Supplementary Material

Surrogate endpoint analysis: An exercise in extrapolation

Stuart G. Baker, Barnett S. Kramer

This supplementary material illustrates computation of the standard error multiplier for the second example mentioned in the main text, namely $k = 10$ historical trials that investigated treatments for advanced colorectal cancer [1-3]. Let τ_S and τ_T denote pre-specified survival times for determining surrogate and true endpoints, respectively. For trial j and randomization group z, let S_{jz} denote the estimated probability of surviving a surrogate event to time τ_S, and let T_{jz} denote the estimated probability of surviving a true event to time τ_T. In this example S_{jz} is the estimated probability of no cancer progression by 6 months, and T_{jz} is the estimated probability of no mortality from any cause by 12 months. Survival probabilities were estimated using a discrete-time Kaplan-Meier estimate. The first seven trials listed in Supplementary Tables 1 and 2 compare fluouracil and leucovorin with fluouracil; the last three trials compare fluouracil and leucovorin with raltitrexed. Ideally, in other applications, there would be a greater variety of treatments among the historical trials.

In this notation, the surrogate result is $\Delta S_j = S_{j1} - S_{j0}$ and the true result is $\Delta T_j = T_{j1} - T_{j0}$. With estimated survival probabilities computed via Kaplan-Meier, a simple approximation for the standard error of the true result is $se\Delta T_j = \{T_{j1} (1 - T_{j0}) / n_{j0} + T_{j1} (1 - T_{j1}) / n_{j1}\}^{1/2}$, where n_{jz} is the sample size in group z [1]. Supplementary Table 1 shows calculations of surrogate and true results based on published data in supplementary material from Baker *et al.* [1].

For the leave-one-out analysis, each historical trial j is successively removed and the following regression model is fit to data from the remaining k - 1 historical trials indexed by u,

$$\Delta T_u = \beta_j \Delta S_u + \varepsilon_j,$$

where ε_j is random error. To give more weight to larger trials, the regression model is fit using weights of $n_{u0} + n_{u1}$ for each remaining historical trial u. Let b_j denote the estimate of β_j. The model result is $\Delta Q_j = Q_{j1} - Q_{j0}$, where $Q_{jz} = b_j S_{jz}$. The estimated variance of the model result is $var\Delta Q_j = Q_{j1}(1 - Q_{j0}) / n_{j0} + Q_{j1}(1 - Q_{j1}) / n_{j1}$. The extrapolation error for historical trial j is the difference between the true result and the model result, namely $E_j = \Delta T_j - \Delta Q_j$. For historical trial j, the mean and variance of the extrapolation error based on the remaining k-1 trials are

$$meanE_j = \Sigma_{u \neq j} E_u / (k - 1),$$
$$varE_j = \Sigma_{u \neq j} (E_u - meanE_j)^2 / (k - 2).$$

The predicted result is the model result plus the mean of the extrapolation error, $\Delta P_j = \Delta Q_j + meanE_j$. The variance of the predicted result equals the variance of the model result plus the variance of the extrapolation error, $var\Delta P_j = var\Delta Q_j + varE_j$, The standard error of the predicted result is $se\Delta P_j = var\Delta P_j^{1/2}$. Supplementary Table 2 presents some of these calculations. From the above calculations,

standard error multiplier $= \Sigma_j \left(se\Delta P_j / se\Delta T_j \right) / k$.

If the questions relating to extrapolation estimation are satisfactorily answered, the predicted result and its 95% confidence interval would be computed for a new trial. The calculations parallel those given above, but using all *k* historical trials to compute the model result.

References

1. Baker, SG, Sargent DJ, Buyse M, Burzykowski T. Predicting treatment effect from surrogate endpoints and historical trials: an extrapolation involving probabilities of a binary outcome or survival to a specific time. *Biometrics*. 2012;68(1):248-257.
2. Buyse M, Burzykowski T, Carroll K, Michiels S, Sargent D, Miller LL, Elfring GL, Pignon JP, Piedbois P. Progression-free survival is a surrogate for survival in advanced colorectal cancer. *J Clin Oncol.* 2007;25(33):5218-5224.
3. Meta-Analysis Group in Cancer (2004). Modulation of fluorouracil by leucovorin in patients with advanced colorectal cancer: an updated meta-analysis. *J Clin Oncol.* 22;(18):3766-3775.

Trial	Surrogate endpoint		True endpoint		Sample size		Surro-gate result	True result	SE of true result
	Group		Group		Group				
	0	1	0	1	0	1			
	S_{i0}	S_{i1}	T_{i0}	T_{i1}	n_{i0}	n_{i1}	ΔS_j	ΔT_j	$se\Delta T_j$
1	0.346	0.686	0.022	0.223	49	99	0.340	0.201	0.047
2	0.319	0.518	0.348	0.518	69	137	0.199	0.170	0.071
3	0.451	0.511	0.352	0.489	91	94	0.060	0.138	0.072
4	0.356	0.417	0.526	0.526	164	326	0.061	0.000	0.048
5	0.466	0.484	0.522	0.570	178	93	0.018	0.047	0.064
6	0.681	0.783	0.433	0.513	157	152	0.102	0.080	0.057
7	0.440	0.618	0.517	0.662	67	68	0.178	0.145	0.084
8	0.264	0.357	0.431	0.471	246	243	0.093	0.040	0.045
9	0.227	0.397	0.371	0.522	216	206	0.169	0.151	0.048
10	0.316	0.326	0.416	0.413	222	212	0.010	-0.003	0.047

Supplementary Table 1. Computation of surrogate and true results. S_{jz} is the estimated probability of no cancer progression by 6 months, and T_{jz} is the estimated probability of no mortality from any cause by 12 months. The first seven trials compare fluouracil and leucovorin with fluouracil; the last three trials compare fluouracil and leucovorin with raltitrexed.

Trial	Slope of regression model	Model result		Extrapolation error		Predicted result	
		Mean	SE	Mean	SE	Estimate	SE
	b_j	ΔQ_j	$var\Delta Q_j^{½}$	$meanE_j$	$varE_j^{½}$	ΔP_j	$var\Delta P_j^{½}$
1	0.79	0.268	0.081	0.015	0.042	0.284	0.091
2	0.699	0.139	0.065	0.0042	0.049	0.144	0.081
3	0.703	0.042	0.069	-0.003	0.038	0.039	0.079
4	0.75	0.046	0.043	0.0130	0.046	0.058	0.063
5	0.72	0.013	0.061	0.0037	0.049	0.016	0.078
6	0.72	0.073	0.057	0.0069	0.050	0.080	0.076
7	0.715	0.127	0.083	0.0057	0.050	0.133	0.097
8	0.749	0.069	0.038	0.0110	0.048	0.080	0.062
9	0.673	0.114	0.039	0.0035	0.049	0.117	0.063
10	0.724	0.0074	0.041	0.0088	0.05	0.016	0.064

Supplementary Table 2. Computation of predicted results. The first seven trials compare fluouracil and leucovorin with fluouracil; the last three trials compare fluouracil and leucovorin with raltitrexed.