Published in final edited form as:

Genet Epidemiol. 2012 September ; 36(6): 538–548. doi:10.1002/gepi.21645.

Power and Sample Size Calculations for SNP Association Studies with Censored Time-to-Event Outcomes

Kouros Owzar^{1,*}, Zhiguo Li¹, Nancy Cox², and Sin-Ho Jung¹

¹Department of Biostatistics and Bioinformatics, Duke University

²Section of Genetic Medicine, Department of Medicine and Department of Human Genetics, University of Chicago

Abstract

For many clinical studies in cancer, germline DNA is prospectively collected for the purpose of discovering or validating Single Nucleotide Polymorphisms associated with clinical outcomes. The primary clinical endpoint for many of these studies are time-to-event outcomes such as time of death or disease progression which are subject to censoring mechanisms. The Cox score test can be readily employed to test the association between a SNP and the outcome of interest. In addition to the effect and sample size, and censoring distribution, the power of the test will depend on the underlying genetic risk model and the distribution of the risk allele. We propose a rigorous account for power and sample size calculations under a variety of genetic risk models without resorting to the commonly used contiguous alternative assumption. Practical advice along with an open-source software package to design SNP association studies with survival outcomes are provided.

Keywords

censoring pharmacogenomics; Cox score test; genetic risk; SNP association study

1 Introduction

Single Nucleotide Polymorphisms (SNP) case-control studies have historically been carried out in the context of binary phenotypes. Power and sample-size calculation methods and software for designing case-control studies have been published extensively (e.g., Michael and Otta (2002); Lange et al. (2004); Hao et al. (2004); De La Vega et al. (2005); Edwards et al. (2005); Skol et al. (2006); Klein (2007); Menashe et al. (2008); Spencer et al. (2009)). For many clinical studies, the primary phenotype of interest is a censored time-to-event outcome, such as time to death, disease progression or drug induced toxicity. The methods and software developed for case-control studies are not applicable in this setting. For illustration, we consider three genome-wide association studies (GWAS) carried out by the Cancer and Leukemia Group B (CALGB). CALGB 80303 (Kindler et al. 2010; Innocenti et al. 2012) is a placebo-controlled randomized phase III study in advanced pancreatic cancer randomizing patients to gemcitabine with or without bevacizumab. A cancer risk phenotype of interest is time of death post randomization. CALGB 90401 (Kelly et al. 2010) is a placebo-controlled phase III study in metastatic prostate cancer. A cancer risk phenotype of interest is time of development of liver metastasis. Finally, CALGB 40101 (Shulman et

*Corresponding Author: Kouros Owzar, Department of Biostatistics and Bioinformatics, Duke University, 2424 Erwin Road, Suite 1102; Room 11074, Durham, NC 27710, 919-681-1829, kouros.owzarduke.edu.

The authors declare no conflicts of interest

al. 2010) is a randomized phase study in early-stage breast cancer randomizing patients to cyclophosphamide or paclitaxel. For the latter arm, the pharmacogenomic phenotype of interest is the time of onset of peripheral neuropathy.

As the event of interest may not have been realized at the time of the analysis, it is subject to a random censoring mechanism. For example, the time of death for a CALGB 80303 patient alive at the time of analysis is unobserved. The only information available is the amount of time the patient has lived since randomization based on the data from the last follow-up. A CALGB 40101 patient who has not experienced a neuropathy episode may have not been treated long enough with the study drug. The time of onset of toxicity for this patient will remain unobservable.

As with any experiment, the conduct of scientifically sound research in this setting necessitates access to rigorous and practical design methodology and tools. In this paper, we will propose rigorous power and sample size calculation formulas on the basis of using the score statistic in a proportional hazards model to test the hypothesis of SNP by outcome association under specific genetic risk models. We derive the power function under quite general settings, only requiring the independent censoring assumption. We begin by discussing specific genetic risk models in the context of survival phenotypes and proceed by providing an outline of our approach to calculating power and sample size in this setting. The technical details are relegated to an appendix. Afterwards, we will outline an extensive simulation investigation to empirically assess our proposed formulas and conclude the paper with a discussion followed by a brief summary.

2 Genetic Disease Model

In these discussions, we assume that the SNP is bi-allelic and denote the genotype outcomes as $\{AA, AB, BB\}$, where B is considered to be the risk allele in the sense that its presence is associated with increased hazard of experiencing the event. Corresponding to these outcomes, we define $G \in \{0, 1, 2\}$ to denote the number of copies of the risk allele B .

Denoting the actual event time by $\tilde{T}$. Due to censoring, what is observed is the random pair (T, δ) where $T = \min\{\tilde{T}, C\}$ and $\delta = \mathbb{I}[\tilde{T} < C]$, where $\mathbb{I}$ is the indicator function defined as $\mathbb{I}[a < b] = 1$ if $a < b$ or 0 otherwise. Here C is the censoring time and δ is the event indicator. The latter realizes the value 1 if the event of interest occurs before time C or 0 otherwise. Let $\bar{F}(t) = \mathbb{P}(\tilde{T} > t)$ denote the survival distribution in the population which may be expressed in terms of the survival functions conditional on genotypes, $\bar{F}_g(t) = \mathbb{P}(\tilde{T} > t | G =$

$g)$, and the genotypic relative frequencies $\pi_g = \mathbb{P}(G = g)$ as the mixture $\bar{F}(t) = \sum_{g=0}^2 \pi_g \bar{F}_g(t)$ for all $t \geq 0$. It is noted these conditional survival functions are the counterparts to penetrances, the conditional probabilities of being affected given the number of copies of the risk allele, in the case-control setting. Two important functions related to a survival function $\bar{F}(t)$ are the hazard function, defined as $\lambda(t) = f(t)/\bar{F}(t)$, where $f(t)$ is the density function of $\tilde{T}$, and the corresponding cumulative hazard function defined as $\Lambda(t) = \int_0^t \lambda(s) ds$. We can express the survival distribution in the population in terms of the conditional cumulative

hazard functions, as $\bar{F}(t) = \sum_{g=0}^2 \pi_g \exp\{-\Lambda_g(t)\}$ for all $t \geq 0$, where Λ_g is the cumulative hazard function corresponding to the conditional survival function $\bar{F}_g$.

In the case-control setting, the null hypothesis is expressed in terms of the equality of the penetrances. In the survival setting, the null hypothesis of no association can be expressed in terms of the conditional cumulative hazard functions as $H_0 : \Lambda_0(t) = \Lambda_1(t) = \Lambda_2(t)$ for all $t \geq 0$. For most genetic risk models relevant in cancer, the risk is a non-decreasing function of

the risk allele. Therefore, it is prudent to power the analyses for ordered alternative hypotheses of the form $\Lambda_0(t) < \Lambda_1(t) < \Lambda_2(t)$ (or $\Lambda_0(t) > \Lambda_1(t) > \Lambda_2(t)$) for some $t > 0$.

In the context of SNP association studies, the power of a test is typically investigated under specific genetic risk models. The canonical effect size in genomic case-control studies is the Genotype Relative-Risk (GRR) defined as the ratio of two penetrances. We will denote the corresponding effect size in the survival setting as the Genotype Hazard-Ratio (GHR), to be denoted by $\Delta(t)$, defined as the ratio of two conditional hazard rates evaluated at time $t > 0$. We define the *recessive* model as $\lambda_1(t) = \lambda_0(t)$ and $\lambda_2(t) = \Delta(t)\lambda_0(t)$ and the *dominant* model as $\lambda_1(t) = \lambda_2(t) = \Delta(t)\lambda_0(t)$. Finally, we define the *additive* model as $\lambda_1(t) = \Delta(t)\lambda_0(t)$ and $\lambda_2(t) = \Delta(t)^2\lambda_0(t)$. More generally, we can generate these genetic risk models using a weight function ω mapping the genotype G into $\omega(G)$ and hence $\lambda_g(t) = \Delta(t)^{\omega(G)}\lambda_0(t)$. The weights $(\omega(0), \omega(1), \omega(2))$ for the recessive, dominant and additive models are $(0,0,1)$, $(0,1,1)$ and $(0,1,2)$ respectively.

The remainder of the discussion will concentrate on the time-independent risk ratio model, i.e., the proportional hazards model. In other words, we will assume that $\Delta(t) = \Delta$ for all $t > 0$. We will also make the assumption of Hardy-Weinberg Equilibrium (HWE). Among other things under this assumption the genotypic relative frequencies are $(1-q)^2$, $2q(1-q)$ and q^2 respectively, where q denotes the relative frequency of the risk allele B . Under the assumption of HWE and the three aforementioned genetics risk models, we can now express the survival distribution of the population as the following mixture as

$$\mathbb{P}[\tilde{T} > t] = (1-q)^2 e^{-\Lambda_0(t)} + 2q(1-q) e^{-\Delta^{\omega(1)}\Lambda_0(t)} + q^2 e^{-\Delta^{\omega(2)}\Lambda_0(t)}.$$

3 Power and Sample Size Formula Derivation

In this section, we will outline our approach for deriving the power and sample size formulas while relegating a complete and rigorous technical account of the details to the Appendix. The sampling distribution of the score test statistic under appropriate regularity conditions is asymptotically normal (Fleming and Harrington 1991). Under the null hypothesis, the asymptotic sampling distribution is standard normal. To calculate the power and sample size, we will have to derive the asymptotic mean of the test statistic under the alternative hypothesis and a given genetic risk model. Complete and accessible accounts on the counting process and empirical process framework used in our approach are provided in Fleming and Harrington (1991) and van der Vaart and Wellner (1996), respectively.

Let $Y_i(t) = \mathbb{I}[T_i \geq t]$ and $N_i(t) = \delta_i \mathbb{I}[T_i \leq t]$ for $i \in \{1, \dots, n\}$. These are usually referred to as the “at risk” and the event processes respectively. Under the proportional hazards assumption that $\Delta(t) \equiv \Delta$, the score statistic for testing for $H_0: \theta = 0$, where $\theta = \log \Delta$, is

$$U_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^\infty \left\{ \omega(G_i) - \frac{s_1(t)}{s_0(t)} \right\} dN_i(t),$$

where $s_k(t) = \sum_{j=1}^n \{\omega(G_j)\}^k Y_j(t) / n$ for $k \in \{0, 1, 2\}$. The test statistic of the score test is defined as $W_n = U_n / \hat{\sigma}_0$, where

$$\widehat{\sigma}_0^2 = \frac{1}{n} \sum_{i=1}^n \int_0^\infty \left[\frac{s_2(t)}{s_1(t)} - \left\{ \frac{s_1(t)}{s_0(t)} \right\}^2 \right] dN_i(t).$$

Denote $S_{\theta^*}^{(k)}(\theta, t) = \mathbb{E}_{\theta^*} [Y(t) \{\omega(G)\}^k e^{\theta \omega(G)}]$ for $k = 0, 1, 2$, where the expectation is taken based on the true effect size θ^* . To derive the asymptotic properties of the test statistic under alternative genetic risk models, we obtain an asymptotic decomposition of the score statistic in the form

$$U_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n [\xi_i(\theta) + \eta_i(\theta) - \mathbb{E}\{\xi_i(\theta) + \eta_i(\theta)\}] + \sqrt{n}\mu(\theta) + o_p(1),$$

as $n \rightarrow \infty$, where $\xi_i(\theta)$ and $\eta_i(\theta)$ are functions of the observed data for subject i (see the Appendix for the exact expressions), and

$$\mu(\theta) = \int_0^\infty \left\{ S_{\theta}^{(1)}(\theta, t) - \frac{S_{\theta}^{(1)}(0, t)}{S_{\theta}^{(0)}(0, t)} S_{\theta}^{(0)}(\theta, t) \right\} d\Lambda_0(t).$$

It follows that

$$U_n - \sqrt{n}\mu(\theta) \rightarrow N(0, \sigma^2(\theta))$$

in distribution, as $n \rightarrow \infty$, where

$$\begin{aligned} \sigma^2(\theta) &= \sigma_1^2(\theta) + \sigma_2^2(\theta) \\ &= \underbrace{\text{var}\{\xi(\theta)\}}_{\sigma_1^2(\theta)} + \underbrace{\text{var}\{\eta(\theta)\} + 2\text{cov}\{\xi(\theta), \eta(\theta)\}}_{\sigma_2^2(\theta)}. \end{aligned}$$

By definition of $\mu(\theta)$, we have $\mu(0) = 0$. Moreover, we show in the Appendix that $\widehat{\sigma}_0^2 \rightarrow \sigma_0^2(\theta)$ in probability as $n \rightarrow \infty$, where

$$\sigma_0^2(\theta) = \int_0^\infty \left[\frac{S_{\theta}^{(2)}(0, t)}{S_{\theta}^{(0)}(0, t)} - \left\{ \frac{S_{\theta}^{(1)}(0, t)}{S_{\theta}^{(0)}(0, t)} \right\}^2 \right] S_{\theta}^{(0)}(\theta, t) d\Lambda_0(t).$$

It can also be seen in the Appendix that $\eta_i(0) = 0$ and $\sigma^2(0) = \text{var}\{\xi(\theta)\} = \sigma_0^2(0)$. Hence $W_n \rightarrow N(0, 1)$ in distribution as $n \rightarrow \infty$ under H_0 . For a two-sided test with a significance level α , we reject H_0 when $|W_n| > z_{1-\alpha/2}$. Given sample size n , the power of the test is given by

$$P(|W_n| > z_{1-\alpha/2}) = P\left\{\frac{\sigma_0(\theta)W_n - \sqrt{n}\mu(\theta)}{\sigma(\theta)} > \frac{\sigma_0(\theta)z_{1-\alpha/2} - \sqrt{n}\mu(\theta)}{\sigma(\theta)}\right\} + P\left\{\frac{\sigma_0(\theta)W_n - \sqrt{n}\mu(\theta)}{\sigma(\theta)} < \frac{-\sigma_0(\theta)z_{1-\alpha/2} - \sqrt{n}\mu(\theta)}{\sigma(\theta)}\right\} \\ \approx 1 - \Phi\left\{\frac{\sigma_0(\theta)z_{1-\alpha/2} - \sqrt{n}\mu(\theta)}{\sigma(\theta)}\right\} + \Phi\left\{\frac{-\sigma_0(\theta)z_{1-\alpha/2} - \sqrt{n}\mu(\theta)}{\sigma(\theta)}\right\},$$

where Φ is the distribution function of $N(0, 1)$. Setting this equal to $1 - \beta$, the sample size needed for a $1 - \beta$ power is

$$n \approx \left\{ \frac{\sigma_0(\theta)Z_{1-\alpha/2} + \sigma(\theta)Z_{1-\beta}}{\mu(\theta)} \right\}^2.$$

To calculate the sample size, one has to calculate the asymptotic mean $\mu(\theta)$, the limit $\sigma_0^2(\theta)$, and the asymptotic variance $\sigma^2(\theta)$. The functions $\mu(\theta)$, $\sigma_0^2(\theta)$ and $\sigma_1^2(\theta)$ are univariate integrals or can be directly expressed in terms of univariate integrals under the independent censoring assumption. Consequently, these quantities can be calculated using numerical integration methods for univariate integrals.

It should be noted that our formula applies to a recessive, dominant or additive model, by choosing the corresponding function $\omega(G)$. We will empirically show that approximating the variance $\sigma^2(\theta)$ by $\sigma_1^2(\theta)$ generally yields an accurate approximation to the power.

4 Simulation Study

In this section, we summarize the results from a comprehensive simulation study. We will assume that $\tilde{F}(t)$ is a mixture of exponential laws and that the censoring distribution is uniform on the interval $(0, \tau_\rho)$. Under one of the three previously listed genetic risk models, the survival probability in the population is expressible as

$$\mathbb{P}[\tilde{T} > t] = (1-q)^2 e^{-\lambda_0 \Delta^{\omega(0)} t} + 2q(1-q) e^{-\lambda_0 \Delta^{\omega(1)} t} + q^2 e^{-\lambda_0 \Delta^{\omega(2)} t}.$$

For a fixed choice of t , $\mathbb{P}[\tilde{T} > t]$ and q , we obtain the baseline risk rate λ_0 by numerically solving the last equation. The shapes $\mathbb{P}[\tilde{T} > t]$ and the three conditional survival functions $S_g(t) = \mathbb{P}[\tilde{T} > t | G = g]$ are illustrated in Figure 1. The parameter τ_ρ in the censoring distribution is chosen so as the $\rho = \mathbb{P}[\tilde{T} < C]$ for a given event rate $\rho \in (0, 1)$. It is trivial to show that

$$\tau_\rho = \left\{ (1-q)^2 + 2q(1-q)e^{-\omega(1)} + q^2 e^{-\omega(2)} \right\} \frac{1 - e^{-\lambda_0 t}}{\rho \lambda}.$$

The data are simulated as follows:

1. Draw (N_0, N_1, N_2) from a multinomial distribution with parameter $(n; (1-q)^2, 2q(1-q), q^2)$.

2. Conditional on these genotype counts, independently draw $\tilde{T}_1, \dots, \tilde{T}_{N_0}$ from $\text{Exp}(\lambda_0)$, $\tilde{T}_{N_0+1}, \dots, \tilde{T}_{N_0+N_1}$ from $\text{Exp}(\lambda_0 \Delta^{\omega(1)})$ and $\tilde{T}_{N_0+N_1+1}, \dots, \tilde{T}_n$ from $\text{Exp}(\lambda_0 \Delta^{\omega(2)})$, where $\text{Exp}(\lambda)$ denotes an exponential law with parameter λ .
3. Draw $C_1, \dots, C_n$ independently from a uniform distribution on the interval $(0, \tau_\rho)$.

For the simulation studies, we will assume $P(\tilde{T} > 1) = 0.5$. In other words, we will assume that the median in the population is one unit of time. Unless indicated otherwise, a nominal two-sided level of $\alpha = 0.01$ is used. For empirical power calculations, $B = 10,000$ simulation replicates are used for each example. The power illustrations considered here are limited to additive risk models.

We begin by studying the type I error control of the score test within this framework for $q \in [0.01, 0.99]$. We consider sample sizes $n \in \{100, 500, 1000\}$ and event rates $\rho \in \{0.5, 0.7, 0.9\}$. The results, illustrated in Figure 1, show that the type I error rate is approximately controlled when the relative allele frequency is not too close to 0 or 1. Next, we investigate the power of the test by calculating the power for detecting a GHR Δ for a given risk allele relative frequency. For each example, the power based on the exact and approximate variance formulas is shown. The empirical power is provided as a reference. We consider event rates $\rho \in \{0.7, 0.9\}$ and relative risk allele frequencies $q \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$. The results, illustrated in Figures 2, 3 and 4 for sample sizes $n = 100, 250$ and 500 respectively, show that the power calculated from our formula is generally very close to the empirical power.

5 Discussion

In this paper, we have presented power and sample size calculations for SNP studies with a time-to-event outcome subject to a random right censoring mechanism. The inference is based on the score test in a proportional hazards model. Our methods do not rely on the contiguous alternative assumption which is made in most sample size formulas in survival analysis (see for example Schoenfeld (1983) and Hsieh and Lavori (2000)). Under the contiguous alternative it is assumed that the effect size (log of GHR in our case) converges to 0 at a rate of $1/\sqrt{n}$. This particular rate is chosen due to theoretical reasons. A sample size and power calculation method developed under the contiguous alternative setting may only be applicable where the effect size is small and therefore may not be suitable for designing GWAS studies.

An obvious concern would be the lack of type I error control in the sparse setting induced by low relative frequency of the minor allele coupled with a potentially low event rate. This is illustrated empirically in Figure 1 by considering control for a nominal two-sided level of 0.01. These results suggest that one should exercise care using asymptotics for relative frequencies outside of the $[0.1, 0.9]$ range. It should be noted that the P -value is uniformly distributed under the null hypothesis. Therefore, there is no loss of generality by investigating the marginal type I error at the 0.01 rather than say the 10^{-8} nominal level. Before conducting any power or sample size calculation, it is prudent to assess the type I error control empirically using a similar simulation exercise. Permutation resampling could be employed to approximate the exact sampling distribution of the score test when the relative allele frequency is small.

The power, at the two-sided level of 0.01, is illustrated for small ($n = 100$), medium ($n = 250$) and large ($n = 500$) sample sizes in figures 2, 3 and 4 respectively. Based on the empirical type I error simulation, the relative frequency of the risk allele is restricted to the interval $[0, 1, 0.9]$. Using the empirical power curve as the reference, the observed asymptotic power based on the exact variance formula is very accurate. The asymptotic

variance formula yields very accurate approximation to the power when the SNP is common $q \in [0, 3, 0.7]$ and the effect size is not too large.

For the sake of illustration, the distributions of $\tilde{T}$ and C have assumed to be a mixture of exponentials and uniform respectively in these discussions. The proposed power and sample size formulas have been derived for arbitrary marginal distributions within the proportional hazards setting.

In the simulation model considered in this paper, we assumed a putative probability $P(\tilde{T} > \tau_\ell)$ for a landmark τ_ℓ . As with any mixture model, we should be cognizant of issue of non-identifiability. The model is identifiable conditional on specifying the three individual components of the mixture. To design these studies, we suggest that τ_ℓ is chosen as a clinically relevant landmark. The median survival for metastatic pancreatic cancer patients is about six months. To design a validation study for CALGB 80303, we can set $\tau_\ell = 6$ months. Then given $P(\tilde{T} > 6) = 0.5$, for a given relative allele frequency, we can obtain the exponential baseline hazard rate λ_0 from the mixture model representation.

We have considered power and sample size calculations in the case of a single SNP. The proposed methodology and tools can be employed to conduct these calculations to design GWAS studies. Specifically, the study can be designed by using a marginal error rate of $K^{-1}\alpha$ where K denotes the number of SNPs to be tested. This approach will provide genome-wide type I error control at the family-wise error rate (FWER) of α . For example, suppose that one is tasked with designing a GWAS study based on a target sample size of $n = 1,000$ patients with usable DNA and time-to-event data. We assume that $K = 1,000,000$ SNPs are to be tested in a variant by variant analysis. The analyses are to be powered for an additive effect model. The power, at the two-sided 0.05 FWER level, is illustrated in Figure 5. In this analysis each variant is tested at the nominal 5×10^{-8} level. These power calculations may be conservative as they do not account for linkage disequilibrium among SNPs typed on the platform. The work to extend the proposed methodology to account for multiple testing within the false-discovery rate (FDR) framework using a similar approach proposed by Jung (2005) for binary outcomes is under way.

We have developed the power and sample size calculations based on the score test. The likelihood ratio and the Wald test can also be employed for the purpose of inference. It should be noted that all three tests are asymptotically equivalent, although their finite sampling distributions may differ. Unlike the likelihood ratio and Wald test, calculating the score statistic does not require numerical optimization. This is an important consideration, especially in the case of GWAS where it may not be practical to monitor the numerical stability of the optimization for every single SNP.

In the case-control setting, the genotype test is used for testing the omnibus hypothesis of any difference among the penetrances. In the survival setting, the corresponding hypothesis of any difference among the three conditional survival functions can be tested using the three-sample log-rank test. The power and sample size formulas proposed by Jung and Hui (2002) can be employed for this case.

While the genotypic relative frequencies are calculated under HWE, the score test yields valid inference, in sense of controlling type I error, in case of deviation from HWE. The major assumption is that the inference is carried out under the framework of mutually independent experimental units where the relative genotypic frequencies, p_0 , p_1 and p_2 are homogenous.

It is common to design GWAS studies powered for additive alternatives. The true genetic risk model for a prognostic SNP may be recessive or dominant. Our proposed formula can

also be used to calculate the power when the hypothesized genetic model differs from the true risk model. This is trivially accomplished by using the true risk scores ($\omega(0)$, $\omega(1)$, $\omega(2)$) when calculating the moments.

The specific form of censoring assumed here is referred to as a non-informative right-censoring mechanism. The methodology can be extended to other censoring mechanisms including left or interval censoring mechanism. Left censoring arises in the cases where the phenotype may not be quantifiable at low levels below background. Interval censoring arises in cases where the time of the event is not directly observable but is known to have occurred during an interval. For example, when following patients for disease progression, the event of interest is thought to have occurred some between the time of the last visit, when the patient's disease was assessed to have been under control, and the time of the current visit, where the patient's disease was assessed to have progressed.

To facilitate and enable the proposed power and sample calculations, we provide an open-source extension package, *survSNP*, for the R statistical environment (R Development Core Team 2011). This package is extensible and its functions have been used to conduct the analyses presented in this paper. Point releases are available through The Comprehensive R Archive Network at <http://cran.r-project.org/web/packages/survSNP/while> the development version is available from <https://bitbucket.org/kowzar/survsnp/>. The package was developed using the Rcpp framework (Eddelbuettel and Francois 2011) and depends on the GNU Scientific Library (Galassi et al. 2009) for numerical integration. The package depends on the *survival* (Therneau 2011) package to calculate the score statistic for the simulation studies. The packages *foreach* (Revolution Analytics 2012), *lattice* (Sarkar 2008), and *xtable* (Dahl 2012) are used to facilitate the production and illustration of the results. These, as illustrated in this paper, can be further enhanced by using the *latticeExtra* (Sarkar and Andrews 2011), *RColorBrewer* (Neuwirth 2007) or *tikzDevice* (Sharpsteen and Bracken 2011) packages.

6 Summary

We have presented rigorous methodology for power and sample size calculations for SNP association studies with censored time-to-event outcomes. Along with practical suggestions for designing prospective experiments and validation studies, we provide an open-source and extensible software package to facilitate the requisite calculations.

Acknowledgments

The authors thank two reviewers for helpful comments. The authors also thank Mr. Chanhee Yi for his programming support. This research was supported in part by Award Number P01CA142538 (KO, ZL, SJ) and Award Number CA33601 (KO, ZL, SJ) from the National Cancer Institute, and the PAAR-Pharmacogenomics of Anticancer Agents Research Group, Award Number U01GM061393 (KO, NC) from the National Institute Of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Dahl, DB. R package version 1.7-0. 2012. *xtable*: Export tables to L^AT_EX or HTML.
- De La Vega F, Gordon D, Su X, Scafe C, Isaac H, Gilbert D, Spier E. Power and sample size calculations for genetic case/control studies using gene-centric snp maps: application to human chromosomes 6, 21, and 22 in three populations. *Human Heredity*. 2005; 60(1):43–60. [PubMed: 16137993]
- Eddelbuettel D, Francois R. Rcpp: Seamless r and c++ integration. *Journal of Statistical Software*. 2011; 40(8):1–18.

- Edwards B, Haynes C, Levenstien M, Finch S, Gordon D. Power and sample size calculations in the presence of phenotype errors for case/control genetic association studies. *BMC Genetics*. 2005; 6(1):18. [PubMed: 15819990]
- Fleming, T.; Harrington, D. Counting processes and survival analysis. Vol. 8. Wiley; New York: 1991.
- Galassi, M.; Davies, J.; Theiler, J.; Gough, B.; Jungman, P.; Alken, P.; Booth, M.; Rossi, F. GNU Scientific Library Reference Manual. 3. 2009.
- Hao K, Xu X, Laird N, Wang X, Xu X. Power estimation of multiple snp association test of case-control study and application. *Genetic Epidemiology*. 2004; 26(1):22–30. [PubMed: 14691954]
- Hsieh F, Lavori P. Sample-size calculations for the Cox proportional hazards regression model with nonbinary covariates. *Controlled Clinical Trials*. 2000; 21:552–560. [PubMed: 11146149]
- Innocenti F, Owzar K, Cox NL, Evans P, Kubo M, Zembutsu H, Jiang C, Hollis D, Mushiroda T, Li L, Friedman P, Wang L, Glubb D, Hurwitz H, Giacomini KM, McLeod HL, Goldberg RM, Schilsky RL, Kindler HL, Nakamura Y, Ratain MJ. A genome-wide association study of overall survival in pancreatic cancer patients treated with gemcitabine in calgb 80303. *Clin Cancer Res*. 2012; 18(2): 577–584.10.1158/1078-0432.CCR-11-1387 [PubMed: 22142827]
- Jung S. Sample size for FDR-control in microarray data analysis. *Bioinformatics*. 2005; 21(14):3097. [PubMed: 15845654]
- Jung S, Hui S. Sample size calculation for rank tests comparing k survival distributions. *Lifetime Data Analysis*. 2002; 8(4):361–373. [PubMed: 12471945]
- Kelly W, Halabi S, Carducci M, George D, Mahoney J, Stadler W, Morris M, Kantoff P, Monk J III, Small E, et al. A randomized, double-blind, placebo-controlled phase iii trial comparing docetaxel, prednisone, and placebo with docetaxel, prednisone, and bevacizumab in men with metastatic castration-resistant prostate cancer (mCRPC): survival results of CALGB 90401. *J Clin Oncol*. 2010; 28(suppl 18):951S.
- Kindler H, Niedzwiecki D, Hollis D, Sutherland S, Schrag D, Hurwitz H, Innocenti F, Mulcahy M, O'Reilly E, Wozniak T, Picus J, Bhargava P, Mayer R, Schilsky R, Goldberg R. Gemcitabine plus bevacizumab compared with gemcitabine plus placebo in patients with advanced pancreatic cancer: phase III trial of the Cancer and Leukemia Group B (CALGB 80303). *J Clin Oncol*. 2010; 28(22):3617. [PubMed: 20606091]
- Klein R. Power analysis for genome-wide association studies. *BMC Genetics*. 2007; 8(1):58. [PubMed: 17725844]
- Lange C, DeMeo D, Silverman E, Weiss S, Laird N. Pbat: tools for family-based association studies. *American Journal of Human Genetics*. 2004; 74(2):367. [PubMed: 14740322]
- Menashe I, Rosenberg P, Chen B. Pga: power calculator for case-control genetic association analyses. *BMC Genetics*. 2008; 9(1):36. [PubMed: 18477402]
- Michael D, Otta N. Power and sample size calculations for case-control genetic association tests when errors are present: application to single nucleotide polymorphisms. *Hum Hered*. 2002; 54:22–33. [PubMed: 12446984]
- Neuwirth, E. R package version 1.0-2. 2007. RColorBrewer: ColorBrewer palettes.
- R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; Vienna, Austria: 2011.
- Revolution Analytics. R package version 1.3.5. 2012. foreach: Foreach looping construct for R.
- Sarkar, D. Lattice: Multivariate Data Visualization with R. Springer; New York: 2008.
- Sarkar, D.; Andrews, F. R package version 0.6-16. 2011. latticeExtra: Extra Graphical Utilities Based on Lattice.
- Schoenfeld D. Sample-size formula for the proportional-hazards regression model. *Biometrics*. 1983; 39(2):499–503. [PubMed: 6354290]
- Sharpsteen, C.; Bracken, C. R package version 0.6.1. 2011. tikzDevice: A Device for R Graphics Output in PGF/TikZ Format.
- Shulman, L.; Cirincione, C.; Berry, D.; Becker, H.; Perez, E.; O'Regan, R.; Martino, S.; Atkins, J.; Hudis, C.; Winer, E.; Cancer; BLG. Four vs. 6 cycles of doxorubicin and cyclophosphamide (AC) or paclitaxel (T) as adjuvant therapy for breast cancer in women with 0–3 positive axillary nodes: CALGB 40101-a 2x2 factorial phase III trial: first results comparing 4 vs. 6 cycles of therapy. San Antonio Breast Cancer Symposium; 2010. p. S6-3

- Skol A, Scott L, Abecasis G, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nature Genetics*. 2006; 38(2):209–213. [PubMed: 16415888]
- Spencer C, Su Z, Donnelly P, Marchini J. Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. *PLoS Genetics*. 2009; 5(5):e1000477. [PubMed: 19492015]
- Therneau, T. R package version 2.36-9. 2011. *survival: Survival analysis, including penalised likelihood*.
- van der Vaart, A.; Wellner, J. Weak convergence and empirical processes. Springer-Verlag; 1996.

A Derivation of the asymptotic distribution of the score statistic

Denote

$$\mathcal{E}(\theta, t) = \sum_{i=1}^n Y_i(t) \omega(G_i) e^{\theta \omega(G_i)} / \sum_{i=1}^n Y_i(t) e^{\theta \omega(G_i)},$$

and $S_{\theta^*}^{(j)}(\theta, t) = \mathbb{E}_{\theta^*} [Y(t) \{\omega(G)\}^j e^{\theta \omega(G)}]$, $j = 0, 1, 2$. Let $e(\theta, t) = S_{\theta}^{(1)}(\theta, t) / S_{\theta}^{(0)}(\theta, t)$, and $e_0(\theta, t) = S_{\theta}^{(1)}(0, t) / S_{\theta}^{(0)}(0, t)$, where $\mathbb{E}_{\theta}$ means taking expectation assuming that the underlying true parameter value is θ . Then the score statistic for testing $H_0 : \theta = 0$ is

$$U_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^{\infty} [\omega(G_i) - \mathcal{E}(0, t)] dN_i(t).$$

Let $M_i(t) = N_i(t) - \int_0^t Y_i(s) e^{\theta \omega(G_i)} d\Lambda_0(s)$, $1 \leq i \leq n$. Then $\{M_i(t), t \geq 0\}$ is a mean zero martingale with respect to filtration $\{\mathcal{F}(t), t \geq 0\}$, where $\mathcal{F}(t)$ is the σ algebra generated by all observed data before time t . Under the alternative hypothesis we can write

$$\begin{aligned} U_n &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^{\infty} [\omega(G_i) - \mathcal{E}(0, t)] dN_i(t) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^{\infty} [\omega(G_i) - \mathcal{E}(0, t)] dM_i(t) + \sqrt{n} \int_0^{\infty} [\mathcal{E}(\theta, t) - \mathcal{E}(0, t)] S_{\theta}^{(0)}(\theta, t) d\Lambda_0(t) + \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^{\infty} [\mathcal{E}(\theta, t) - \mathcal{E}(0, t)] \{Y_i(t) e^{\theta \omega(G_i)} - S_{\theta}^{(0)}(\theta, t)\} d\Lambda_0(t). \end{aligned}$$

Since by law of large numbers $\mathcal{E}(\theta, t) \rightarrow e(\theta, t)$ and $\mathcal{E}(0, t) \rightarrow e_0(\theta, t)$ in probability as $n \rightarrow \infty$, by the previous expression and empirical process theory (van der Vaart and Wellner, 1996), we obtain that

$$\begin{aligned} U_n &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^{\infty} [\omega(G_i) - e_0(\theta, t)] dM_i(t) + \sqrt{n} \int_0^{\infty} [e(\theta, t) - \mathcal{E}(0, t)] S_{\theta}^{(0)}(\theta, t) d\Lambda_0(t) \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^{\infty} [e(\theta, t) - e_0(\theta, t)] \{Y_i(t) e^{\theta \omega(G_i)} - S_{\theta}^{(0)}(\theta, t)\} d\Lambda_0(t) + o_p(1) \\ &\equiv A_n + B_n + o_p(1), \end{aligned}$$

where

$$A_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^\infty [\omega(G_i) - e_0(\theta, t)] dM_i(t) + \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^\infty [e(\theta, t) - e_0(\theta, t)] \{Y_i(t) e^{\theta \omega(G_i)} - S_\theta^{(0)}(\theta, t)\} d\Lambda_0(t),$$

and

$$B_n = \sqrt{n} \int_0^\infty [\mathcal{E}(\theta, t) - \mathcal{E}(0, t)] S_\theta^{(0)}(\theta, t) d\Lambda_0(t).$$

The details of the proof of the above decomposition can be obtained from the authors. We further decompose B_n as

$$\begin{aligned} & \sqrt{n} \int_0^\infty [\mathcal{E}(\theta, t) - \mathcal{E}(0, t)] S_\theta^{(0)}(\theta, t) d\Lambda_0(t) \\ &= \sqrt{n} \int_0^\infty [\mathcal{E}(\theta, t) - e(\theta, t)] S_\theta^{(0)}(\theta, t) d\Lambda_0(t) - \sqrt{n} \int_0^\infty [\mathcal{E}(0, t) - e_0(\theta, t)] S_\theta^{(0)}(\theta, t) d\Lambda_0(t) + \sqrt{n} \mu(\theta), \end{aligned}$$

where

$$\sqrt{n} \mu(\theta) = \sqrt{n} \int_0^\infty [e(\theta, t) - e_0(\theta, t)] S_\theta^{(0)}(\theta, t) d\Lambda_0(t).$$

By a similar argument by empirical process theory, for which the details are omitted again, the first term in the decomposition of B_n is

$$\begin{aligned} & \sqrt{n} \int_0^\infty [\mathcal{E}(\theta, t) - e(\theta, t)] S_\theta^{(0)}(\theta, t) d\Lambda_0(t) \\ &= \sqrt{n} \int_0^\infty \frac{S_\theta^{(0)}(\theta, t)}{\sum_{j=1}^n Y_j(t) e^{\theta Z_j} / n} \left\{ \frac{1}{n} \sum_{i=1}^n Y_i(t) \omega(G_i) e^{\theta \omega(G_i)} - S_\theta^{(1)}(\theta, t) \right\} d\Lambda_0(t) \\ & \quad - \sqrt{n} \int_0^\infty \frac{S_\theta^{(1)}(\theta, t)}{\sum_{j=1}^n Y_j(t) e^{\theta Z_j} / n} \left\{ \frac{1}{n} \sum_{i=1}^n Y_i(t) e^{\theta \omega(G_i)} - S_\theta^{(0)}(\theta, t) \right\} d\Lambda_0(t) \\ &= \sqrt{n} \int_0^\infty \left\{ \frac{1}{n} \sum_{i=1}^n Y_i(t) \omega(G_i) e^{\theta \omega(G_i)} - S_\theta^{(1)}(\theta, t) \right\} d\Lambda_0(t) \\ & \quad - \sqrt{n} \int_0^\infty \frac{S_\theta^{(1)}(\theta, t)}{S_\theta^{(0)}(\theta, t)} \left\{ \frac{1}{n} \sum_{i=1}^n Y_i(t) e^{\theta \omega(G_i)} - S_\theta^{(0)}(\theta, t) \right\} d\Lambda_0(t) + o_p(1). \end{aligned}$$

Similarly, for the second term we have

$$\begin{aligned} & \sqrt{n} \int_0^\infty \{\mathcal{E}(0, t) - e_0(\theta, t)\} S_\theta^{(0)}(\theta, t) d\Lambda_0(t) \\ &= \sqrt{n} \int_0^\infty \frac{S_\theta^{(0)}(\theta, t)}{S_\theta^{(0)}(0, t)} \left\{ \frac{1}{n} \sum_{i=1}^n Y_i(t) \omega(G_i) - S_\theta^{(1)}(0, t) \right\} d\Lambda_0(t) \\ & \quad - \sqrt{n} \int_0^\infty \frac{S_\theta^{(0)}(\theta, t) S_\theta^{(1)}(0, t)}{\{S_\theta^{(0)}(0, t)\}^2} \left\{ \frac{1}{n} \sum_{i=1}^n Y_i(t) - S_\theta^{(0)}(0, t) \right\} d\Lambda_0(t) + o_p(1). \end{aligned}$$

Finally we obtain the following decomposition for U_n :

$$U_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n [\xi_i(\theta) + \eta_i(\theta) - \mathbb{E}\{\xi_i(\theta) + \eta_i(\theta)\}] + \sqrt{n}\mu(\theta) + o_p(1),$$

where $\xi_i(\theta) = \int_0^\infty \{\omega(G_i) - e_0(\theta, t)\} dM_i(t)$ and

$$\eta_i(\theta) = \int_0^\infty \{\omega(G_i) e^{\theta\omega(G_i)} + a_1(\theta, t) e^{\theta\omega(G_i)} + a_2(\theta, t) \omega(G_i) + a_3(\theta, t)\} Y_i(t) d\Lambda_0(t),$$

with

$$a_1(\theta, t) = -\frac{S_\theta^{(1)}(0, t)}{S_\theta^{(0)}(0, t)}, \quad a_2(\theta, t) = -\frac{S_\theta^{(0)}(\theta, t)}{S_\theta^{(0)}(0, t)}, \quad \text{and } a_3(\theta, t) = \frac{S_\theta^{(0)}(\theta, t) S_\theta^{(1)}(0, t)}{\{S_\theta^{(0)}(0, t)\}^2}.$$

It follows that

$$U_n - \sqrt{n}\mu(\theta) \rightarrow N(0, \sigma^2(\theta))$$

in distribution, as $n \rightarrow \infty$, where $\sigma^2(\theta) = \text{var}\{\xi(\theta) + \eta(\theta)\}$.

Next, we derive the limit of $\widehat{\sigma}_0^2$. By law of large numbers, we have

$$\begin{aligned} \widehat{\sigma}_0^2 &= \frac{1}{n} \sum_{i=1}^n \int_0^\infty \left[\frac{s_2(0, t)}{s_0(0, t)} - \left\{ \frac{s_1(0, t)}{s_0(0, t)} \right\}^2 \right] dN_i(t) \\ &= \frac{1}{n} \sum_{i=1}^n \int_0^\infty \left[\frac{s_2(0, t)}{s_0(0, t)} - \left\{ \frac{s_1(0, t)}{s_0(0, t)} \right\}^2 \right] dM_i(t) + \int_0^\infty \left[\frac{s_2(0, t)}{s_0(0, t)} - \left\{ \frac{s_1(0, t)}{s_0(0, t)} \right\}^2 \right] s_0(\theta, t) d\Lambda_0(t) \\ &\rightarrow \int_0^\infty \left[\frac{S_\theta^{(2)}(0, t)}{S_\theta^{(0)}(0, t)} - \left\{ \frac{S_\theta^{(1)}(0, t)}{S_\theta^{(0)}(0, t)} \right\}^2 \right] S_\theta^{(0)}(\theta, t) d\Lambda_0(t) \\ &\equiv \sigma_0^2(\theta), \end{aligned}$$

in probability, as $n \rightarrow \infty$. By the above results, we have

$$\frac{\{\sigma_0(\theta) W_n - \sqrt{n}\mu(\theta)\}}{\sigma(\theta)} \rightarrow N(0, 1)$$

in distribution, as $n \rightarrow \infty$.

B Calculation of asymptotic variances

Assume independence between $\tilde{T}$ and C given G .

B.1 Variance of $\xi(\theta)$

The variance of $\xi(\theta)$ can be expressed as

$$\begin{aligned}\text{var}\{\xi(\theta)\} &= \text{var} \left[\int_0^\infty \{\omega(G) - e_0(\theta, t)\} dM(t) \right] = E_\theta \int_0^\infty \{\omega(G) - e_0(\theta, t)\}^2 Y(t) e^{\theta\omega(G)} d\Lambda_0(t) \\ &= \lambda_0 \sum_{g=0}^2 \pi_g e^{\theta\omega(g)} \int_0^\infty \{\omega(g) - e_0(\theta, t)\}^2 e^{-e^{\theta\omega(g)} \Lambda_0(t)} \bar{F}_c^g(t) d\Lambda_0(t),\end{aligned}$$

where $\pi_g = P(G = g)$, and $\bar{F}_c^g(t)$ is the survival function of the censoring variable C given $G = g$.

B.2 Variance of $\eta(\theta)$

To calculate the variance of $\eta(\theta)$, we need to calculate $E\{\eta^2(\theta)\}$ and $E\{\eta(\theta)\}$. Denoting $F(t, c|G = g)$ to be the distribution function of $(\tilde{T}, C)$ given $G = g$, for $E\{\eta^2(\theta)\}$ we can write it as

$$\sum_{g=0}^2 \pi_g \int_0^\infty \int_0^\infty \left\{ \int_0^{t \wedge c} d_g(\theta, u) d\Lambda_0(u) \right\}^2 dF(t, c|G = g) = \sum_{g=0}^2 \pi_g \int_0^\infty \int_0^\infty \left\{ \int_0^{t \wedge c} d_g(\theta, u) d\Lambda_0(u) \right\}^2 \lambda_0(t) e^{\theta\omega(g)} e^{-e^{\theta\omega(g)} \Lambda_0(t)} f_c^g(c) dc dt, \quad (1)$$

where $f_c^g(c)$ is the density function of C given $G = g$ and

$$d_g(\theta, u) = \omega(g) e^{\theta\omega(g)} + a_1(\theta, u) e^{\theta\omega(g)} + a_2(\theta, u) \omega(g) + a_3(\theta, u).$$

The integral on the right hand side of (1) is equal to

$$\begin{aligned}& \int_0^\infty \int_0^\infty \left\{ \int_0^{t \wedge c} d_g(\theta, u) d\Lambda_0(u) \right\}^2 \lambda_0(t) e^{\theta\omega(g)} e^{-e^{\theta\omega(g)} \Lambda_0(t)} f_c^g(c) dc dt \\ &= \int_{t \leq c, a \leq c \leq b} \left\{ \int_0^t d_g(\theta, u) d\Lambda_0(u) \right\}^2 \lambda_0(t) e^{\theta\omega(g)} e^{-e^{\theta\omega(g)} \Lambda_0(t)} f_c^g(c) dc dt \\ &+ \int_{t > c, a \leq c \leq b} \left\{ \int_0^c d_g(\theta, u) d\Lambda_0(u) \right\}^2 \lambda_0(t) e^{\theta\omega(g)} e^{-e^{\theta\omega(g)} \Lambda_0(t)} f_c^g(c) dc dt \\ &= e^{\theta\omega(g)} \int_0^\infty \left\{ \int_0^t d_g(\theta, u) d\Lambda_0(u) \right\}^2 \lambda_0(t) e^{-e^{\theta\omega(g)} \Lambda_0(t)} \bar{F}_c^g(t) dt \\ &+ \int_0^\infty \left\{ \int_0^c d_g(\theta, u) du \right\}^2 e^{-e^{\theta\omega(g)} \Lambda_0(c)} f_c^g(c) dc.\end{aligned}$$

The $E\{\eta(\theta)\}$ term can be calculated as follows:

$$E\{\eta(\theta)\} = \sum_{g=0}^2 \pi_g \int_0^\infty d_g(\theta, t) \bar{F}_c^g(t) e^{-e^{\theta\omega(g)} \Lambda_0(t)} d\Lambda_0(t).$$

B.3 Covariance between $\xi(\theta)$ and $\eta(\theta)$

The covariance can be expressed as

$$\begin{aligned}
& \sum_{g=0}^2 \pi_g \int_0^\infty \int_0^\infty \left[\{\omega(g) - e_0(\theta, t)\} I(t \leq c) - \lambda_0 e^{\theta\omega(g)} \int_0^{t \wedge c} \{\omega(g) - e_0(\theta, u)\} du \right] \\
& \quad \times \left\{ \int_0^{t \wedge c} d_g(\theta, u) du \right\} dF(t, c | G=g) \\
& = \sum_{g=0}^2 \pi_g \int_0^\infty \int_0^\infty \left[\{\omega(g) - e_0(\theta, t)\} I(t \leq c) - \lambda_0 e^{\theta\omega(g)} \int_0^{t \wedge c} \{\omega(g) - e_0(\theta, u)\} du \right] \\
& \quad \times \left\{ \int_0^{t \wedge c} d_g(\theta, u) du \right\} e^{\theta\omega(g)} \lambda_0(t) e^{-e^{\theta\omega(g)} \Lambda_0(t)} f_c^g(c) dt dc \\
& = C_1(g) + C_2(g),
\end{aligned}$$

where

$$\begin{aligned}
C_1(g) &= \sum_{g=0}^2 \pi_g \int \int_{0 \leq t \leq c} \left[\{\omega(g) - e_0(\theta, t)\} - \lambda_0 e^{\theta\omega(g)} \int_0^t \{\omega(g) - e_0(\theta, u)\} du \right] \times \left\{ \int_0^t d_g(\theta, u) du \right\} e^{\theta\omega(g)} \lambda_0(t) e^{-e^{\theta\omega(g)} \Lambda_0(t)} f_c^g(c) dt dc \\
&= \sum_{g=0}^2 \pi_g \int_0^\infty \left[\{\omega(g) - e_0(\theta, t)\} - \lambda_0 e^{\theta\omega(g)} \int_0^t \{\omega(g) - e_0(\theta, u)\} du \right] \times \left\{ \int_0^t d_g(\theta, u) du \right\} e^{\theta\omega(g)} \lambda_0(t) e^{-e^{\theta\omega(g)} \Lambda_0(t)} \bar{F}_c^g(t) dt,
\end{aligned}$$

and

$$C_2(g) = - \sum_{g=0}^2 \pi_g \int_0^\infty \lambda_0 e^{\theta\omega(g)} \int_0^c \{\omega(g) - e_0(\theta, u)\} du \times \left\{ \int_0^c d_g(\theta, u) du \right\} e^{-e^{\theta\omega(g)} \Lambda_0(c)} dc.$$

All the above variances and covariance can be expressed by integrals of univariate functions.

Figure 1.

Illustration of the mixture distributions used in the examples. The survival function for the population, $S(t) = \mathbb{P}[\tilde{T} > t]$ (solid line), and the conditional survival probabilities, $S_g(t) = \mathbb{P}[\tilde{T} > t | G = g]$ for $g = 0, 1, 2$ (dotted lines), are drawn. The baseline hazard rate λ_0 is calculated based on setting $\mathbb{P}[\tilde{T} > 6] = 0.5$ and $\text{GHR} = 2$. The relative risk allele frequency is denoted by q .

Figure 2.

Illustration of the empirical type I error rate for the score test. The empirical rate is to be compared to a nominal two-sided level of 0.01. The sample sizes are $n = 1,000$ (top panel), $n = 500$ (middle panel) and $n = 100$ (bottom panel). Within each panel, uniform event rates of $\rho \in \{0.5, 0.7, 0.9\}$ are considered. Each example is based on $B = 10,000$ simulation replicates.

Figure 3.

Illustration of the empirical power for the score test at the nominal two-sided level of 0.01 based on a sample size of $n = 100$. For each example, the asymptotic power based on the exact and approximate variance formulas, and the empirical power based on $B = 10,000$ simulation replicates, are provided. The relative risk allele frequency and the uniform event rates are denoted by q and ρ respectively.

Figure 4.

Illustration of the empirical power for the score test at the nominal two-sided level of 0.01 based on a sample size of $n = 250$. For each example, the asymptotic power based on the exact and approximate variance formulas, and the empirical power based on $B = 10,000$ simulation replicates, are provided. The relative risk allele frequency and the uniform event rates are denoted by q and ρ respectively.

Figure 5.

Illustration of the empirical power for the score test at the nominal two-sided level of 0.01 based on a sample size of $n = 500$. For each example, the asymptotic power based on the exact and approximate variance formulas, and the empirical power based on $B = 10,000$ simulation replicates, are provided. The relative risk allele frequency and the uniform event rates are denoted by q and ρ respectively.

Figure 6.

Illustration of power calculations for a planned GWAS study based on a planned sample size of $n = 1,000$. A event rate of $p = 0.75$ is assumed. For a given risk relative allele frequency q the power is provided based on the exact variance formula as a function of Δ . The analyses are powered for an additive risk model. It is assumed that 1,000,000 variants are to be tested for association with the time-to-event outcome. The study is designed to conservatively control the FWER at the $\alpha = 0.05$ level by testing each variant at the 5×10^{-8} level.