Surrogate Endpoints and Causal Methods

By Ariel A. Abad

Keywords: *surrogate endpoints, causal inference, individual causal effects, information theory*

Abstract: The increasing cost of drug development has raised the demand for the use of surrogate endpoints in the evaluation of new drugs in clinical trials. However, over the years it has become clear that surrogate endpoints need to be statistically evaluated and deemed valid, before they can be used as substitutes of “true” endpoints in clinical studies. Nowadays, two paradigms, based on causal inference and meta-analysis, dominate the scene. In this article, we describe the general theoretical framework and some of the methodologies introduced in the causal inference paradigm.

1 The Evaluation of Surrogate Endpoints Based on Causal Inference

Two main paradigms are prevalent in the evaluation of surrogate endpoints, the so-called causal-inference and meta-analytic paradigms (*see* **Surrogate Endpoints**). In the former, individual causal treatment effects are often the primary building block for the analysis in a single-trial setting (STS), whereas in the latter, expected causal treatment effects, that is, the averages of the individual causal effects across all patients within the trial populations, are used to carry out the validation exercise (*see* **Surrogate Markers**).

In the causal inference framework, it is assumed that information on both the surrogate S and true endpoint T is available from a single clinical trial. In this setting, the inclusion and exclusion criteria of the trial characterized a unique, well-defined, and fixed population in which the validation exercise is framed. The fundamental quantities used for the assessment of surrogacy are the so-called individual causal treatment effects ΔT and ΔS . In the Neyman–Rubin “potential outcomes framework” of causality, an individual causal treatment effect ΔT and/or ΔS is defined for each individual patient in terms of two “potential outcomes.” Each patient has one outcome that would manifest if the patient were exposed to the treatment and another outcome that would manifest if he were exposed to the control. The “individual causal treatment effect” (ΔT and/or ΔS) is the difference between these two potential outcomes. However, this individual-level causal treatment effects are often unobservable because individual patients typically receive the treatment or the control but not both. This is the so-called fundamental problem of causal inference and constitutes one of the biggest challenges one faces when using causal inference models in this context, essentially, the estimands of interest are not identifiable from the data^[1].

Katholieke Universiteit Leuven, Leuven, Belgium

Wiley StatsRef: Statistics Reference Online, © 2014–2017 John Wiley & Sons, Ltd.
This article is © 2017 John Wiley & Sons, Ltd.
DOI: 10.1002/9781118445112.stat08027

Validation methods for surrogate endpoints have been classified into the so-called causal effects (CEs) and causal association (CA) frameworks^[2]. In the former, the causal treatment effect on the true endpoint is assessed when the candidate surrogate is fixed at different values, whereas in the latter, the association between the causal treatment effects on the surrogate and true endpoint is the focus of the analysis. For binary true endpoints, several methods have been developed within both the CA and CE frameworks^[3–5]. In fact, in a CE framework, the so-called causal effect predictiveness (CEP) surface, has been proposed for categorical and continuous surrogate endpoints^[3]. The CEP quantifies how well causal treatment effects on the surrogate predict causal treatment effects on the clinical endpoint. In order to achieve identification of the CEP, these authors assumed that the surrogate endpoint is constant in the control group, and that the value of the surrogate potential outcome for the new treatment in the control group could be predicted using baseline covariates. In addition, within a CA framework, a Bayesian modeling approach was proposed to estimate the associative proportion for binary surrogate endpoints, under the assumption of monotonicity^[4]. The previous work was extended to accommodate missing data as well^[5].

Working in a CA framework, a new metric of surrogacy, the so-called individual causal association (ICA), has been introduced^[6,7]. When both endpoints are continuous and normally distributed the ICA is quantified using the Pearson correlation coefficient $\rho_{\Delta} = \text{corr}(\Delta T, \Delta S)$ ^[6]. An extension of the ICA based on information theory, more precisely, based on the mutual information and the entropies of the individual causal treatment effects, has been developed for binary outcomes^[7]. The ICA always lies in the unit interval and has a simple and appealing interpretation in terms of uncertainty reduction. It takes value 1 if there is a deterministic relationship between ΔT and ΔS and, therefore, ΔS predicts ΔT without error, that is, knowing what the treatment does to the surrogate brings full information about what it will do to the true endpoint. In addition, when the ICA equals zero, both treatment effects are independent, and knowing what the treatment does to the surrogate brings no information whatsoever about what it will do to the true endpoint. The identifiability issues are addressed using a two-step Monte Carlo procedure. The procedure determines, in a first step, the region of the parametric space of the distribution of the potential outcomes that is compatible with the data at hand (the so-called Γ_D region) and, subsequently, in a second step, a Monte Carlo approach is implemented to study the behavior of the ICA on this region. Essentially, in the Monte Carlo step, the value of the ICA is computed for several points uniformly sampled on Γ_D and later summarized using frequency distributions and histograms. These frequency distributions and histograms describe the behavior of the ICA across all scenarios compatible with the data. Moreover, within a CE framework, the so-called surrogate predictive function (SPF) is defined as the conditional distribution of the individual causal treatment effect on the true endpoint, given the individual causal treatment effect on the surrogate^[8]. It has been shown that the ICA and the SPF provide complementary and useful information for the evaluation exercise^[8]. The SPF suffers from the same identifiability issues as the ICA and the authors tackled the problem using the Monte Carlo approach previously mentioned. The relationship between the causal inference and meta-analytic paradigms has been studied in Ref. 6.

The lack of user-friendly software is a challenge frequently encountered by data analysts when working with surrogate endpoints. The problem is compounded by the fact that the most sound approaches are oftentimes rather complex and difficult to implement. The R package *Surrogate*, freely available at <http://cran.r-project.org/web/packages/Surrogate/>, implements the most modern and sophisticated methodologies available for that purpose. In the package, several examples and data sets are included with a user-friendly help file^[9].

Related Articles

07135
03732
03920

Causal Inference; Causal Diagrams; Causality.

Wiley StatsRef: Statistics Reference Online, © 2014–2017 John Wiley & Sons, Ltd.
This article is © 2017 John Wiley & Sons, Ltd.
DOI: 10.1002/9781118445112.stat08027

References

- [1] Holland, P.W. (1986) Statistics and causal inference. *J. Am. Stat. Assoc.* **81**, 945–960.
- [2] Joffe, M.M. and Greene, T. (2009) Related causal frameworks for surrogate outcomes. *Biometrics* **65** (2), 530–538.
- [3] Gilbert, P.B. and Hudgens, M.G. (2008) Evaluating candidate principal surrogate endpoints. *Biometrics* **64**, 1146–1154.
- [4] Li, Y., Taylor, J.M.G., and Elliott, M.R. (2010) A Bayesian approach to surrogacy assessment using principal stratification in clinical trials. *Biometrics* **58**, 21–29.
- [5] Elliott, M.R., Li, Y., and Taylor, J.M.G. (2013) Accommodating missingness when assessing surrogacy via principal stratification. *Clin. Trials* **10**, 363–377.
- [6] Alonso, A., Van der Elst, W., Molenberghs, G., Buyse, M., and Burzykowski, T. (2015) On the relationship between the causal-inference and meta-analytic paradigms for the validation of surrogate endpoints. *Biometrics* **70**, 15–24.
- [7] Alonso, A., Van der Elst, W., Molenberghs, G., Buyse, M., and Burzykowski, T. (2016) An information-theoretic approach for the evaluation of surrogate endpoints based on causal inference. *Biometrics* **72** (3), 669–677.
- [8] Alonso, A. and Van der Elst, W. (2016) Assessing a surrogate predictive value: a causal inference approach. *Stat. Med.* **36** (7), 1083–1098.
- [9] Alonso, A., Bigirimurame, T., Burzykowski, T., Buyse, M., Molenberghs, G., Muchene, L., Perualila, N.J., Shkedy, Z., and Vander Elst, W. (2017) *Applied Surrogate Endpoint Evaluation Methods with SAS and R*, Chapman and Hall/CRC, Boca Raton, FL.