Published in final edited form as:

Clin Trials. 2015 August ; 12(4): 317–322. doi:10.1177/1740774514561046.

Surrogacy Assessment Using Principal Stratification With Multivariate Normal and Gaussian Copula Models

Jeremy M.G. Taylor^a, Anna S.C. Conlon^a, and Michael R. Elliott^{a,b}

Jeremy M.G. Taylor: jmgmt@umich.edu

^aDepartment of Biostatistics, School of Public Health, University of Michigan, 1420 Washington Heights, Ann Arbor, MI, USA

^bSurvey Methodology Program, Institute for Social Research, Ann Arbor, MI, USA

Abstract

Background—The validation of intermediate markers as surrogate markers (S) for the true outcome of interest (T) in clinical trials offers the possibility for trials to be run more quickly and cheaply by using the surrogate endpoint in place of the true endpoint.

Purpose—Working within a principal stratification framework, we propose causal quantities to evaluate surrogacy using a Gaussian copula model for an ordinal surrogate and time-to-event final outcome. The methods are applied to data from four colorectal cancer clinical trials where S is tumor response and T is overall survival.

Methods—For the Gaussian copula model, a Bayesian estimation strategy is used and, as some parameters are not identifiable from the data, we explore the use of informative priors that are consistent with reasonable assumptions in the surrogate marker setting to aid in estimation.

Results—While there is some bias in the estimation of the surrogacy quantities of interest, the estimation procedure does reasonably well at distinguishing between poor and good surrogate markers.

Limitations—Some of the parameters of the proposed model are not identifiable from the data, and therefore assumptions must be made in order to aid in their estimation.

Conclusions—The proposed quantities can be used in combination to provide evidence about the validity of S as a surrogate marker for T .

Keywords

Bayesian estimation; Causal inference; Gaussian copula; Potential outcomes; Surrogate endpoint

1 Introduction

There is much interest in the use of intermediate outcome variables as surrogate endpoints (S) for the true outcome of interest (T) in randomized clinical trials, as they may allow trials to be run more quickly and inexpensively by basing the evaluation of the treatment effect on the earlier measured surrogate. Many previous methods of surrogate validation rely on models for the conditional distribution of T given the treatment (Z) and S . However, S is a post-randomization variable, and unobserved, simultaneous predictors of S and T may exist.

When such confounders exist, these methods will not have a causal interpretation. Therefore, there has been much recent work in the area of surrogacy assessment under the principal surrogacy [1] approach, which looks at the distribution of the potential outcomes of T conditional on principal strata based on the potential outcomes of S , which are pre-randomization variables [2–4]. Treatment effect estimates that condition on these principal strata are therefore causal estimates. In this paper, using the principal surrogacy framework we describe surrogate validation measures for multivariate normal surrogate and outcome data, with extensions to non-normally distributed data through the use of a Gaussian copula model.

We consider the entire joint distribution of the potential outcomes of S and the potential outcomes of T , denoted by $S(Z)$ and $T(Z)$ for $Z = 0; 1$. We first consider the scenario where both $S(Z)$ and $T(Z)$ are continuous and their joint distribution is multivariate normal [5]. We then relax the multivariate normality assumption and consider the scenario of an ordinal categorical surrogate marker and a censored time-to-event true outcome [6]. Once parameter estimates from the joint distribution of $(S(0), S(1), T(0), T(1))$ are obtained, we examine various causal quantities that may aid in the assessment of S as a surrogate marker for T . As only two of the four potential outcomes are observed for each subject $((S(0), T(0))$ for those in the control arm and $(S(1), T(1))$ for those in the treatment arm), the models are not fully identifiable from the data. We therefore use a Bayesian estimation strategy and propose some prior distributions that are consistent with reasonable assumptions in the surrogacy setting to aid in estimation. We apply the method to data from an advanced colorectal cancer clinical trial where tumor response is assessed as a surrogate for overall survival.

2 Measures of Surrogacy

In the setting of multivariate normal surrogate and outcome variables, we previously developed a model [5] for the joint distribution of the potential outcomes of S and T , given by:

$$\begin{pmatrix} S(0) \\ S(1) \\ T(0) \\ T(1) \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_{S_0} \\ \mu_{S_1} \\ \mu_{T_0} \\ \mu_{T_1} \end{pmatrix}, \begin{pmatrix} \sigma_{S_0}^2 & \rho_s \sigma_{S_0} \sigma_{S_1} & \rho_{00} \sigma_{S_0} \sigma_{T_0} & \rho_{01} \sigma_{S_0} \sigma_{T_1} \\ & \sigma_{S_1}^2 & \rho_{10} \sigma_{S_1} \sigma_{T_0} & \rho_{11} \sigma_{S_1} \sigma_{T_1} \\ & & \sigma_{T_0}^2 & \rho_t \sigma_{T_0} \sigma_{T_1} \\ & & & \sigma_{T_1}^2 \end{pmatrix} \right).$$

We focus on the conditional distribution of $[T(0); T(1) | S(0); S(1)]$ and derive quantities from this distribution to determine surrogacy. Specifically, we can consider $E(T(1) | T(0) | S(1) | S(0) = s)$, which is linear in s and equal to $\gamma_0 + \gamma_1 s$, where

$$\gamma_0 = (\mu_{T_1} - \mu_{T_0}) - \left(\frac{\rho_{11} \sigma_{S_1} \sigma_{T_1} - \rho_{10} \sigma_{S_1} \sigma_{T_0} - \rho_{01} \sigma_{S_0} \sigma_{T_1} + \rho_{00} \sigma_{S_0} \sigma_{T_0}}{\sigma_{S_0}^2 + \sigma_{S_1}^2 - 2\rho_s \sigma_{S_0} \sigma_{S_1}} \right) (\mu_{S_1} - \mu_{S_0})$$

$$\gamma_1 = \left(\frac{\rho_{11} \sigma_{S_1} \sigma_{T_1} - \rho_{10} \sigma_{S_1} \sigma_{T_0} - \rho_{01} \sigma_{S_0} \sigma_{T_1} + \rho_{00} \sigma_{S_0} \sigma_{T_0}}{\sigma_{S_0}^2 + \sigma_{S_1}^2 - 2\rho_s \sigma_{S_0} \sigma_{S_1}} \right).$$

The value of γ_0 is then is a measure of the “associative effect” [1], which should be near zero for a valid surrogate, indicating that the average treatment effect on T is near zero when there is no treatment effect on S , and γ_1 is a measure of the “dissociative effect” which should be non-zero for a valid surrogate, indicating that there is an expected treatment effect on T when there is a treatment effect on S . We can also consider the entire curve of $E(T(1)|S(1)|S(0)=s)$ [3], which provides estimates of the expected treatment effect within principal strata of S that may be clinically interesting. For a good surrogate the curve should pass through the origin, be monotonic and be much different from zero at large values of $|s|$.

In the setting of an ordinal categorical S , and a time-to-event T , a Gaussian copula model is used to model the joint distribution of the two potential outcomes for each of S and T [6]. We assume that $S(0)$ and $S(1)$ arise from latent underlying standard normal Gaussian random variables, denoted by $\tilde{S}(0)$ and $\tilde{S}(1)$, respectively, such that:

$$S(Z) = \begin{cases} 1_Z, & \text{if } \tilde{S}(Z) \in (-\infty, \alpha_{1Z}) \\ 2_Z, & \text{if } \tilde{S}(Z) \in (\alpha_{1Z}, \alpha_{2Z}) \\ \vdots & \\ C_Z, & \text{if } \tilde{S}(Z) \in (\alpha_{(C-1)Z}, \infty) \end{cases}$$

where $\alpha_{1Z} < \alpha_{2Z} < \dots < \alpha_{(C-1)Z}$ are unknown cut points with $\alpha_{0Z} = -\infty$, $\alpha_{CZ} = \infty$ and let $\Phi^{-1}\{P(S(Z) = k_Z)\} = \alpha_{kZ}$. We assume a Weibull distribution for T , with hazard function

$\lambda(T(Z)) = \left(\frac{\gamma_{T_Z}}{\lambda_{T_Z}}\right) \left(\frac{T(Z)}{\lambda_{T_Z}}\right)^{\gamma_{T_Z}-1}$. We denote the marginal cumulative distributions of $\tilde{S}(0)$, $\tilde{S}(1)$, $T(0)$ and $T(1)$ by $F_{\tilde{S}(0)}$, $F_{\tilde{S}(1)}$, $F_{T(0)}$ and $F_{T(1)}$, respectively, and assume that their joint cumulative distribution is generated by the Gaussian copula function:

$F_{\tilde{\mathbf{y}}}(\tilde{\mathbf{y}}) = \Phi_4\{\Phi_1^{-1}(F_{\tilde{S}(0)}(\tilde{S}(0))), \Phi_1^{-1}(F_{\tilde{S}(1)}(\tilde{S}(1))), \Phi_1^{-1}(F_{T(0)}(T(0))), \Phi_1^{-1}(F_{T(1)}(T(1))) | \Gamma\}$ where $\tilde{\mathbf{y}} = (S(0), S(1), T(0) \text{ and } T(1))$ and Φ_P is the standard P-variate normal distribution,

$$\Gamma = \begin{pmatrix} 1 & \rho_s & \rho_{00} & \rho_{01} \\ & 1 & \rho_{10} & \rho_{11} \\ & & 1 & \rho_t \\ & & & 1 \end{pmatrix}$$

with correlation matrix

In this setting, we can again consider measures of the associative and dissociative effect which can be measured by $E(\log(T(1)=T(0))|S(1)|S(0)=s)$. This conditional expectation provides an estimate of the dissociative effect when $s = 0$ and a measure of the associative effect at each value of s when $s \neq 0$. We can also consider the entire curve of $E(\log(T(1)=T(0))|S(1)|S(0)=s)$ versus s .

3 Estimation

In both the multivariate normal model and the Gaussian copula model there are four correlation parameters that are unidentifiable from the observed data ($\rho_s, \rho_t, \rho_{10}, \rho_{01}$). Therefore, a Bayesian approach was developed [5,6] with reasonable priors in the surrogate evaluation setting used to aid in estimation. Unobserved potential outcomes are treated as missing data and imputed from the appropriate posterior distribution at each iteration of the

Markov chain. Non-informative priors are placed on the parameters that define the mean and variance in the multivariate normal model, and on λ_{TZ} , γ_{TZ} and the α_{KZ} 's in the Gaussian copula model, all of which are fully identified parameters. Marginal priors are placed on each of the correlation parameters of the covariance matrix and the use of different prior assumptions are explored. For each of these there is the additional assumption that the correlation matrix must be positive definite. The priors considered for the correlation parameters include:

- a. Jointly uniform prior such that for each of the six correlations $p(\rho) \sim \text{Unif}(-1, 1)$
- b. Jointly uniform prior such that for each of the six correlations $p(\rho) \sim \text{Unif}(0, 1)$
- c. Uniform over the region where all ρ 's ≥ 0 , $\rho_{01} < \min(\rho_{00}, \rho_{11}, \rho_s, \rho_t)$ and $\rho_{10} < \min(\rho_{00}, \rho_{11}, \rho_s, \rho_t)$
- d. Beta priors such that:
 - $p(\rho_{11}) \sim \text{Unif}(0, 1)$
 - $p(\rho_{00}) \sim \text{Unif}(0, 1)$
 - $p(\rho_{10})$ and $p(\rho_{01}) \sim \text{Beta}(3a_0, 3 - 3a_0)$ such that $P(\rho_{01} \geq \rho_{10} | \hat{\rho}_{00}, \hat{\rho}_{11}) = 0.80$
 - $p(\rho_s)$ and $p(\rho_t) \sim \text{Beta}(3a_1, 3 - 3a_1)$ such that $P(\rho_s \geq \rho_t | E[\rho_{10}]) = 0.80$

where $\hat{\rho}_{00}$, and $\hat{\rho}_{11}$, are the Pearson correlation coefficients when S and T are normally distributed, and the polyserial correlation coefficients when S is ordinal and T is a time-to-event. These are estimable from the observed data. $E[\rho_{10}]$ is the expected value under the $\text{Beta}(3a_0, 3 - 3a_0)$ distribution. In simulation studies [5,6] we found that prior d performed the best in terms of bias and mean squared error for both the multivariate normal model and the Gaussian copula model.

4 Applications

To illustrate the use of the Gaussian copula model, we consider the use of the ordinal variable “tumor response” as a surrogate for overall survival in advanced colorectal cancer. Tumor response and overall survival are common endpoints of interest in cancer clinical trials, and there is a large literature on the use of tumor response as a surrogate marker for overall survival [7–9]. We apply our estimation method to data from four separate meta-analyses in advanced colorectal cancer [10]. Each of the four studies is a combination of separate randomized trials. Study 1 consists of eight trials that compare uracilpyrimidines given as a bolus intravenous injection (5FU bolus) to 5FU modulated by methotrexate with a total of 1,168 subjects [11], Study 2 consists of four trials that compare 5FU bolus to hepatic arterial infusion of 5FU with a total of 244 patients [12], Study 3 consists of 9 trials that compare 5FU bolus to 5FU modulated by leucovorin with a total of 1,378 patients [13] and Study 4 consists of six trials that compare 5FU bolus to 5FU given in continuous infusion with a total of 1,216 patients [14]. Within each study we pool the data from the separate trials. The results from fitting the Gaussian copula model to Study 4 have been previously presented [6]. The inclusion of these additional studies allows an assessment of whether the

causal surrogacy quantity, $E(\log(T(1)=T(0))|S(1); S(0))$ is similar across different treatment comparisons.

Patients in each trial were followed with tumor response and survival time recorded. Tumor response was defined by one of four categories: complete response (CR), partial response (PR), stable disease (SD) and progressive disease (PD). In our analysis, the true endpoint T is survival time, defined as the time from randomization to death from any cause and the surrogate end point S is tumor response, defined as a categorical variable with $S = 1; 2; 3; 4$ for PD, SD, PR, and CR, respectively. The binary treatment indicator for treatment Z is set to 0 for control treatment and 1 for experimental treatment. Tumor response was measured after approximately three to six months of follow-up in advance of the recorded survival time.

There are modest effects of the treatment on S and T in the 4 studies. The odds ratio for response (PR or CR) in the treatment vs. control arm was 2.05 (95% CI: 1.46, 2.88) for Study 1, 3.89 (95% CI: 2.17, 6.96) for Study 2, 2.37 (95% CI: 1.74, 3.23) for Study 3, and 1.84 (95% CI: 1.35, 2.51) for Study 4. The median survival time was longer for those in the treatment group (10.7 months, 16.3 months, 11.6 months and 12.1 months for Studies 1, 2, 3 and 4, respectively) than for those in the control group (9.1 months, 12.2 months, 11.0 months and 11.3 months for Studies 1, 2, 3 and 4, respectively) for all four studies, with an estimated hazard ratio of 0.88 (95% CI: 0.78, 0.99) in Study 1, 0.82 (95% CI: 0.64, 1.07) in Study 2, 0.98 (95% CI: 0.87, 1.10) in Study 3, and 0.89 (95% CI: 0.79, 1.00) in Study 4 for the treatment vs. control group. Figure 1 (a) and Figure 1 (b) provide a plots of $E[\log(T(1)=T(0))|S(1) S(0) = s]$ vs. $S(1) S(0) = s$ for each of the four studies using prior a and using prior d, respectively. Figure 2 provides plots of $E[\log(T(1)=T(0))|S(1) S(0) = s]$ vs. $S(1) S(0) = s$ with the 95% credible interval for the curve for each of the four studies using prior d.

Under both prior a and prior d Figure 1 shows that $E[\log(T(1)=T(0)) | S(1) S(0) = 0]$ is approximately 0 for all four studies, indicative of a good principal surrogate, with $E[\log(T(1)=T(0)) | S(1) S(0) = s]$ increasing in s . We would therefore conclude that tumor response appears to be a moderately good principal surrogate for overall survival, and the fact that the four estimated curves are quite similar suggests that this is a quantity that may generalize from one treatment comparison to the next. Figure 2 shows that there is very little uncertainty around $S(1) S(0) = 0$, however the credible intervals become much larger at either end of the curve when $S(1) S(0) = 3$ and $S(1) S(0) = 3$. Table 1 provides the four by four table of $E[\log(T(1)=T(0)) | S(0); S(1)]$ for every pair of values for $S(0)$ and $S(1)$ for each of the four studies from the estimation done using Beta priors, as well as the average fraction of the population within each of the cells across all of the iterations of the Markov chain. If S is a good principal surrogate, then $E[\log(T(1)=T(0)) | S(0); S(1)]$ in the 16 cells of this table should be similar across these four studies, even though the estimated proportion of data in each cell may differ. The estimates and uncertainty intervals don't contradict the hypothesis that $E[\log(T(1)=T(0)) | S(0); S(1)]$ is similar across studies. If $E[\log(T(1)=T(0)) | S(0); S(1)]$ is similar across studies, then it may be possible to use this information to predict treatment effects on overall survival in a future trial based only on the surrogate outcome if there is a

way to predict the missing counterfactual values of S , such as by using baseline covariate information.

5 Discussion

We have developed methods for surrogacy assessment within the principal stratification framework when the joint distribution of the potential outcomes of S and T is multivariate normal, with an extension to non-normally distributed data through the use of a Gaussian copula model. Our simulation results [5,6] suggest that the estimation procedure is able to distinguish valid principal surrogates from invalid ones. As some parameters of the proposed model are not identifiable from the data, certain assumptions must be made in order to aid in their estimation. A Bayesian estimation strategy is used, which allows the use of context specific prior distributions on the unidentified parameters to be explored. The priors that were placed on the unidentified parameters seem reasonable in the setting that we are considering. The use of other priors or other context specific assumptions about unidentified parameters or the conditional relationships between the potential outcomes could be made to accommodate the specific setting of the trial of interest [4].

Acknowledgments

This paper is a summary of a conference presentation and includes research, some of which has previously been published in *Biostatistics* [5] and *Statistical Methods in Medical Research* [6]. The authors are grateful to Marc Buyse for providing the data and to the Meta-Analysis Group in Cancer for permission to use their individual patient data.

Funding

This research was supported by NIH grants CA 129102 and T32 CA 083654.

References

1. Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics*. 2002; 58:21–29. [PubMed: 11890317]
2. Li Y, Taylor JMG, Elliott MR. A Bayesian approach to surrogacy assessment using principal stratification in clinical trials. *Biometrics*. 2010; 66:523–531. [PubMed: 19673864]
3. Gilbert PB, Hudgens MG. Evaluating candidate principal surrogate endpoints. *Biometrics*. 2008; 64:1146–1154. [PubMed: 18363776]
4. Zigler CM, Belin TR. A Bayesian approach to improved estimation of causal effect predictiveness for a principal surrogate endpoint. *Biometrics*. 2012; 68:922–932. [PubMed: 22348277]
5. Conlon ASC, Taylor JMG, Elliott MR. Surrogacy assessment using principal stratification when surrogate and outcome measures are multivariate normal. *Biostatistics*. 2014; 15:266–283. [PubMed: 24285772]
6. Conlon ASC, Taylor JMG, Elliott MR. Surrogacy assessment using principal stratification and a Gaussian copula model. *Statistical Methods in Medical Research*. 2014 NIHMS #607289. 10.1177/0962280214539655
7. Ellenberg SS, Hamilton JM. Surrogate endpoints in clinical trials: cancer. *Statistics in Medicine*. 1989; 8:405–413. [PubMed: 2727464]
8. Torri V, Simon R, Russek-Cohen E, et al. Statistical model to determine the relationship of response and survival in patients with advanced ovarian cancer treated with chemotherapy. *Journal of the National Cancer Institute*. 1992; 84:407–414. [PubMed: 1531682]
9. Buyse M, Piedbois P. On the relationship between response to treatment and survival. *Statistics in Medicine*. 1996; 15:2797–2812. [PubMed: 8981687]

10. Buyse M, Thirion P, Carlson RW, Burzykowski T, Molenberghs G, Piedbois P. for the Meta-Analysis Group in Cancer. Tumour response to first line chemotherapy improves the survival of patients with advanced colorectal cancer. *Lancet*. 2000; 365:373–378. [PubMed: 10972369]
11. Advanced Colorectal Cancer Meta-Analysis Project. Meta-analysis of randomized trials testing the biochemical modulation of 5-fluorouracil by methotrexate in metastatic colorectal cancer. *Journal of Clinical Oncology*. 1994; 12:960–969. [PubMed: 8164048]
12. Meta-analysis Group in Cancer. Reappraisal of hepatic arterial infusion in the treatment of nonresectable liver metastases from colorectal cancer. *Journal of the National Cancer Institute*. 1996; 88:252–258. [PubMed: 8614003]
13. Advanced Colorectal Cancer Meta-Analysis Project. Modulation of 5-fluorouracil by leucovorin in patients with advanced colorectal cancer: evidence in terms of response rate. *Journal of Clinical Oncology*. 1992; 10:896–903. [PubMed: 1534121]
14. Advanced Colorectal Cancer Meta-Analysis Project. Efficacy of intravenous continuous infusion of 5-fluorouracil compared with bolus administration in patients with advanced colorectal cancer. *Journal of Clinical Oncology*. 1998; 16:301–308. [PubMed: 9440757]

Figure 1.

$E[\log(T(1)/T(0)) | S(1) - S(0) = s]$ vs. $S(1) - S(0)$ from 4 studies assessing tumor response (1 =progressive disease, 2 =stable disease, 3 =partial response, 4 =complete response) as a surrogate for overall survival.

Figure 2.

$E[\log(T(1)/T(0)) | S(1) - S(0) = s]$ vs. $S(1) - S(0)$ and 95% credible intervals from 4 studies assessing tumor response (1 = progressive disease, 2 = stable disease, 3 = partial response, 4 = complete response) as a surrogate for overall survival

$E[\log(T(1)=T(0)) | S(0); S(1)]$ for every pair of $S(0)$ and $S(1)$ and population fraction within each cell for colorectal cancer data analyzed using Beta priors

Table 1

Study	S(0)	S(1)			
		1	2	3	4
1	1	0.19 (0.27) 23.9%	0.72 (0.20) 18.6%	0.88 (0.37) 6.0%	0.97 (0.72) 1.1%
2		-0.31 (0.69) 23.8%	0.31 (0.54) 14.2%	0.72 (1.14) 13.2%	1.50 (2.41) 0.2%
3		0.04 (0.20) 42.8%	0.69 (0.40) 11.8%	1.12 (0.68) 5.3%	1.61 (1.30) 0.3%
4		-0.11 (0.10) 35.7%	0.69 (0.11) 12.5%	1.22 (0.19) 4.2%	1.75 (0.44) 0.4%
1	2	-0.72 (0.22) 14.3%	-0.11 (0.13) 16.7%	0.07 (0.33) 7.0%	0.22 (0.59) 1.7%
2		-0.54 (1.74) 5.0%	-0.07 (0.68) 7.9%	0.47 (0.49) 16.7%	2.00 (1.87) 0.3%
3		-0.99 (0.73) 9.5%	-0.28 (0.33) 9.3%	0.28 (0.23) 8.9%	0.96 (0.65) 1.0%
4		-0.57 (0.12) 12.3%	0.09 (0.07) 12.6%	0.61 (0.10) 8.2%	1.20 (0.24) 1.2%
1	3	-1.20 (0.45) 2.5%	-0.53 (0.29) 3.5%	-0.34 (0.41) 1.8%	-0.12 (0.70) 0.5%
2		-1.27 (2.36) 1.2%	-0.34 (1.69) 2.5%	0.27 (0.58) 13.2%	1.16 (1.50) 1.0%
3		-1.69 (1.38) 1.2%	-0.98 (0.90) 2.0%	-0.36 (0.46) 4.0%	0.40 (0.42) 1.0%
4		-0.83 (0.21) 2.0%	-0.30 (0.12) 3.6%	0.17 (0.10) 4.2%	0.69 (0.16) 1.2%
1	4	-1.47 (0.85) 0.6%	-0.91 (0.53) 0.9%	-0.81 (0.63) 0.6%	-0.62 (0.84) 0.3%
2		-1.84 (2.38) 0.02%	-0.99 (2.18) 0.04%	-0.22 (1.85) 0.4%	0.19 (1.46) 0.3%
3		-1.68 (1.76) 0.2%	-1.46 (1.45) 0.4%	-0.92 (0.82) 1.3%	-0.24 (0.55) 0.9%
4		-1.08 (0.46) 0.1%	-0.60 (0.28) 0.4%	-0.17 (0.17) 0.8%	0.36 (0.19) 0.6%