_G^2 / (\sigma_G^2 + \sigma_e^2)$. Rewriting the phenotypic covariance to be in terms of h^2 and the phenotypic variance, σ_p^2 , $\boldsymbol{\Sigma}_p = [\sigma_p^2 h^2] \mathbf{K} + [\sigma_p^2 (1 - h^2)] \mathbf{I}_N$ with log-likelihood:

$$l(\boldsymbol{\beta}, \sigma_p^2, h^2) = c - \frac{1}{2} \log(|\boldsymbol{\Sigma}_p|) - \frac{1}{2} (\mathbf{y} - \mathbf{X}'\boldsymbol{\beta})' \boldsymbol{\Sigma}_p^{-1} (\mathbf{y} - \mathbf{X}'\boldsymbol{\beta}), \quad (2)$$

where log here and throughout is the natural-logarithm.

2.1 Eigendecomposition of the Kinship matrix

Eigendecomposition of twice the kinship matrix can ease computation and lead to tractable analytic manipulation of the likelihood, (2). The eigendecomposition of $\mathbf{K}$ is written as $\mathbf{U}\mathbf{S}\mathbf{U}'$, where $\mathbf{U}$ is an $N \times N$ matrix with the eigenvectors of $\mathbf{K}$ as its columns and $\mathbf{S}$ is a $N \times N$ diagonal matrix with the diagonal elements corresponding to the N eigenvalues of $\mathbf{K}$, λ_i . We will consider $\mathbf{K}$ which are positive definite or positive semi-definite, which places a positive or non-negative constraint on λ_i , respectively.

Using this decomposition of $\mathbf{K}$ we can rotate the vector of trait values, $\mathbf{y}$, by $\mathbf{U}'$, under which the rotated trait, $\mathbf{U}'\mathbf{y}$ has a diagonal covariance matrix, $\mathbf{D}$, and mean vector, $\mathbf{U}'\mathbf{X}'\boldsymbol{\beta}$ giving the likelihood:

$$l(\beta, \sigma_p^2, h^2) = c - \frac{1}{2} \log(|\mathbf{D}|) - \frac{1}{2} (\mathbf{U}'\mathbf{y} - \mathbf{U}'\mathbf{X}'\beta)' \mathbf{D}^{-1} (\mathbf{U}'\mathbf{y} - \mathbf{U}'\mathbf{X}'\beta). \quad (3)$$

Both the inverse and log determinant of $\mathbf{D}$ can be easily calculated. The i^{th} diagonal element of $\mathbf{D}$ is $[\sigma_p^2(1+h^2(\lambda_i-1))]$, so then $\log(|\mathbf{D}|) = N \log(\sigma_p^2) + \sum_{i=1}^N \log(1+h^2(\lambda_i-1))$. This result has been previously described [21], and has been used recently by others [13].

2.2 Expected Likelihood Ratio Test

The expected (log) likelihood ratio test (ELRT) is one method to derive statistical power for general hypothesis testing [4]. The ELRT is the expectation of the twice the difference of the log-likelihoods evaluated at the true parameter and the null-parameter values, respectively. An analogous statistic is widely used in genetic linkage analysis [17], where it is known as the ELOD (expected LOD score), and can be characterized by evaluating the effective number of informative meioses [9]. In our setting, the null hypotheses will specify some heritability, h_0^2 and the data will be generated with some true heritability, h^2 . For the detection of heritability, $h_0^2=0$ is the null hypothesis of interest, but the more general form allows for analyses of precision through the construction of confidence intervals.

Using the log-likelihood form in (3) evaluated at h_0^2 , noting that c and $\log(|\mathbf{D}_0|)$ do not depend on the trait data, we need only calculate the expectation of quadratic-form,

$\mathbb{E}[(\mathbf{U}'\mathbf{y} - \mathbf{U}'\mathbf{X}'\beta)' \mathbf{D}_0^{-1} (\mathbf{U}'\mathbf{y} - \mathbf{U}'\mathbf{X}'\beta)]$ where $\mathbf{D}_0$ is the matrix $\mathbf{D}$ evaluated at the null hypothesis heritability, h_0^2 . When σ_p^2 is known, this expectation will simplify as:

$$\mathbb{E}[(\mathbf{U}'\mathbf{y} - \mathbf{U}'\mathbf{X}'\beta)' \mathbf{D}_0^{-1} (\mathbf{U}'\mathbf{y} - \mathbf{U}'\mathbf{X}'\beta)] = \text{trace}(\mathbf{Q}\mathbf{D}_0^{-1}) + \mu' \mathbf{D}_0^{-1} \mu,$$

where $\mathbf{Q} = \text{Var}(\mathbf{U}'(\mathbf{y} - \mathbf{X}'\beta)) = \mathbf{U}'\Sigma_p\mathbf{U} = \mathbf{U}'\mathbf{U}\mathbf{D}\mathbf{U}'\mathbf{U} = \mathbf{D}$ and $\mu = \mathbb{E}[\mathbf{U}'(\mathbf{y} - \mathbf{X}'\beta)] = \mathbf{0}$.

Thus the expectation of the quadratic form is the trace of the matrix product of two diagonal matrices, $\text{trace}(\mathbf{D}\mathbf{D}_0^{-1})$, and is equal to $\sum_{i=1}^N \{[1+h^2(\lambda_i-1)][1+h_0^2(\lambda_i-1)]^{-1}\}$. This results in expected log-likelihood:

$$\mathbb{E}[l(h_0^2)] = c - \frac{N}{2} \log(\sigma_p^2) - \frac{1}{2} \sum_{i=1}^N \log(1+h_0^2(\lambda_i-1)) - \frac{1}{2} \sum_{i=1}^N \frac{1+h^2(\lambda_i-1)}{1+h_0^2(\lambda_i-1)}. \quad (4)$$

We apply (4) and obtain:

$$ELRT = 2\mathbb{E}[l(h^2)] - 2\mathbb{E}[l(h_0^2)] = -\sum_{i=1}^N \log(1+h^2(\lambda_i-1)) - (N-1) + \sum_{i=1}^N \log(1+h_0^2(\lambda_i-1)) + \sum_{i=1}^N \frac{1+h^2(\lambda_i-1)}{1+h_0^2(\lambda_i-1)}.$$

(5)

The case $h_0^2=0$ is considered by Blangero [3]. In this case, most of the terms disappear or cancel, leading him to the result:

$$ELRT = -\sum_{i=1}^N \log(1+h^2(\lambda_i-1)).$$

(6)

Blangero [3] notes that (6) depends only on the true heritability, h^2 and the N eigenvalues of $\mathbf{K}$. ELRT and therefore the power are highest for a fixed h^2 when there are many small eigenvalues, λ_i . Simplification of (5) for other values of h_0^2 is generally not possible. For (6), when $h_0^2=0$, different ELRT values must be used, due to the null parameter value being on the boundary of the parameter space. In this situation, the correct likelihood ratio test cut-off has been discussed extensively in the literature [3, 22].

Using the ELRT in the form (5) under a large enough sample size, we can calculate the power for testing $h^2=h_0^2$. The power function for testing $h^2=h_0^2$ at the α significance level is $1 - F(x_\alpha; \nu, \xi)$, where F is the cumulative distribution function of a non-central $\chi^2_{\nu, \xi}$ distribution. Here, $\nu = 1$ degrees of freedom and non-centrality parameter, $\xi = ELRT - 1$, evaluated at x_α , the $1 - \alpha$ quantile for a $\chi^2_{\nu=1, \xi=0}$ distribution with 1 degree of freedom, and non-centrality parameter of zero. If h_0^2 is strictly greater than zero, $\nu = 1$ and a computed ELRT from (5) of about 6, 9, and 12 will have approximately 60%, 80% and 90% power respectively at a 0.05 significance level. While the general null hypothesis case may be of interest in some situations, one of the useful aspects of considering $h^2 > 0$ comes from the ability to derive expected precision of heritability estimates by inverting the ELRT. We will examine this in Section 5.3.

3 Approximations to the ELRT by Taylor Series Expansion

The ELRT of the form (5) requires knowledge of all N eigenvalues to calculate. We provide approximations to (5) which rely on one or two simple summary functions of the eigenvalues of $\mathbf{K}$, which allow for easier interpretation and communication of the effect of the relatedness on power and precision.

3.1 Approximation A

One property of the matrix $\mathbf{K}$ is that for non-inbred individuals the diagonal consists solely of ones. Since the sum of the eigenvalues of $\mathbf{K}$ is the trace of the matrix [11, pp. 545–6], this

shows that the mean of the eigenvalues, $\bar{\lambda}$, then will also be one. We will use this property to expand a second-order Taylor series for (5) around $\lambda_i = \bar{\lambda} = 1$.

$$\begin{aligned} ELRT \approx N[-\log(1+h^2(\bar{\lambda}-1))+\log(1+h_0^2(\bar{\lambda}-1))+\frac{1+h^2(\bar{\lambda}-1)}{1+h_0^2(\bar{\lambda}-1)}] \\ - (N-1) + \frac{(h^2-h_0^2)^2(\bar{\lambda}-1)}{(h^2\bar{\lambda}-h^2+1)(h_0^2\bar{\lambda}-h_0^2+1)^2} \sum_{i=1}^N (\lambda_i-\bar{\lambda}) \\ + \frac{(h^2-h_0^2)^2(4h_0^2h^2\bar{\lambda}-2h^2h_0^2\bar{\lambda}^2-2h^2h_0^2-h_0^2\bar{\lambda}+h_0^2+1)}{2(h_0^2\bar{\lambda}-h_0^2+1)^3(h^2\bar{\lambda}-h^2+1)^2} \sum_{i=1}^N (\lambda_i-\bar{\lambda})^2. \end{aligned}$$

When $\bar{\lambda} \approx 1$, the first two terms which do not involve $\sum_{i=1}^N (\lambda_i-\bar{\lambda})^2$ will be zero or small, which simplifies to an Approximate ELRT (AELRT),

$$ELRT \approx \frac{(N-1)V_{\lambda}(h^2-h_0^2)^2}{2} + 1 = AELRT_A, \quad (7)$$

where V_{λ} is the variance of the eigenvalues, $\frac{1}{N-1} \sum_{i=1}^N (\lambda_i-1)^2$. As noted for (6), power comes from eigenvalues that are small, and this approximation may not adequately capture this feature because V_{λ} is dominated by the large eigenvalues.

The remainder term in a Taylor series can be used to estimate the error of the approximation.

The third order term of Taylor series of the ELRT is $-\frac{1}{3}(h^2+2h_0^2)(h^2-h_0^2)^2 \sum_{i=1}^N (\lambda_i-1)^3$. Since both $h^2 \geq 0$ and $h_0^2 \geq 0$, the term outside the sum will always be negative, with the sign of the error being dictated by the sign of the third-central moment of the eigenvalues of $\mathbf{K}$. Due to the constraint on λ_i , the third central moment will almost always be positive, which in turn implies that this approximation will be greater than or equal to the true ELRT for all h^2 and h_0^2 . The main exception to this is when the eigenvalues are symmetric about the mean, which would most commonly occur when a pedigree consisted of pairs of related individuals who are otherwise unrelated to all other members in the study. Under this scenario, the third central moment, $\sum_{i=1}^N (\lambda_i-1)^3=0$, and the remainder would depend on the fourth-order term, $\frac{1}{4}(h^2+2h^2h_0^2+3h_0^2)(h^2-h_0^2)^2 \sum_{i=1}^N (\lambda_i-1)^4$, which is always positive, indicating that the approximation will be less than or equal to the true ELRT for all h^2 and h_0^2 . In both cases, in addition to the third or fourth central moment of the eigenvalues, the error will be proportional to the squared difference between the h^2 and h_0^2 suggesting good behaviour local to the true heritability, provided that the third-central moment is not too large. In complex pedigrees scenarios, the eigenvalues of $\mathbf{K}$ can be relatively large, which can in turn result in large values for the third central moment. In such scenarios, this approximation to the ELRT is likely not precise, and an alternative approach is needed.

3.2 Approximation B and C

As noted for (7), the approximation will tend to perform poorly when the study pedigrees are complex, as the positive-(semi)-definite matrix $\mathbf{K}$ puts a constraint on the small eigenvalues, but not on the larger ones. Provided $\mathbf{K}$ is strictly positive-definite, we can generally improve on (7) by basing the approximation around the natural logarithm of the eigenvalues [7]. Writing $q_i = \log(\lambda_i)$, we can rewrite (5) in terms of q_i as:

$$ELRT = -\sum_{i=1}^N \log(1+h^2(\exp(q_i)-1)) - (N-1) + \sum_{i=1}^N \log(1+h_0^2(\exp(q_i)-1)) + \sum_{i=1}^N \frac{1+h^2(\exp(q_i)-1)}{1+h_0^2(\exp(q_i)-1)}. \quad (8)$$

Using the same approach as in Section 3.1, we expand a Taylor series around $\log(\gamma)$, where γ is the geometric mean of the eigenvalues, $\lambda_{\bar{r}}$. In this case, after simplifying, we derive an approximation for the ELRT as:

$$ELRT \approx N[-\log(\frac{g_{\gamma}(h)}{g_{\gamma}(h_0)}) + \frac{g_{\gamma}(h)}{g_{\gamma}(h_0)} - 1] + 1 + \frac{(N-1)V_q\gamma}{2} [\frac{h_0^2(h_0^2-1)}{g_{\gamma}(h_0)^2} + \frac{h^2(h^2-1)}{g_{\gamma}(h)^2} - \frac{\gamma(h_0^2\gamma+h_0^2-1)(h^2-h_0^2)}{g_{\gamma}(h_0)^3}] = AELRT_B, \quad (9)$$

with $g_{\gamma}(h) = 1 + h^2(\gamma - 1)$ and $V_q = \frac{1}{N-1} \sum_{i=1}^N (q_i - \log(\gamma))^2$, the variance of the log eigenvalues. Under this approximation, V_q is no longer dominated by the large eigenvalues as was the case for V_{λ} in approximation A. While (9) is much more complex than (7), it is in general a better approximation to the true ELRT in most scenarios, as we will demonstrate. It is also important to note that while it is algebraically more complicated than (7), it depends on only one additional parameter, the geometric mean of the eigenvalues, γ .

The remainder term of this Taylor series is also more complicated than the approximation discussed in Section 3.1. In general, the absolute error will be of order $\sum_{i=1}^N (q_i - \log(\gamma))^3$, but whether it over- or under- estimates the exact ELRT, equation (5), will also depend on γ , h^2 and h_0^2 . As we will show in the next section, both over- and under- estimation can occur. As was the case with approximation A, (9) is generally going to be a good approximation local to the true heritability.

It is worth noting that for $\gamma \approx 1$, a useful first-order expansion of (9) around $\gamma = 1$ yields an approximate ELRT of $(N-1)V_q(h^2-h_0^2)^2/2 = AELRT_C$, which is of the same form as approximation A with the variance of the log eigenvalues replacing the variance of the eigenvalues. This form will be used to derive analytical results in subsequent sections. As noted in Section 3.1, the Approximation A will likely overestimate the true ELRT. Since V_q in most contexts, use of V_q will provide a better approximation in most situations.

4 An Effective Sample Size for Heritability Studies, Approximation

ALERT_{ESS}

Approximations to the ELRT provides a means to estimate an effective sample size relative to some reference study design. This is a measure of how many of a particular reference design or unit would have approximately equal ELRT and therefore power to the design of interest. Given the wide use of sibpair studies in genetic epidemiology, we have chosen unrelated sibpairs as our reference study unit.

To derive an effective sample size of sibpairs, N_r , for our target design with N individuals, we can equate the approximate ELRT of our design with the approximate ELRT of N_r individuals in sibpairs ($N_r/2$ pairs). Using this approach for approximation A,

$$\frac{AELRT_{Ar}}{(N_r-1)V_{\lambda}^*(h^2-h_0^2)^2+1} = \frac{AELRT_A}{(N-1)V_{\lambda}(h^2-h_0^2)^2+1} = \frac{(N-1)V_{\lambda}}{V_{\lambda}^*} + 1 \quad (10)$$

where V_{λ}^* is the variance of the eigenvalues for sibpairs. Based on this approximation we can see that the effective sample size is approximately equal to the ratio of V_{λ} and V_{λ}^* times the study sample size, N . One can use the same approach with approximation B, but these effective sample sizes will be a function of both h^2 and h_0^2 which make them more difficult to use. We can also make use of $AELRT_C$, when γ is close to 1, and replace V_{λ} and V_{λ}^* in (10) with V_q and V_q^* , respectively to get $N_r = [(N-1)V_q/V_q^*] + 1$. This latter approximation works reasonably well, in most scenarios. Using an approximation based on the effective sample size in (10), will be equivalent to using approximation A directly in most scenarios and will therefore approximate the exact ELRT, (5), well only when approximation A does, (which we shall see is primarily in non-complex pedigree scenarios), while using an effective sample size based on $AELRT_C$ and V_q will provide better approximation in scenarios where it performs well. We will denote the approximate ELRT based on $N_r/2$ effective sibpairs when $N_r = [(N-1)V_q/V_q^*] + 1$ as $AELRT_{ESS}$

Given a effective sample size, N_r for a sibpair design, we can write out the ELRT from (5), plugging in the two eigenvalues from this design, 1/2 and 3/2 as:

$$\begin{aligned} AELRT_{ESS} &= \sum_{i=1}^{N_r/2} ELRT(\text{One Sibpair}) \\ &= \frac{N_r}{2} \left[\log\left(\frac{4-h_0^4}{4-h^4}\right) + \frac{8-h_0^2h^2}{4-h_0^4} \right] - (N_r-1). \end{aligned} \quad (11)$$

Using an effective sample size provides a means to communicate the size of a study on a familiar and standardized scale, and another method to approximate the the exact *ELRT*, (5) without the need to know all N eigenvalues of $\mathbf{K}$. As we shall see in Section 5, this method

provides a quite accurate approximation of the exact ELRT, (5), and is easy to interpret and calculate.

5 Comparing simulated and exact ELRT to the Approximations

Given the four approximations from Sections 3 and 4, we compared the analytic form of (5) and the ELRT estimated under simulation to $AE\text{LRT}_A$ (7), $AE\text{LRT}_B$ (9), $AE\text{LRT}_C$ and $AE\text{LRT}_{ESS}$ (11).

5.1 Overview of Pedigrees Used

We used three study designs chosen to contrast the approximations. The first study design used 1500 unrelated sibpairs, for a total sample size of 3000. The second study design is from a complex pedigree, distributed as part of a previous Genetics Analysis Work Shop (GAW13) [2]. This simulated pedigree roughly corresponds to those individuals with measured phenotypes in the Framingham Heart Study (FHS) by combining both of the cohorts (separate recruitment periods) from simulation replicate 1 [8]. This is a simulated pedigree, but has many of the same features as the actual FHS pedigree, with sample size, $N = 2860$. The last pedigree originates from a study of a neuro-degenerative disease from an isolated community in Guam [20]. There is a high degree of relatedness among individuals in this study; the median number of relatives (non-zero kinship between individuals) per study subject was 132 (range: 1–267). This contrasts with the other two studies where study participants were related to only one other person in the study in the sibpair example by definition, and had a median of eight relatives in the GAW13 example (range: 2–45).

One complication arose in the Guam example, where the small sample size (320) reduces the applicability of the asymptotic results that the likelihood ratio test and ELRT are based on. This is not an issue with the approximations per se, but rather the fit of the exact ELRT (5) to the mean LRT observed under simulation. In this paper we focus on large sample sizes, and evaluate the approximations' fit to the empirical ELRT under situations where the asymptotic result are most applicable. In the Guam example this meant reproducing the identical design three times so that the study had three large pedigrees of the original size (320) for a total of $N = 960$ subjects, with the three pedigrees being unrelated to each other. We refer to this dataset as Triple Guam.

For each of the three studies, we report the RSPs from approximation A (V_λ) and approximation B (γ and V_q) in Table 1. The dataset with the most structure, Triple Guam, has the highest values for both V_λ and V_q , and the dataset with the least structure, 1500 sibpairs, has the lowest values of V_λ and V_q . The distributions of eigenvalues and log-eigenvalues are shown in Figure 2. In the GAW13 and Triple Guam pedigrees the distribution of the eigenvalues is asymmetric, but the skew of the distribution is reduced by taking logarithms (Figure 2).

In order to calculate $AE\text{LRT}_{ESS}$, we need to calculate an effective sample size from (10), using either V_λ or V_q . In the GAW13 pedigree the effective sample size using (10) is about 6396 sibpairs ($N_r = 12792$ individuals), whereas using the effective sample size based on V_q and V_q^* from Table 1, the effective sample size is about 3590 sibpairs ($N_r = 7180$).

individuals). This can be interpreted as 3590 sibpairs have approximately the same ELRT and approximately the same power as the GAW13 pedigree using the latter approximation. Similarly, in the Triple Guam pedigree the effective sample size using V_λ is about 5036 sibpairs ($N_r = 10072$ individuals), whereas using the effective sample size based on V_q is about 1511 sibpairs ($N_r = 3022$ individuals). The estimates of effective sample size based on V_λ versus V_q differ drastically in both studies. In general, the approximation based on V_q is far more accurate, and will be used to calculate $ALERT_{ESS}$ throughout the remaining discussion. It is noteworthy that often in complex pedigrees, the effective number of individuals will be larger than the actual size of the study. This illustrates the utility of using complex pedigrees, where the structured dependence within a large family can be leveraged to estimate heritability more efficiently than in a large number of pairs of unrelated siblings.

5.2 ELRT and Power Simulation Study

For values of the true heritability, h^2 , 0.3, 0.5 and 0.7, J datasets were simulated under an additive polygenic model (see equation 1) for pedigrees in Table 1 for a fixed intercept, $\beta = 3$ and environmental variance, $\sigma_e^2 = 64$ (The ELRT and approximations do not depend on the specific values of β and σ_e^2). For each of the $j = 1, 2, \dots, J$ simulated dataset, the maximum likelihood estimate $\hat{h}^2$ was determined, and the likelihood ratio test statistic under a several different null-values, $h_0^2: L_j(h_0^2) = 2l(\hat{h}^2) - 2l(h_0^2)$. For a given h_0^2 the likelihood ratio tests were averaged over the J simulations to get an estimate of the ELRT at the particular value of $h_0^2: \bar{L}(h_0^2) = \frac{1}{J} \sum_{j=1}^J L_j(h_0^2)$. These estimates of the ELRT were then compared against the exact ELRT, (5) and the approximations.

For the three simulation values ($h^2 = 0.3, 0.5$ and 0.7), Figure 3 shows the exact ELRT and the approximations under the three pedigrees of Table 1 and the simulation base estimates of ELRT over the $J = 10000$ simulations. For the pedigree of sibpairs, all approximations provide a good fit to the true ELRT and simulated ELRT. Approximation A in this instance underestimates the true ELRT. This is consistent with the remainder term to the approximation discussed in Section 3.1, for related pairs of individuals who are otherwise unrelated to all other subjects in the study sample. This contrasts with the GAW13 and Triple Guam pedigrees, where approximation A severely overestimates the true ELRT. This is also expected given that the eigenvalue distributions has positive third central moment in both cases. Approximation B provides a good fit in most instances, particularly local to the true heritability, and when true heritability is not near the boundaries of the parameter space at 0 and 1. Further, it is worth noting that approximation B can both over and under estimate the true ELRT for different values of null-heritability but under the same pedigree and true heritability. Approximation C is generally as or more accurate than approximation B, which is noteworthy given its relatively simple form. Unfortunately, no consistency on when approximation C over- or under- estimates the exact ELRT is apparent. Over the scenarios we examined, $ALRT_B$ and $ALRT_C$ seem to bound the true $ELRT$, and computing both of these values may give some assessment of how accurate these approximations may be. In general, $ALRT_{ESS}$ was most accurate, but was observed to both over- and under-estimate the exact ELRT.

In Figure 4, we use the ELRT results when the true heritability, h^2 , is 0.3 and 0.7 to to assess how well the approximations will translate to statistical power. For a given level of true heritability, for each simulated dataset, a likelihood ratio test statistics,

$L_j(h_0^2) = 2l(\hat{h}^2) - 2l(h_0^2)$ was computed over a grid of null values, and the proportion of times the test was rejected at each null value at a significance level ($\alpha = 0.05$) was estimated.

When $h_0^2 = h^2$, the proportion of time the null was rejected was between 0.049 and 0.054 over each of the six plots. As can be seen in Figure 4, while approximation A tends to severely over-estimate power, and approximation B may underestimate power for some null-hypotheses, approximations C and ESS provide fairly accurate assessments of power over the scenarios we examined. This observation is consistent with our results from Figure 3

In order to determine the properties of the our preferred approximation on power, the maximum and minimum deviations from the true exact power were calculated to understand the worst case scenarios for over- and underestimation of the true power, respectively. For a fixed true heritability, the true power function was calculated for the GAW13 and Triple Guam studies, and denoted $\phi_{h^2}(h_0^2)$. Similarly, the same power function was approximated using $ELRT_{ESS}$ for the estimated number of effective sibpair individuals, N_e , and denoted $\phi_{h^2}(h_0^2)$. For a particular value of h^2 , the maximum and minimum values of $\Delta(h^2, h_0^2) = \tilde{\phi}_{h^2}(h_0^2) - \phi_{h^2}(h_0^2)$, denoted $\max(h^2)$ and $\min(h^2)$, respectively, was determined over a range of h_0^2 values, [0.01, 0.8]. This process was repeated over a range of true heritability values, h^2 , over the interval [0.01, 0.8].

As can be seen in Figure 5, over a range of values of true heritability, the power approximation based on $ELRT_{ESS}$ approximates the exact theoretical power very well, with the maximum overestimation of power, never exceeding 6.5% in either study, while the maximum underestimation was negligible over a large range of interior values (0.2–0.7) of h_0^2 . The peculiar behaviour for small true heritability values, is due primarily to the presence of the boundary on the parameter space at $h^2 = 0$. This would suggest that some care may be required when the true heritability is thought to be low (< 0.2), where the power approximation is likely to be an underestimate of the theoretical power under (5). When the true heritability is at an intermediate level $h^2 \in [0.2, 0.7]$, or when the primary concern is overestimating power, the approximation based on $ELRT_{ESS}$ is particularly useful. The values in Figure 5 represent the worst case scenarios, and the exact amount of over- or under- estimation will be much less over a large range of null hypotheses.

5.3 Deriving Approximate Expected Confidence Interval Widths for h^2

A likelihood ratio test rejection region can be inverted to derive confidence intervals for the true parameter of interest [6, pp. 420–5]. In the same way, the ELRT approximations, (7) and (9) can be inverted to derive expected confidence interval widths, which may be understood as the expected precision of heritability estimates resulting from repeated sampling from a particular pedigree design.

In Figure 3, a lower horizontal line at 3.84 is present, $\chi_{1,0,(1-\alpha)=0.95}^2$, indicating the 0.95 quantile of a chi-squared random variable with one degree of freedom, which is the cut-off

typically used to construct 95% confidence intervals by inverting the rejection region of a likelihood ratio test. The interior of the intersection of the convex curves with this horizontal line provides an estimate of the expected 95% confidence interval for h^2 . For approximation A, (7), we can provide an analytic formula for an expected $(1 - \alpha) \times 100\%$ confidence

interval: the solution for h_0^2 to $(h^2 - h_0^2)^2 V_\lambda (N-1)/2 = \chi_{1,0,1-\alpha}^2$ is $h^2 \pm z_{\alpha/2} \sqrt{\frac{2}{V_\lambda (N-1)}}$, where $z_{\alpha/2}$ is the $1 - \frac{\alpha}{2}$ quantile from a standard Normal distribution.

Using approximation B to construct expected confidence interval widths can be done numerically with standard root solving software to solve $[AELRT_B - \chi_{1,0,(1-\alpha)=0.95}^2] = 0$ for h_0^2 . One can construct expected confidence interval widths numerically in a similar fashion, using $ALERT_{ESS}$ in place of $AELRT_B$. Because approximation A will produce intervals that are generally too narrow, the solutions to these equation is likely to be in the intervals (0,

$h^2 - z_{\alpha/2} \sqrt{\frac{2}{V_\lambda (N-1)}}$) and $(h^2 + z_{\alpha/2} \sqrt{\frac{2}{V_\lambda (N-1)}}, 1)$ provided the solutions for h_0^2 do exist in (0,1). Alternatively, we can apply the first order approximation, $AELRT_C$, from Section 3.2 which results in an expected confidence interval width similar to the form produced under

approximation A, but with V_q replacing V_λ : $h^2 \pm z_{\alpha/2} \sqrt{\frac{2}{V_q (N-1)}}$.

We conducted a simulation study using the same models, parameters and pedigrees as was done in Section 5 varying true heritability from 0.2 to 0.8. For each level of true heritability, we calculated the 0.025 and 0.975 simulation quantiles for h^2 , along with the average upper and lower limits of 95% confidence intervals based on Wald intervals from $J = 10000$ simulations iterations on each pedigree. An estimate of these quantities based on approximations A, B, C and ESS were also determined over the interval [0.2, 0.8]. The plots are presented in Figure 6. As can be seen in Figure 6, the 95% confidence intervals width can vary considerably over the range of values we examined. This is true for both the empirically estimated widths, estimated via the simulated quantiles of h^2 or the empirical Wald intervals, as well as the the exact theoretical widths determined via (5). In the Triple Guam example, the maximum 95% confidence intervals width (occurring around $h^2 = 0.3$) is over 60% wider than the narrowest width we considered at true heritability 0.8.

Approximations A and C have constant width across all values of true heritability, and are not able to approximate the changing width behaviour, but with the approximation based on $AELRT_C$ performing much better than approximation A for most values of h^2 .

Approximation B captures the changing width behaviour quite well for large values (> 0.4) of h^2 , but severely overestimates the width for small h^2 . The effective sibpair approximation does capture the change in confidence intervals width over the range of values of h^2 , but is too wide for some values of h^2 , and too narrow in others. For the complex pedigrees (GAW13 and Triple Guam), the ESS approximation overestimates the width for small (< 0.2) and large (> 0.7) values of h^2 , and moderately underestimates widths for intermediate values ($0.2 < h^2 < 0.7$). This is consistent with what would be expected based on Figure 5. In Figure 6, we have also inserted the maximum 95% confidence interval width for the ESS approximation (confidence intervals width at $h^2 = 0$, which may be useful as a conservative estimate of precision, which seems to bound or be very close to the maximum width in the GAW13 and Triple Guam examples.

When comparing study designs, these interval approximations provide useful measures of the precision of heritability estimates under general pedigrees. Note however, that these are formulae for an approximation to expected confidence intervals. They should not be used in data analyses, where one can construct confidence intervals from the trait and pedigree data directly.

6 Discussion

We have developed several approximations to the ELRT for heritability analysis in pedigrees. These provide useful tools to describe the relatedness between individuals in a study sample with reference to the impact on statistical power and precision of estimates of narrow-sense heritability. The approximations use one or two simple functions of the eigenvalues of twice the kinship matrix, which we refer to as the relatedness summary parameters (RSPs), and work reasonably well even on very complex pedigrees. Further we describe how a target study may be described in terms of an effective sample size of a reference study (e.g., unrelated sibpairs). This effective sample size provides an accurate approximation to the power of the study, and can be easily interpreted as the number of unrelated sibpairs which would be required to have equivalent power to the target study. A summary of the different approximations and their properties is presented in Table 2.

Our general approach is an extension of [3] which determined the exact ELRT for testing for positive heritability in pedigrees, using the eigendecomposition of twice the kinship matrix. Others have also considered this null hypothesis, and derived approximations to related quantities such as the non-centrality parameter for simple and small pedigrees [15, 19]. These approaches in general are difficult to apply to complex pedigrees, and provide little intuition of how relatedness between individuals in a study can affect statistical power. Our approach considers a null hypothesis of arbitrary heritability. While the most relevant null hypothesis is often $h^2 = 0$, the more general null provides estimates of the expected precision of heritability estimates by inverting the critical region of the general test.

The first approximation, (7), in general does not provide accurate enough approximation to the exact ELRT, (5), for complex pedigrees. A similar approximation was recently developed independently by Visscher and Goddard [23] using an alternative method, but these methods are likely to perform poorly in complex pedigrees, where the eigenvalues are not symmetrically distributed. Approximation B however, provides a generally more accurate assessment of ELRT even in very complex pedigrees, but can perform poorly when the true heritability is small. Construction of expected confidence intervals from approximation B will depend on the true heritability. While this complicates the derivation of an analytic formula for these intervals, it corresponds with what we observe under simulation, where the difference between the 0.025 and 0.975 empirical quantiles of the heritability estimates, $\hat{h}^2$, varied as the true heritability changed. A viable alternative to approximation B is approximation C, which has advantages in terms of its simple form and can be more accurate than $AELRT_B$, particularly for small values of h^2 . Approximations B and C are also restricted to kinship matrices that are strictly positive-definite, which may preclude their use in some pedigrees. Finally, the simplification resulting from the eigendecomposition of twice the kinship matrix does not generally extend to the case with additional variance

components, such as a common environmental effect for individuals living in the same household. This likely limits the approaches of this paper to the case of additive genetic models. Further, we focused on situations with large sample sizes. In cases where sample sizes are smaller, the underlying asymptotic results and approximations have not been evaluated and some caution in using both is merited.

The most promising approach may be the effective sample method, which when using V_q provides remarkably accurate approximations to the ELRT, power and precision. The level of accuracy of this approximation will depend on the particular pedigree, and the situations and concerns of the particular application. When approximating power, the risk of overestimation seems small or negligible. Underestimation of power can occur when true heritability is small (<0.2) or large (>0.7). The effective sample size is on a scale which is readily interpretable, in much the same way we understand study sample size in studies involving independent units. Further, given the effective sample size and the reference study population's V_q^* , one can recover V_q for the study of interest, if desired. The variance of the log-eigenvalue, V_q in this context seems to be related to the efficiency of the target design on a per-subject basis, while the product of the sample size (N) and V_q is approximately proportional to the ELRT, and more directly related to the study's power. Each of these components may be of interest. The efficiency aspect is likely to be more of interest to someone considering what types of pedigrees to ascertain to minimize the actual study sample size or cost, whereas the NV_q would be more of interest to someone considering more rigorous or expensive follow-up on pedigrees already ascertained. As can be seen particularly in Figures 5 and 6, $AELRT_{ESS}$ does have some deficiencies in its ability to approximate both power and precision. These deficiencies are likely most relevant when h^2 is small, as other situations are rare (large h^2), or the error is negligible (intermediate levels of h^2). The lack of fit for small h^2 may be attributed to using such a simple reference design, sibpairs, where the resultant approximate ELRT, $AELRT_{ESS}$, lacks some of the features of the more complex pedigrees it is trying to approximate. Using alternative reference designs (e.g., trios or nuclear families) to capture some of these features may be a more accurate method and is an area of future work.

These approximations' approaches may be relevant in other types of analysis, such as association testing in related individuals, or using inferred relatedness to estimate heritability [14, 24, 25], where a similar type of dependence between individuals exists. Further applications may be in the area of meta-analysis, where the appropriate weighting of study heritability estimates may also be functions of the RSPs.

Recently, there has been increasing concern about the replication of research results, particularly in genetic research [16]. While this problem likely has many causes, one source may simply be lack of statistical power [5, 18]. In the context of heritability studies, it can be difficult to communicate how subjects are related, especially for large complex studies (Figure 1). The approximations in this paper provide means to effectively communicate the relatedness of a study population as it pertains to power, without the need to compromise confidentiality by providing full pedigree information. Reporting each of the actual eigenvalues is an impractical alternative, as most presentation formats (e.g., research articles) will be unable to accommodate a list of N numbers, unless N is very small or the

pedigree structure is trivial. For example, in the GAW13 pedigree, $N = 2860$ and there are 1040 unique eigenvalues. The RSPs or effective sample sizes on the other hand, can be reported concisely alongside other characteristics of a study population, such as average age or gender distribution in order to summarize the relatedness within the sample (Figure 1d). We hope they will allow for increased understanding of the role relatedness plays in statistical power and precision in heritability analysis on pedigrees.

Acknowledgments

We would like to thank the Genetic Analysis Workshop (GAW) 13 organizers for allowing us to use simulated pedigree data in this paper. We also thank Dr. Ellen Wijsman for permitting us to use the Guam pedigree structure, as well as the two anonymous referees who provided insightful comments to our paper. This research was supported in part by National Institutes of Health grants GM046255 and GM099568 and a Pacific Institute for the Mathematical Sciences postdoctoral fellowship.

References

1. Ahn, C.; Moonseoung, H.; Zhang, S. Sample Size Calculations for Clustered and Longitudinal Outcomes in Clinical Research. Vol. 1. CRC Press; 2014.
2. Almasy L, Amos CI, Bailey-Wilson JE, Cantor RM, Jaquish CE, Martinez M, Neuman RJ, Olson JM, Palmer LJ, Rich SS, et al. Genetic analysis workshop 13: analysis of longitudinal family data for complex diseases and related risk factors. *BMC Genetics*. 2003; 4(Suppl 1):S1. [PubMed: 14975069]
3. Blangero J, Diego VP, Dyer TD, Almeida M, Peralta J, Kent JW Jr, Williams JT, Almasy L, Göring H, et al. A kernel of truth: statistical advances in polygenic variance component models for complex human pedigrees. *Advances in Genetics*. 2012; 81:1–31. [PubMed: 23419715]
4. Brown BW, Lovato J, Russell K. Asymptotic power calculations: description, examples, computer code. *Statistics in Medicine*. 1999; 18(22):3137–3151. [PubMed: 10544312]
5. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, McCarthy MI, Ouwehand WH, Samani NJ, et al. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007; 447(7145):661–678. [PubMed: 17554300]
6. Casella, G.; Berger, RL. *Statistical Inference*. 2. Duxbury; Pacific Grove, CA: 2002.
7. Daniels MJ, Kass RE. Shrinkage estimators for covariance matrices. *Biometrics*. 2001; 57(4):1173–1184. [PubMed: 11764258]
8. Daw EW, Morrison J, Zhou X, Thomas DC. Genetic analysis workshop 13: simulated longitudinal data on families for a system of oligogenic traits. *BMC Genetics*. 2003; 4(Suppl 1):S3. [PubMed: 14975071]
9. Edwards JH. The interpretation of lod scores in linkage analysis. *Human Gene Mapping*. 1976; 3:289–293.
10. Fisher RA. XV. - The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*. 1918; 52(2):399–433.
11. Harville, DA. *Matrix algebra from a statistician's perspective*. Springer; New York, NY: 2008.
12. Henderson CR. Simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics*. 1976; 32(1):69–83.
13. Lippert C, Listgarten J, Liu Y, Kadie CM, Davidson RI, Heckerman D. Fast linear mixed models for genome-wide association studies. *Nature Methods*. 2011; 8(10):833–835. [PubMed: 21892150]
14. Makowsky R, Pajewski NM, Klimentidis YC, Vazquez AI, Duarte CW, Allison DB, de Los Campos G. Beyond missing heritability: prediction of complex traits. *PLoS Genetics*. 2011; 7(4):e1002051.
15. Martin N, Eaves L, Kearsey M, Davies P. The power of the classical twin study. *Heredity*. 1978; 40(1):97–116. [PubMed: 272366]

16. Moonesinghe R, Khoury MJ, Janssens ACJW. Most published research findings are false but a little replication goes a long way. *PLoS Medicine*. 2007; 4(2):e28. [PubMed: 17326704]
17. Ott, J. Analysis of human genetic linkage. JHU Press; Baltimore, MD: 1999.
18. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Meitinger T, Braund P, Wichmann HE, et al. Genomewide association analysis of coronary artery disease. *New England Journal of Medicine*. 2007; 357(5):443–453. [PubMed: 17634449]
19. Schork N, Schork M. The relative efficiency and power of small-pedigree studies of the heritability of a quantitative trait. *Human Heredity*. 1993; 43(1):1–11. [PubMed: 8514319]
20. Sieh W, Choi Y, Chapman NH, Craig UK, Steinbart EJ, Rothstein JH, Oyanagi K, Garruto RM, Bird TD, Galasko DR, et al. Identification of novel susceptibility loci for Guam neurodegenerative disease: challenges of genome scans in genetic isolates. *Human Molecular Genetics*. 2009; 18(19): 3725–3738. [PubMed: 19567404]
21. Thompson EA, Shaw RG. Pedigree analysis for quantitative traits: variance components without matrix inversion. *Biometrics*. 1990; 46:399–413. [PubMed: 2364130]
22. Visscher PM. A note on the asymptotic distribution of likelihood ratio tests to test variance components. *Twin Research and Human Genetics*. 2006; 9(04):490–495. [PubMed: 16899155]
23. Visscher PM, Goddard ME. A general unified framework to assess the sampling variance of heritability estimates using pedigree or marker-based relationships. *Genetics*. 2015; 199(1):223–232. [PubMed: 25361897]
24. Visscher PM, Medland SE, Ferreira MA, Morley KI, Zhu G, Cornes BK, Montgomery GW, Martin NG. Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLoS Genetics*. 2006; 2(3):e41. [PubMed: 16565746]
25. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*. 2011; 88(1):76–82. [PubMed: 21167468]

Fig. 2.
Histograms of eigenvalues and log eigenvalues for each of the example study designs

Fig. 3.

ELRT as a function of null-heritability, h_0^2 for the three studies in Table 1. Circles indicate mean (simulated) likelihood ratio test at a specific null heritability value, orange line is the exact ELRT according to equation (5) and the red, blue, green and magenta lines are approximations A, B, C and ESS respectively. True heritabilities (h^2) run from left to right 0.3, 0.5, 0.7. The lower dotted horizontal line is at $\chi^2_{1,0,(1-\alpha)=0.95}$ used to construct expected CIs in Section 5.3

Fig. 4.

Power as a function of null-heritability, h_0^2 at $h^2 = 0.3$ or 0.7 for the three studies in Table 1 at a $\alpha = 0.05$ significance level. Circles indicate mean (simulated) power from a likelihood ratio test at the specific null heritability value, orange line is the exact theoretical power using the exact ELRT according to equation (5) and the red, blue and green lines are approximations A, B, and C respectively. Magenta line represents power based on the $ELRT_{ESS}$ approximation.

Fig. 5. Maximum and Minimum Deviations of power based on $ELRT_{ESS}$ from the Exact Theoretical Power based on the exact ELRT, (5) for the GAW13 (left) and Triple Guim (right) Examples considering h^2 and h_0^2 values [0.01, 0.8].

Fig. 6.

95% Confidence Interval (CI) Width as a function of true heritability. Black solid circles represent the distance between the 0.025 and 0.975 quantiles of $\hat{h}^2$ under simulation. Black line represents the mean width of a 95% Wald Interval for h^2 , under simulation. Orange line represents the 95% confidence interval from the exact ELRT. Red, blue, green and magenta lines represent the approximated 95% confidence interval width using approximations, A, B, C and ESS, respectively. The dotted pink line is the maximum 95% confidence interval width under the ESS approximation.

Table 1

Relatedness Summary Parameters (RSPs) for Several Example Pedigrees

Pedigree	N	V_{λ}	γ	V_q
1500 Unrelated Sibpairs	3000	0.25	0.87	0.30
GAW13 simulated pedigree	2860	1.12	0.66	0.76
Triple Guam	960	2.62	0.57	0.95

Table 2

Summary of Approximations

	AELRT_A	AELRT_B	AELRT_C	AELRT_{ESS}
Based On	Eigenvalues	Log-Eigenvalues	Log-EigenValues	Effective Sample size
Expected Confidence Interval Formula	$h^2 \pm z_{\alpha/2} \sqrt{\frac{2}{V_{\lambda}(N-1)}}$	Requires a numerical solver	$h^2 \pm z_{\alpha/2} \sqrt{\frac{2}{V_q(N-1)}}$	Requires a numerical solver
General Properties: Power	Overestimate for complex pedigrees	Can overestimate and underestimate	Can overestimate and underestimate	Generally accurate for intermediate levels of heritability. See Figure 5.
General Properties: Confidence Interval Width	Generally too narrow for complex pedigrees	Very wide for small h^2 , a little too narrow for large h^2	Too narrow for small h^2 ; Too Wide for large h^2	Generally Accurate. See Figure 6.