Statistical Methods in Medical Research 2008; 1–4

Editorial

Surrogate endpoints: wishful thinking or reality?

Generally, before a new drug can be accepted for the use in clinical practice, its efficacy and safety needs to be rigorously assessed in a series of clinical trials. This process of testing a new therapy can (and, in fact, does) take many years. One of the reasons is the use of long-term clinical endpoints like clinical progression or survival.

However, recent advances in the understanding of the biological mechanisms of disease development have resulted in the emergence of a large number of potentially effective new agents. There is also increasing public pressure for promising new drugs to receive marketing approval as rapidly as possible, in particular for life threatening diseases such as cancer. For these reasons, there is an urgent need to find ways of shortening the duration of cancer clinical trials.

A possible solution to this problem is to replace the endpoint of interest, the 'true' endpoint, by another one, a 'surrogate' endpoint, which might be measured earlier or more frequently. However, before a surrogate can replace a true endpoint, it should be *validated*. This means that it should be checked whether the use of the surrogate leads to correct conclusions about the effect of the treatment on the true endpoint.

The validation of a candidate surrogate endpoint is not straightforward. Merely establishing a correlation between both endpoints is not sufficient.[1] Formal methods, allowing for validation of surrogate endpoints, have become the subject of intensive research over the past decades.[2] Until recently, the statistical approaches developed for this purpose were based on the definition of a surrogate proposed by Prentice,[3] according to which a surrogate endpoint is 'a response variable for which a test of the null hypothesis of no relationship to the treatment groups under comparison is also a valid test of the corresponding null hypothesis based on the true endpoint.' The methods assumed the availability of data from a single trial.[3–5] These methods suffer from numerous drawbacks: some of them are too stringent to be of practical value, while others are based on non-testable assumptions.[6] To overcome these limitations, 'meta-analytic' validation approaches have recently been developed.[7–10] They are based on an alternate definition, according to which 'a surrogate endpoint is expected to predict clinical benefit (or harm or lack of benefit or harm)'.[11] These methods use large databases from multiple randomized clinical trials and aim at measuring directly the association between the treatment effects on the surrogate and the true endpoint.

At the 33rd International Biometric Conference, which took place on 16–21 July, 2006, in Montreal, a special Topic Contributed Session 'Surrogate Endpoints: Wishful Thinking or Reality?' was devoted to the issue of surrogate endpoint validation. Each speaker was provided with two datasets, containing data from multiple randomized clinical trials in colorectal cancer. The speakers were asked to evaluate, using different

10.1177/0962280208081866

Table 1 Clinical trials included in the early colon cancer (ACCENT) dataset

Study	Accrual period	Treatment*	Number of patients
NSABP C01	1977–1983	MOF	349
		Surgery only*	375
NCCTG-78-48-52	1978–1984	FU + LEV	121
		Surgery only*	126
INT-0035	1984–1987	FU + LEV	457
		Surgery only*	469
NSABP C02	1984–1988	PVI of FU	438
		Surgery only*	458
NSABP C03	1987–1989	FU + LV	519
		MOF*	523
NCCTG-87-46-51	1988–1989	FU + LV	255
		Surgery only*	153
NSABP C04	1989–1990	FU + LEV*	693
		FU + LV	693
		FU + LV + LEV	697
NCCTG-89-46-51	1989–1991	FU + LEV × 1 year*	228
		FU + LEV × 6 months	230
		FU + LEV + LV × 1 year	232
		FU + LEV + LV × 6 months	225
NSABP C05	1991–1994	FU + LV*	1070
		FU + LV + Interferon	1066
NCCTG-91-46-53	1993–1998	High dose LEV + FU/LV	437
		Standard dose LEV + FU/LV*	441

(Abbreviations—NSABP: National Surgical Adjuvant Breast and Bowel Project, NCCTG: North Central Cancer Treatment Group, INT: Intergroup. FU: fluorouracil, LEV: levamisole, LV: leucovorin, MOF: semustine, vincristine, and fluorouracil, PVI: portal vein infusion. Control arm is indicated by *.)

meta-analytic approaches, the validity of progression-free survival (PFS) as a surrogate for overall survival (OS) in colorectal cancer.

The first dataset (ACCENT data) contained data for 10,255 patients included in 10 early colon cancer trials. It is a subset of data analysed by Sargent *et al.*[12] Table 1 presents a short overview of sample sizes and treatments used in each study. The trials accrued patients between 1977 and 1998 (median follow-up 10.4 years). A more detailed description is provided by Sargent *et al.*[12]

The second (MAGIC) dataset included data for 3,089 patients enrolled in 10 advanced colorectal cancer trials. All trials had a 5FU+leucovorin treatment group (Table 2). Seven of them compared 5FU + leucovorin with 5FU alone (1744 patients), while the remaining three compared 5FU + leucovorin with raltitrexed (1345 patients). They accrued patients between 1981 and 1990 (median follow-up 30.4 months). A meta-analysis of trials comparing 5FU+leucovorin with 5FU was previously reported.[13] The other three trials were those carried out for the registration of the new drug tomudex (studies 3, 10, and 12 in Cunningham *et al*).[14]

Table 2 Clinical trials included in the advanced colorectal cancer (MAGIC) dataset

Study	Treatment*	Number of patients
Crema	FU + LV	100
	FU*	50
NCCTG	FU + LV	142
	FU*	70
Siena	FU + LV	94
	FU*	91
EORTC	FU + LV	165
	FU*	166
SWOG	FU + LV	178
	FU*	93
SAKK	FU + LV	152
	FU*	158
HECOG	FU + LV	70
	FU*	68
TCCSG-EU1	FU + LV*	248
	Raltitrexed	247
TCCSG-US	FU + LV*	210
	Raltitrexed	217
TCCSG-EU2	FU + LV*	216
	Raltitrexed	223

(Abbreviations—NCCTG: North Central Cancer Treatment Group, EORTC: European Organization for Research and Treatment of Cancer, SWOG: Southwest Oncology Group, SAKK: Schweizerische Arbeitsgruppe fur Klinische Krebsforschung, HECOG: Hellenic Cooperative Oncology Group, TCCSG: Tomudex Colorectal Cancer Study Group. FU: fluorouracil, LV, leucovorin. Control arm is indicated by *.)

The papers published in this issue of *Statistical Methods in Medical Research* present details of the analyses and conclusions given during the IBC session. They offer an overview of the meta-analytic approaches currently available for the validation of surrogate endpoints. Also, they allow comparisons of the relative merits of the different approaches. Accompanying papers, written by our colleagues clinicians, offer a complementary view on the validation of surrogate endpoints from a clinical practice perspective.

References

1 Baker SG, Kramer BS. A perfect correlate does not make a surrogate. *BioMed Central Medical Reseach Methodology* 2003; 3: 16.

2 Burzykowski T, Molenberghs G, Buyse M. *The Evalutation of Surrogate Endpoints*. Springer, 2005.

3 Prentice RL. Surrogate endpoints in clinical trials: definitions and operational criteria. *Statistics in Medicine* 1989; **8**: 431–40.
4 Freedman LS, Graubard BI, Schatzkin A. Statistical validation of intermediate endpoints for chronic diseases. *Statistics in Medicine* 1992; **11**: 167–78.
5 Buyse M, Molenberghs G. The validation of surrogate endpoints in randomized experiments. *Biometrics* 1998; **54**: 1014–29.
6 Molenberghs G, Buyse M, Geys H, Renard D, Burzykowski T, Alonso A. Statistical challenges in the evaluation of surrogate endpoints in randomized trials. *Controlled Clinical Trials* 2002; **23**: 607–25.
7 Daniels MJ, Hughes MD. Meta-analysis for the evaluation of potential surrogate markers. *Statistics in Medicine* 1997; **16**: 1515–1527.
8 Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H. The validation of surrogate endpoints in meta-analysis of randomized experiments. *Biostatistics* 2000; **1**: 49–67.
9 Gail MH, Pfeiffer R, van Houwelingen HC, Carroll RJ. On meta-analytic assessment of surrogate outcomes. *Biostatistics* 2000; **1**: 231–46.
10 Baker SG. A simple meta-analytic approach for using a binary surrogate endpoint to predict the effect of intervention on true endpoint. *Biostatstics* 2000; **7**: 58–70.
11 Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clinical Pharmacology and Therapy* 2001; **69**: 89–95.
12 Sargent DJ, Wieand HJ, Haller DG, Gray R. Disease-free survival versus overall survival as a primary endpoint for adjuvant colon cancer studies: individual patient data form 20,898 patients on 18 randomized trials. *Journal of Clinical Oncology* 2005; **23**: 8664–70.
13 Meta-Analysis Group In Cancer. Modulation of fluorouracil by leucovorin in patients with advanced colorectal cancer: an updated meta-analysis. *Journal of Clinical Oncology* 2004; **22**: 3766–75.
14 Cunningham D, Zalcberg J, Maroun J, James R, Clarke S, Maughan TS, Vincent M, Schulz J, Gonzalez BM, Facchini T. Efficacy, tolerability and management of raltitrexed (Tomudex^{TM}) monotherapy in patients with advanced colorectal cancer: a review of phase II/III trials. *European Journal of Cancer* 2002; **38**: 478–86.

Tomasz Burzykowski,
Center for Statistics, Hasselt University,
Agoralaan D, B-3590 Diepenbeek, Belgium.
E-mail:tomasz.burzykowski@uhasselt.be