Web-based Supporting Materials for Power/Sample Size Calculations for Assessing Correlates of Risk in Clinical Efficacy Trials

Peter B. Gilbert, Holly E. Janes, Yunda Huang

Appendix A: Unbiased Biomarker Characterization Accounting for the Sampling Design of the CoR Study

Consider a 2-phase sampling design (without-replacement) with K participant strata defined by variables measured in all study participants. Let N^*_{1k} (N^*_{0k}) be the number of vaccine recipient cases (controls) in stratum k at-risk at τ (i.e., with $Y^\tau = 0$), and N_{1k} (N_{0k}) be the numbers *observed* to be at-risk at τ (i.e., with $X^\tau = 0$), with $N^*_z \equiv \sum_{k=1}^K N^*_{zk}$ and $N_z \equiv \sum_{k=1}^K N_{zk}$ for $z = 0, 1$. The unstarred quantities are not observed (unless there is no dropout by τ) but their expectations can easily be estimated by the numbers of randomized subjects observed to be cases and controls multiplied by an estimate of the probability of primary endpoint occurrence by τ (e.g., a Kaplan-Meier estimate). Let n_{1k} (n_{0k}) be the number of vaccine recipient cases (controls) in stratum k observed to be at-risk at τ from whom immune responses are measured at τ. In practice n_{1k} is set to include all N_{1k} subjects who have available specimens at τ (typically slightly less than N_{1k}). Different approaches may be taken to choose the n_{0k}; for example one approach achieves an overall case-control ratio $r \equiv \sum_{k=1}^K n_{0k} / \sum_{k=1}^K n_{1k}$ with r in the range of 2 to 5, where the n_{0k} may all equal $r \times n_{1k}$ or may upweight certain strata judged to be important.

A consideration for the sampling design is that vaccine trials with a correlates objective also have the objective to characterize the immunogenicity of the vaccine. To represent the trial population this analysis should provide unbiased descriptive and inferential analysis for the population of vaccine recipients at-risk at τ (possibly

within strata) not conditioning on case status. Both approaches can straightforwardly be used to provide inference on parameters of interest using all $n_1 \equiv \sum_{k=1}^{K} n_{1k}$ and $n_0 \equiv \sum_{k=1}^{K} n_{0k}$ subjects, for example by using inverse probability weighting. However, for graphical analysis, the prospective case-cohort approach straightforwardly provides a correct random sample, whereas the outcome-dependent sampling plan does not. This problem can be remedied by defining each $n_{1k}^{IS} \leq n_{1k}$ $(k = 1, \cdots, K)$ to be the number of cases included in the immunogenicity characterization analysis selected to maintain a controls:cases ratio of sampled subjects equal to the controls:cases ratio of the entire study cohort, i.e., to satisfy the constraint

$$\frac{n_{0k}}{n_{1k}^{IS}} = \frac{\widehat{E}[N_{0k}]}{\widehat{E}[N_{1k}]}. \tag{1}$$

The estimates $\widehat{E}[N_{0k}]$ and $\widehat{E}[N_{1k}]$ are determined independently of considerations of the immunogenicity and correlates studies, and any choices of n_{0k} and n_{1k}^{IS} satisfying (1) will allow unbiased immunogenicity analysis within each covariate subgroup k. While this approach provides unbiased immunogenicity analysis for each stratum k separately, if certain strata k are over-sampled it may provide biased analysis for the overall study population. We can obtain unbiased analysis of the overall population by including the immune response data from all n_0^* controls and from $n_1^* \equiv f_k n_{1k}^{IS}$ cases, where the constants $f_k \leq 1$ are selected to achieve each $f_k n_{1k}^*$ being equal to an integer and $n_0^*/n_1^{IS} = \widehat{E}[N_0]/\widehat{E}[N_1]$, where $n_1^{IS} \equiv \sum_{k=1}^{K} n_{1k}^{IS}$.

One way to implement the above approach is to first choose the n_{0k} $(k = 1, \cdots, K)$ to achieve adequate power for the overall correlates analysis, which determines the n_{1k}^{IS} by equation (1) (rounding to the nearest integer). Then, if necessary for the overall analysis, add the second fix on top of this fix. This discussion shows that it is straightforward to conduct an unbiased immunogenicity characterization study regardless of whether the correlates analysis uses prospective case-cohort or retrospective 2-phase sampling.

Appendix B: Selected Mathematical Details of Power Calculations

Computing Sensitivity, Specificity, False Positives, and False Negatives

Given inputs σ^2_{obs}, ρ, P_0^{lat}, P_2^{lat}, P_0, and P_2, the following steps yield Sens, Spec FP^1, FP^2, FN^1, and FN^2 defined in the main manuscript.

1. Set $\sigma^2_e = (1-\rho)\sigma^2_{obs}$ and solve for θ_2 in the equation $P_2^{lat} = P(X^* > \theta_2)$: $\theta_2 = \sqrt{\rho}\sigma_{obs}\Phi^{-1}(1-P_2^{lat})$. Similarly solve for θ_0 in the equation $P_0^{lat} = P(X^* \leq \theta_0)$: $\theta_0 = \sqrt{\rho}\sigma_{obs}\Phi^{-1}(P_0^{lat})$.

2. Simulate a large number M of realizations of X^* and S^* from normal distributions $N(0, \rho\sigma^2_{obs})$ and $N(0, \sigma^2_{obs})$, respectively (e.g., $M = 100,000$).

3. With $P_2(\theta_2) \equiv P(S^* > \theta_2)$ and $P_0(\theta_0) \equiv P(S^* \leq \theta_0)$, determine the cut-points θ_2 and θ_0 that solve equations

$$P_2 = Sens * P_2^{lat} + FP^2 * P_1^{lat} + FP^1 * P_0^{lat}$$

and

$$P_0 = Spec * P_0^{lat} + FN^2 * P_1^{lat} + FN^1 * P_2^{lat}$$

in the main manuscript, which are the solutions to

$$P_2(\theta_2) = Sens(\theta_2)P_2^{lat} + FP^2(\theta_2)P_1^{lat} + FP^1(\theta_2)P_0^{lat} \quad (1)$$

$$P_0(\theta_0) = Spec(\theta_0)P_0^{lat} + FN^2(\theta_0)P_1^{lat} + FN^1(\theta_0)P_2^{lat}. \quad (2)$$

The solution θ_2 is obtained by estimating ($P_2(\theta_2)$, $Sens(\theta_2)$, $FP^1(\theta_2)$, $FP^2(\theta_2)$) for each of the M realizations and picking the $\theta_2 = \theta$ that gives the closest solution. Similarly the solution θ_0 is obtained by estimating ($P_0(\theta_0)$, $Spec(\theta_0)$, $FN^1(\theta_0)$, $FN^2(\theta_0)$) for each of the M realizations and picking the $\theta_0 = \theta$ that gives the closest solution.

4. Output the resulting solutions θ_2 and θ_0 together with $P_2(\theta_2), Sens(\theta_2)$, $FP^1(\theta_2), FP^2(\theta_2)$ evaluated at the solution θ_2 and $P_0(\theta_0), Spec(\theta_0)$, $FN^1(\theta_0), FN^2(\theta_0)$ evaluated at the solution θ_0.

Solutions α^{lat} and β^{lat} for a Continuous Biomarker

Given fixed (VE, $risk_0$, $P^{lat}_{lowestVE}$, VE_{lowest}), β^{lat} in the model of Section 2.4 in the main article can be expressed as a function of α_{lat} by fixing $x = \nu$. This yields

$$\beta^{lat} = \frac{1}{\nu}\left[logit\left(risk_1^{lat}(\nu)\right) - \alpha^{lat}\right]. \tag{3}$$

Plugging (3) into the last formula in Section 2.4 for overall VE yields a zero-equation $U(\alpha^{lat}) = 0$ in one unknown variable α^{lat},

$$U(\alpha^{lat}) = (1 - VE) - \frac{P^{lat}_{lowestVE} * risk_1^{lat}(\nu) + \int_\nu^\infty D(x;\alpha^{lat})\phi(x/(\sqrt{\rho}\sigma_{obs}))dx}{risk_0} \tag{4}$$

where $D(x;\alpha^{lat}) \equiv A(x;\alpha^{lat})/\left[(1 - risk_1^{lat}(\nu))^{x/\nu} + A(x;\alpha^{lat})\right]$ with $A(x;\alpha^{lat}) \equiv exp\{\alpha^{lat} * (1 - x/\nu)\} * \left[risk_1^{lat}(\nu)\right]^{x/\nu}$. Equation (4) can be solved by a one-dimensional line search. Then, β^{lat} is solved by plugging α^{lat} into equation (3).

Appendix C: Estimation of the Noise Level of a Biomarker

As described in model

$$S^* = X^* + e, \quad X^* \sim N(0, \sigma^2_{tr}), \quad e \sim N(0, \sigma^2_e)$$

of the main article (Section 2.3), the continuous-readout biomarker S^* is often measured with protection-irrelevant error, denoted by e. Typically, the error is due to two major independent sources of variability: assay-related error, e_{assay}, and trial-related error, e_{trial}. We suppose that

$$e = e_{assay} + e_{trial}, \quad e_{assay} \sim N(0, \sigma^2_{assay}), \quad e_{trial} \sim N(0, \sigma^2_{trial}), \text{ and } e_{assay} \perp e_{trial}.$$

Consequently, $\sigma^2_e = \sigma^2_{assay} + \sigma^2_{trial}$ and $\rho = 1 - \sigma^2_{assay}/\sigma^2_{obs} - \sigma^2_{trial}/\sigma^2_{obs}$. We describe how the proportion of variability due to trial-related error, $\pi_t = \sigma^2_{trial}/\sigma^2_{obs}$, and the

proportion of variability due to assay-related error, $\pi_a = \sigma^2_{assay}/\sigma^2_{obs}$, can be estimated with data from the CoR study or from external studies.

The component of variability σ^2_{trial} represents variability due to trial-related factors such as differing specimen collection or storage practices. Typically, a study protocol controls these factors to some extent, but some variation still exists. Another common source of trial-related error is deviation in the time of specimen collection from the target time. Most protocols place "windows" around the targeted time for specimen collection, e.g. 7–10 days, and so variation in timing within the allowable window is to be expected. Deviation in specimen collection from the target time affects the biomarker readout by creating variability in the interval between treatment administration and specimen collection.

When data related to these trial-related factors are available in the CoR study, and if it can be assumed that these factors are not collinear with other factors in the CoR study influencing the biomarker readout, such as subject characteristics or assay conditions, the trial-related proportion of variability $\pi_t = \sigma^2_{trial}/\sigma^2_{obs}$ can be estimated in the following way. Under a linear model with dependent variable S^* and trial-related factors as independent terms, the ratio of the regression sum of squares to the total sum of squares is an estimate of π_t.

The term σ^2_{assay} represents the sum of two types of non-systematic components of variability: one is generated when a biomarker is repeatedly measured under the same assay conditions (e.g., by the same technician using the same instrument on the same day), and the other is generated when a biomarker is assessed under different assay conditions. The second component of variability can be assumed to equal zero when all specimens in the CoR study are assessed under the same assay condition. Ideally, the assay-related proportion of variability, $\pi_a = \sigma^2_{assay}/\sigma^2_{obs}$, would also be estimated using data from the CoR study. However, it is often infeasible to obtain the necessary data, given limitations on specimen volume, especially when multiple assay conditions

are involved. Therefore, an external validation study is commonly employed. The validation study should examine all assay-related factors introducing variability in the CoR study. In general, the ideal design of the study is a full factorial randomized block design. An example is shown in Web Table 1 for two assay factors A and B. For concreteness, suppose factor A is the assay technician and factor B is the assay instrument. Each study subject represents one block, and the specimens from each subject are assayed by each technician and using each instrument, and are replicated at least twice. Typically specimens from at least 3 subjects are included in the validation study, and are chosen so that their biomarker readouts span the range of expected levels of response. Under this design, π_a can be estimated using a linear model with dependent variable S^* and technician and instrument factors, in addition to a subject identifier, as independent variables. The ratio of the sum of squares due to technician and instrument to the total sum of squares is an estimate of π_a. Application of this π_a estimate to the CoR study is then valid assuming that the proportion of variability due to technician and instrument is the same in the validation and CoR studies. A stronger condition, not required but further supporting the transfer of the π_a estimate to the CoR study, is that the distribution of biomarker readout in the validation study matches that in the CoR study. The final step is to calculate the estimate of the assay measurement error as $\widehat{\rho} = 1 - \widehat{\pi}_a - \widehat{\pi}_t$.

Web Table 1. Ideal experimental design for estimating $\pi_a = \sigma^2_{assay}/\sigma^2_{obs}$ given two assay-related factors, technician and assay instrument, that introduce variability in the assay readout, S^*. Here S^*_{ijkh} is the biomarker readout when the ith technician performs the assay ($i = 1, 2, ...a$) and the jth instrument is used ($j = 1, 2, ..., b$) for the kth study subject or block ($k = 1, ..., n$) and the hth replicate ($h = 1, 2, ...m$).

Subject (block)	1			...			n		
Technician	1	...	a	1	...	a	1	...	a
Instrument									
1									
...					S^*_{ijkh}				
b									

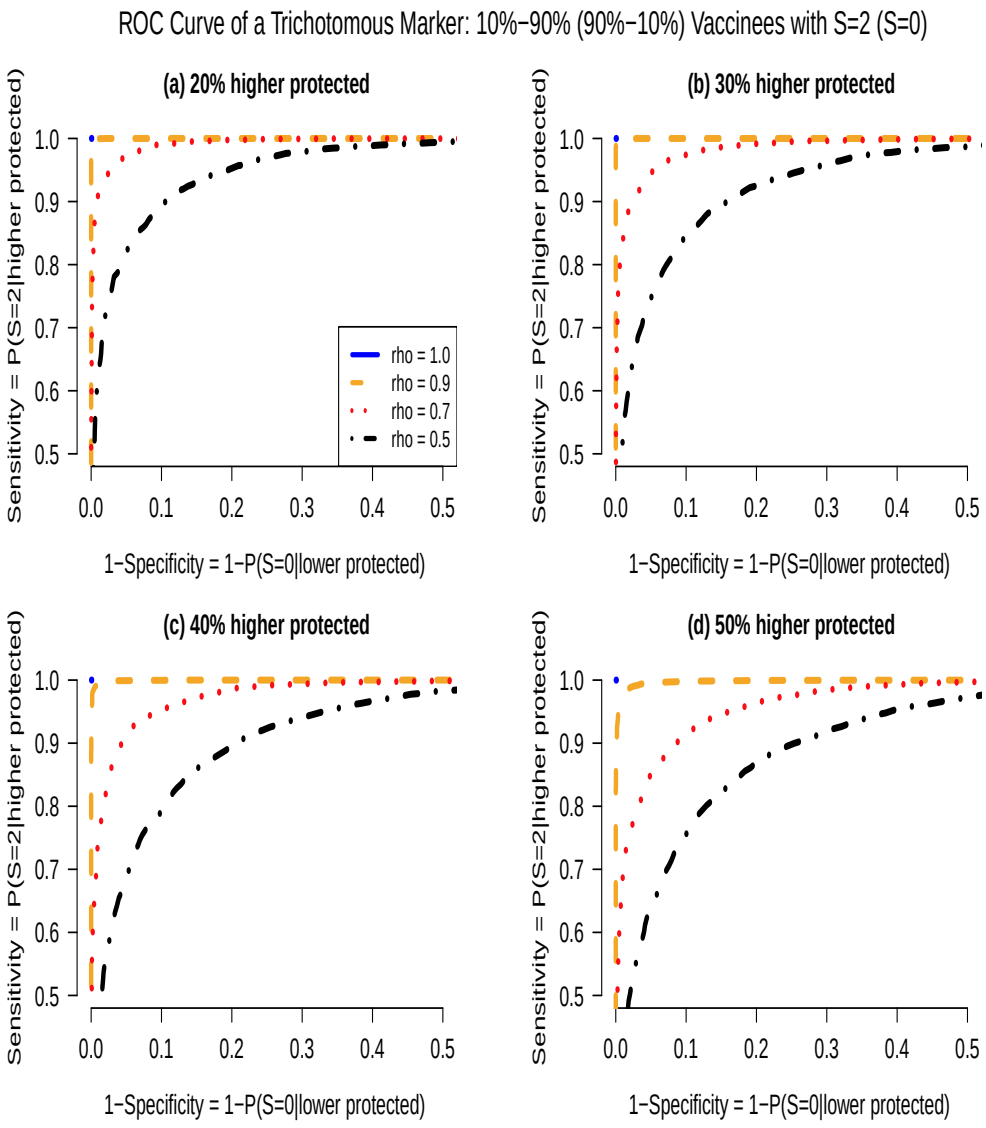
Appendix D: Additional Power Figures for the Illustrations in Section 4

Web Figure Legends

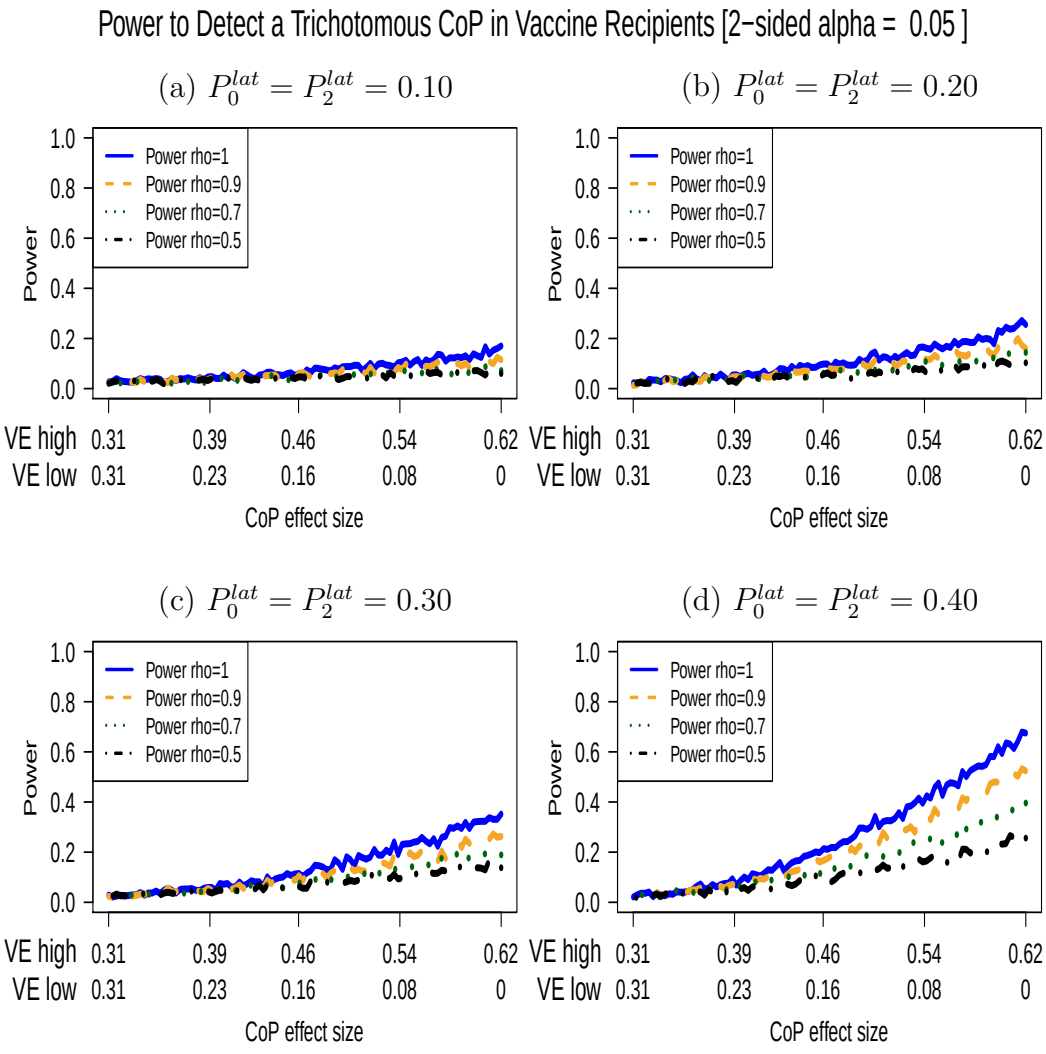
Web Supporting Figure 1. Sensitivity and specificity for $P_2 = 1 - P_0$ ranging from 0.10 to 0.90, for four scenarios of ρ.

Web Supporting Figure 2. Correlates of Protection (VE curve) power calculations for a trichotomous biomarker S for the completed RV144 HIV vaccine efficacy trial with $n_{cases} = 41$ and $n_{controls} = 205$ (1-sided $\alpha = 0.025$-level Wald test), for four scenarios of ρ. This figure is based on the same simulation study as Figure 4.

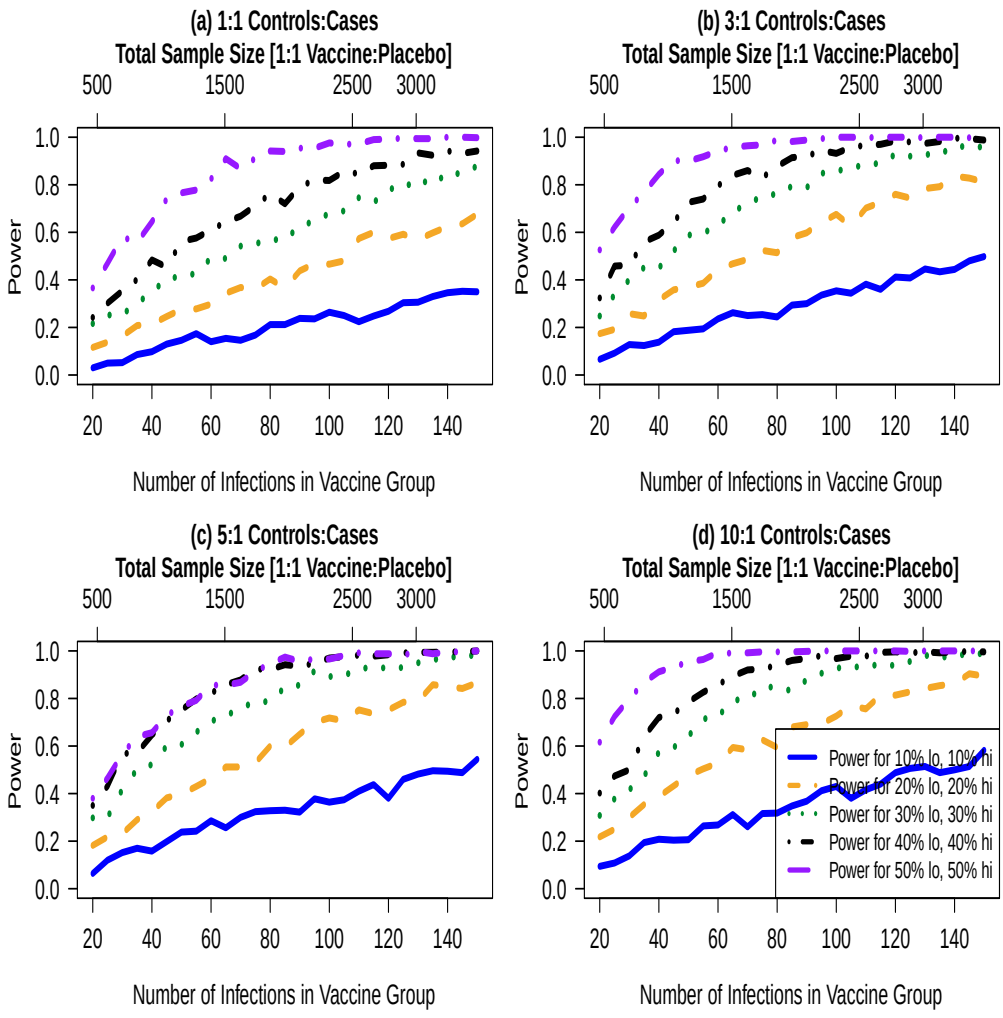
Web Supporting Figure 3. Power curves versus total sample size for a trichotomous biomarker S to plan a 2-arm HIV vaccine efficacy trial with equal allocation randomization to vaccine versus placebo, 4% annual placebo incidence, 5% annual dropout incidence, overall $VE = 0.50$, and correlate of protection effect size $VE_0^{lat} = 0.25$, $VE_1^{lat} = 0.50$, $VE_2^{lat} = 0.75$ for $\rho = 0.9$. Each panel shows power for $P_0 = P_0^{lat} = P_2^{lat} = P_2$ ranging from 0.1 to 0.5, for controls:cases allocation ratios (a) 1:1, (b) 3:1, (c) 5:1, and (d) 10:1. This figure is based on the same simulation study as Figure 6.



Web Supporting Figure 1



Web Supporting Figure 2



Web Supporting Figure 3

Appendix E: Summary of Implementation of the Methods in R

The without-replacement sampling version of the methods are implemented in the R package CoRpower posted at the first author's website

`http://faculty.washington.edu/peterg/programs.html?`

The R package includes code implementing the examples of Section 4 of the manuscript.