Published in final edited form as:

Stat Med. 2009 February 15; 28(4): 716–719. doi:10.1002/sim.3503.

RESPONSE TO ANDREW DUNNING’S COMMENT ON “EVALUATING A SURROGATE ENDPOINT AT THREE LEVELS, WITH APPLICATION TO VACCINE DEVELOPMENT”

Peter B. Gilbert, Li Qin, and Steven G. Self

Fred Hutchinson Cancer Research Center and Department of Biostatistics, University of Washington, Seattle, Washington, 98109, U.S.A

SUMMARY

In replying to Andrew Dunning’s helpful letter, we endeavor to clarify the value of incorporating baseline participant predictors of the potential surrogate endpoint and of the clinical endpoint into the evaluation of a surrogate endpoint. Such baseline predictors are useful for the evaluation of surrogates defined within each of the statistical and principal surrogate frameworks, which are complementary.

Keywords

Clinical trial; Counterfactual; Immune correlate; Potential outcomes; Principal surrogate

We welcome the opportunity to respond to Dr. Dunning’s valuable comments. We hope our reply will amplify key points about surrogate endpoint evaluation that were insufficiently discussed in our article.

On terminology, we enthusiastically welcome the suggestion of Sadoff and Wittes [1] to rename our proposed second and third tiers of immune correlates of protection to a “specific surrogate of protection (SoP)” and a “general SoP,” respectively, which convey whether the surrogate predicts vaccine efficacy (VE) for the same setting as studied in the vaccine trial or for generalized settings. Consensus terminology is important for the advancement of science; for example in the 17th century algebra left forward after Descartes coined the conventions to represent unknown variables by x , y , etc. and known constants by a , b , etc.; and calculus left forward after Leibniz coined the integral symbol $\int$ and the differential dx .

Dr. Dunning first suggests that our article “might also have introduced a condition that the surrogate quantitatively predict the efficacy of the vaccine.” We agree that this is an important criterion for checking surrogate value. Our definitions of a specific principal SoP and a general SoP in Table I explicitly included this criterion, and although we did not list it for a specific statistical SoP, it is equally important for it. Moreover, we discussed the role of this prediction criterion more directly in Qin et al. [2] and Gilbert and Hudgens [3].

*Correspondence to: Peter Gilbert, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Seattle, Washington, 98109, U.S.A. Phone: (206) 667-7299, pgilbert@scharp.org.

Specifically, for the same influenza vaccine trial example used in the current article [4], Qin et al. estimated the predicted VE against hospitalization with Weiss strain A based on Weiss strain A titers and fitted values of a logistic regression model, and compared the predicted VE to a nonparametric estimate of VE that ignored the immune response data, and conducted the same exercise for PR8 strain A titers. The prediction was quite accurate for Weiss strain A titers and not for PR8 strain A titers, which we take as evidence supporting that Weiss strain A titers have greater surrogate value. While this analysis was based on the statistical surrogate framework, wherein VE is predicted based on models for $\Pr(Y = 1|S = s, Z = z)$, $z = 0, 1$, Gilbert and Hudgens [3] suggested a parallel analysis using the principal surrogate framework, wherein VE is predicted based on models for $\Pr(Y(z) = 1|S(1) = s)$, $z = 0, 1$. Therefore VE may be predicted using either the statistical or principal surrogate frameworks. Comparing their prediction accuracy may inform about the relative fidelity of the two approaches.

Dr. Dunning secondly suggests that, when surrogate evaluation is based on the prediction criterion, then “the criteria for a statistical surrogate and a principal surrogate are substantially equivalent.” This equivalency holds only if a certain strong assumption holds, however, which we discussed in Section 2.4.2. This strong assumption may be stated as follows:

Key Assumption for the Statistical Surrogate Framework: No Unaccounted for Baseline Simultaneous Predictors of the Biomarker S and the Clinical Endpoint Y

Formally, this key assumption is that the accounted for baseline covariates contain all the *common causes* of S and Y as defined by Pearl [5]. As discussed in Section 2.4.2, the statistical surrogate approach evaluates the same causal effects as the principal surrogate approach under this assumption. The untestability of this assumption is a main motivator for the principal surrogate framework. The assumption holds in the example Dr. Dunning uses (his Table I) to illustrate his statement about the “substantial equivalency” of the two frameworks. However, as we now show, if the key assumption does not hold, such that an unmeasured baseline covariate X is introduced into the example that is associated with both S and Y , then a perfect principal surrogate may fail to be a perfect statistical surrogate, and vice versa. This will demonstrate that without the key assumption the frameworks are not substantially equivalent.

To make the example as simple as possible we suppose X is binary with levels 0 and 1, with $\Pr(X = 1) = .5$, and there is a 1:1 randomization to vaccine:placebo. The same conclusions hold for a general distribution of X and a general randomization allocation, although we do need to assume that X is balanced in the vaccine and placebo groups. We use Dunning’s notation, expanded to acknowledge X :

$$\begin{aligned}\pi_{i1}^x &\equiv \Pr(Y(1)=1|S(1)=i-1, S(0)=0, X=x), \\ \pi_{i2}^x &\equiv \Pr(Y(0)=1|S(1)=i-1, S(0)=0, X=x), \\ \phi^x &\equiv \Pr(S(1)=0, S(0)=0|X=x),\end{aligned}$$

for $i = 1, 2$ and $x = 0, 1$. The principal strata estimands $VE(0, 0)$ and $VE(1, 0)$ equal

$$V E(0, 0) = 1 - \frac{\pi_{11}^1 \phi^1 + \pi_{11}^0 \phi^0}{\pi_{12}^1 \phi^1 + \pi_{12}^0 \phi^0}, \quad (1)$$

$$V E(1, 0) = 1 - \frac{\pi_{21}^1 (1 - \phi^1) + \pi_{21}^0 (1 - \phi^0)}{\pi_{22}^1 (1 - \phi^1) + \pi_{22}^0 (1 - \phi^0)}. \quad (2)$$

$$\pi_{11}^1 \phi^1 + \pi_{11}^0 \phi^0 = \pi_{12}^1 \phi^1 + \pi_{12}^0 \phi^0 \quad (3)$$

A perfect principal surrogate satisfies $V E(0, 0) = 0$ and $V E(1, 0) = 1$, where the first condition is equivalent to

$$\pi_{11}^1 \phi^1 + \pi_{11}^0 \phi^0 = \pi_{12}^1 \phi^1 + \pi_{11}^0 \phi^0, \quad (4)$$

and the second condition is equivalent to

$$\pi_{21}^1 = \pi_{21}^0 = 0. \quad (5)$$

Now, the condition for a perfect statistical surrogate, $\Pr(Y = 1|S = 0, Z = 1) = \Pr(Y = 1|S = 0, Z = 0)$, is equivalent to

$$\pi_{11}^1 + \pi_{11}^0 = \pi_{12}^1 \phi^1 + \pi_{12}^0 \phi^0 + \pi_{22}^1 (1 - \phi^1) + \pi_{22}^0 (1 - \phi^0). \quad (6)$$

Dunning's condition that "the efficacy predicted by the surrogate equals the observed efficacy" is

$$\left(\frac{\phi^1 + \phi^0}{2} \right) = \frac{\pi_{11}^1 \phi^1 + \pi_{11}^0 \phi^0 + \pi_{21}^1 (1 - \phi^1) + \pi_{21}^0 (1 - \phi^0)}{\pi_{12}^1 \phi^1 + \pi_{12}^0 \phi^0 + \pi_{22}^1 (1 - \phi^1) + \pi_{22}^0 (1 - \phi^0)}. \quad (7)$$

The LHS is one minus the predicted vaccine efficacy based on the measurement S , under the simple model that vaccine takers (with $Z = 1$ and $S = 1$) are 100% protected and vaccine non-takers (with $Z = 1$ and $S = 0$) are completely unprotected. The RHS, one minus the "observed efficacy," is $\Pr(Y(1) = 1)/\Pr(Y(0) = 1) = \Pr(Y = 1|Z = 1)/\Pr(Y = 1|Z = 0)$.

We now show that when (7) holds, the criteria (4) and (5) for a perfect principal surrogate imply the criterion (6) for a perfect statistical surrogate if X is independent of at least one of S or Y , but not if X depends on both. Assuming the three conditions (4), (5), and (7), straightforward algebra shows that (6) holds if and only if

$$\pi_{11}^1 (1 - \phi^1) + \pi_{11}^0 (1 - \phi^0) = \pi_{22}^1 (1 - \phi^1) + \pi_{22}^0 (1 - \phi^0). \quad (8)$$

If X is independent of S , then $\phi^x = \phi$ for $x = 0, 1$, and (8) clearly holds; and if X is independent of Y , then $\pi_{ij}^x = \pi_{ij}$ for $i = 1, 2, j = 1, 2$, and $x = 0, 1$, and (8) clearly holds. If X is correlated with both S and Y , however, the equation does not simplify, and (8) may not hold.

Vice versa, it is straightforward to show that if S is a perfect statistical surrogate (i.e., (6) holds) and (7) holds, then the average causal sufficiency condition for a perfect principal surrogate holds (i.e., $VE(1, 0) = 1$) if X is independent of at least one of S and Y , but not otherwise. This illustrates the non-equivalency of the principal and statistical surrogate frameworks under the prediction criterion (7) in the presence of an unmeasured simultaneous predictor of S and Y . Of course, if X were measured, then the above analysis could be stratified within levels of X , and Dunning's equivalency results would be recovered.

Lastly, Dr. Dunning points to some useful statistical methods for evaluating a specific statistical SoP. Following from the above discussion, we stress that these methods only provide a valid evaluation of an immune correlate at tier 2 if the key assumption about no accounted for simultaneous baseline predictors holds. If this assumption is dubious, then the cited methods more appropriately belong to the class of tools for addressing correlates of risk at tier 1. Furthermore, for the cited methods to be used effectively for evaluating an immune correlate at tier 2, it is critically important to include all baseline covariates that plausibly predict both S and Y . Uncertainty about the key assumption motivates conduct of a sensitivity analysis, which may include use of the complementary approach that evaluates a specific principal SoP.

In conclusion, historically, statistical surrogate based methods for evaluating a SoP have sometimes been applied without controlling for all baseline covariates that plausibly could predict both the immune response and the clinical endpoint. Given the necessity of the key assumption for these methods to provide meaningful evaluation of a specific SoP, we stress that it should be standard practice to do so, just as in observational studies it is standard practice to include all potential confounders. Furthermore, the key assumption suggests that the overarching research agenda for identifying immunological SoPs should include investigations into baseline predictors of immunogenicity and of clinical endpoints. This could motivate, for example, expanded research into genomic and proteomic predictors of vaccine immunogenicity. The fruit of such research would also feed into the methodology for evaluating specific principal SoPs.

The statistical and principal surrogate frameworks are complementary, because they are based on different assumptions, and one or the other may be more favorable for certain settings. For example the statistical surrogate framework may be especially useful when the immune response has broad variability in the control arm and there is good knowledge and data about what are the baseline simultaneous predictors, whereas the principal surrogate framework may be especially useful when the immune response has no variability in the control arm, and/or when an innovative trial design is used that crosses control subjects over to the vaccine arm in order to measure the immune response to vaccine. Where possible both approaches should be applied, as a comparison of the results and the plausibility of the assumptions may lead to insights beyond those achieved by either approach alone.

Acknowledgments

This work is supported by NIH grant 2 R01 AI54165-04.

References

1. Sadoff JC, Wittes J. Correlates, surrogates, and vaccines. *The Journal of Infectious Diseases*. 2007; 196:1279–1281. [PubMed: 17922389]
2. Qin L, Gilbert PB, Corey L, McElrath J, Self SG. A framework for assessing immunological correlates of protection in vaccine trials. *The Journal of Infectious Diseases*. 2007; 196:1304–1312. [PubMed: 17922394]
3. Gilbert PB, Hudgens MG. Evaluating candidate principal surrogate endpoints. *Biometrics*. in press.
4. Salk JE, Menke WJ Jr, Francis T Jr. A clinical, epidemiological and immunological evaluation of vaccination against epidemic influenza. *American Journal of Hygiene*. 1943; 42:57–93.
5. Pearl, J. *Causality: Models, Reasoning, and Inference*. Cambridge University Press; 2000.