ELSEVIER

Journal of Clinical Epidemiology ■ (2015) ■

Journal of Clinical Epidemiology

REVIEW ARTICLE

Meta-analyses of randomized controlled trials show suboptimal validity of surrogate outcomes for overall survival in advanced colorectal cancer

Oriana Ciani[a,b,*], Marc Buyse[c,d], Ruth Garside[e], Jaime Peters[a], Everardo D. Saad[f], Ken Stein[a], Rod S. Taylor[a]

[a]*Institute of Health Research, University of Exeter Medical School, Veysey Building, Salmon Pool Lane, Exeter EX2 4SG, UK*
[b]*Centre for Research on Health and Social Care Management, Bocconi University, via Rontgen, 1, Milan 20141, Italy*
[c]*Interuniversity Institute for Biostatistics and Statistical Bioinformatics (I-BioStat), Hasselt University, Universitaire Campus, 3590 Diepenbeek, Belgium*
[d]*International Drug Development Institute (IDDI), 185 Alewife Brook Parkway, Suite 410, Cambridge, MA 02138, USA*
[e]*European Centre for Environment and Human Health University of Exeter Medical School Knowledge Spa, Royal Cornwall Hospital, Truro TR1 3HD, UK*
[f]*Dendrix Research, Rua Joaquim Floriano, 72/24, Sao Paulo 04534-000, Brazil*

Accepted 25 February 2015; Published online xxxx

Abstract

Objectives: To quantify and compare the treatment effects on three surrogate end points, progression-free survival (PFS), time to progression (TTP), and tumor response rate (TR) vs. overall survival (OS) based on a meta-analysis of randomized controlled trials (RCTs) of drug interventions in advanced colorectal cancer (aCRC).

Study Design and Setting: We systematically searched for RCTs of pharmacologic therapies in aCRC between 2003 and 2013. Trial characteristics, risk of bias, and outcomes were recorded based on a predefined form. Univariate and multivariate random-effects meta-analyses were used to estimate pooled summary treatment effects. The ratio of hazard ratios (HRs)/odds ratios (ORs) and difference in medians were used to quantify the degree of difference in treatment effects on the surrogate end points and OS. Spearman ρ, surrogate threshold effect (STE), and R^2 were also estimated across predefined trial-level covariates.

Results: We included 101 RCTs. In univariate and multivariate meta-analyses, we found larger treatment effects for the surrogates than for OS. Compared with OS, treatment effects were on average 13% higher when HRs were measured and 3% to 45% higher when ORs were considered; differences in median PFS/TTP were higher than on OS by an average of 0.5 month. Spearman ρ ranged from 0.39 to 0.80, mean R^2 from 0.06 to 0.65, and STE was 0.8 for HR_{PFS}, 0.64 for HR_{TTP}, or 0.28 for OR_{TR}. The stratified analyses revealed high variability across all strata.

Conclusion: None of the end points in this study were found to achieve the level of evidence (ie, mean $R^2_{trial} > 0.60$) that has been set to select high or excellent correlation levels by common surrogate evaluation tools. Previous surrogacy relationships observed between PFS and TTP vs. OS in selected settings may not apply across other classes or lines of therapy.

Keywords: Surrogate outcome; Colorectal cancer; PFS; TTP; Tumor response; Health technology assessment

1. Introduction

Surrogate end points have been defined as biomarkers or intermediate outcomes that can substitute for a final patient-relevant end point to successfully measure the effect of health interventions [1]. In colorectal cancer, the second commonest cause of cancer-related mortality in high-income countries [2], predictive end points for overall survival (OS) are needed to accelerate the availability of promising new therapies for patients. A number of surrogate end points for OS in clinical oncology trials have been proposed, including progression-free survival (PFS), time to progression (TTP), and tumor response rate (TR) [3–5]. However, to ensure that these surrogate end points provide the same answer as the final end point (OS) about the experimental therapy, they should undergo a process of surrogate validation [6]. Several authors have dealt with the validation of PFS [7–12], TTP [8,10,13], or TR [13,14] as surrogate end points for OS in advanced colorectal cancer

Conflict of interest: O.C. received a PhD scholarship from the Peninsula College of Medicine and Dentistry. Other coauthors have no conflict of interest to declare.

* Corresponding author. Evidence Synthesis & Modelling for Health Improvement, Institute of Health Research, University of Exeter Medical School, Veysey Building, Salmon Pool Lane, Exeter EX2 4SG, UK. Tel.: 44 7575918306; fax: +39 02 5836.2598.

E-mail address: O.Ciani@exeter.ac.uk (O. Ciani).

http://dx.doi.org/10.1016/j.jclinepi.2015.02.016
0895-4356/

What is new?

- The meta-analyses showed that treatment effect sizes were always larger for the surrogate end points than for overall survival (OS). The stratified analyses revealed high variability across all strata.
- Progression-free survival (PFS), time to progression (TTP), and tumor response rate (TR) have been proposed as surrogate for OS in advanced colorectal cancer (aCRC); however, previous surrogacy relationship observed in selected aCRC therapies may not directly apply across other classes or lines of therapy.
- None of the end points in this study were found to achieve the level of evidence that has been set to select high or excellent correlation levels by common surrogate evaluation tools. Where PFS and TTP are deemed acceptable surrogates for OS, policy makers still need to consider that the anticipated treatment effect on OS is likely to be smaller than that observed on the surrogate measure when weighing up the evidence in their licensing and coverage decisions. TR should not be used as a surrogate end point for OS when evaluating the efficacy of drug interventions in aCRC.

(aCRC) over the last decade. Although most of the studies are of high quality, some are not based on systematic review of the available evidence, either because they were based on opportunistically available individual-patient data (IPD) [7,9,11,12,14] or focused on subgroups of trials and therapies [8,13] and did not, therefore, provide a comprehensive examination of the issue. The present study seeks to overcome these limitations by systematically looking at all available randomized controlled trials (RCTs), across drug classes and lines of therapy, and considering different approaches to surrogate validation, with the primary aim of quantifying and comparing treatment effects on surrogates and on OS.

2. Methods

We conducted and reported this systematic review in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement [15].

2.1. Data sources and selection strategy

We searched the following databases from 2003 to January 31, 2013: MEDLINE, EMBASE (via OVID), and the Cochrane Central Register of Controlled Trials. A copy of the bibliographic searches is provided in the Supplementary Material at www.jclinepi.com. We limited our searches to the last 10 years of drug interventions in metastatic colorectal cancer to limit the heterogeneity in our sample and, at the same time, to reflect current clinical practice in most developed countries. We checked citations in identified studies and systematic reviews already known to the authors [16] as additional sources of potentially eligible trials.

Trials were included if they were RCTs in advanced or metastatic colorectal cancer assessing a pharmacologic therapy against either a placebo or other drug therapy. Trials had to report OS and either PFS or TTP or TR. We excluded adjuvant setting trials and trials assessing radiotherapy, supportive-care drugs, other nonantineoplastic drugs, nondrug treatments, and trials that were stopped early, with accrual rate less than 70% of the target sample size. When multiple publications of the same RCT were available, only the most recent one reporting both surrogate and final end points was included. Titles and abstracts were screened independently by two reviewers, and disagreements were resolved by full-text retrieval and, when necessary, involvement of a third reviewer.

2.2. Data extraction

One reviewer extracted the data using a standardized form, and a second reviewer independently checked the extraction. Information collected included: general characteristics of the trial (ie, study design, sample size), patient characteristics (ie, median age, performance status), treatments under comparison, risk of bias assessment (using the Cochrane Collaboration tool [17]), and treatment effects on OS and PFS, TTP, or TR. In multiarm trials, all available between-arm comparisons were recorded. OS was defined as the time from randomization to death from any cause, with patients censored when they are last seen alive or when they are lost to follow-up [18]; PFS was defined as the time between randomization and tumor progression (however defined) or death from any cause; and TTP as the time between randomization and tumor progression (however defined), with censoring of patients who died without prior documentation of progression. Tumor response is based on objective tumor measurements by imaging methods that allow the classification of patients with a complete or partial confirmed best response as responders. Responses are usually determined according to the Response Evaluation Criteria in Solid Tumors guidelines [19] or the World Health Organization recommendations [20].

For OS, PFS, and TTP, the hazard ratio (HR) and median survival time, together with the 95% confidence intervals (CIs) for each arm, were recorded whenever available. The number of events (ie, deaths or tumor progressions or tumor responses) were also recorded to estimate odds ratios (ORs).

2.3. Statistical analyses

We derived the sample size for this present study based on a previous publication comparing the treatment effects