Progression-free survival as surrogate and as true end point: insights from the breast and colorectal cancer literature

E. D. Saad1*, A. Katz2, P. M. Hoff2 & M. Buyse3
1Dendrix Research; 2Oncology Center, Hospital Sírio-Libanês, Sao Paulo, Brazil and 3International Drug Development Institute, Louvain-la-Neuve, Belgium

Received 28 July 2009; accepted 14 October 2009

Significant achievements in the systemic treatment of both advanced breast cancer and advanced colorectal cancer over the past 10 years have led to a growing number of drugs, combinations, and sequences to be tested. The choice of surrogate and true end points has become a critical issue and one that is currently the subject of much debate. Many recent randomized trials in solid tumor oncology have used progression-free survival (PFS) as the primary end point. PFS is an attractive end point because it is available earlier than overall survival (OS) and is not influenced by second-line treatments. PFS is now undergoing validation as a surrogate end point in various disease settings. The question of whether PFS can be considered an acceptable surrogate end point depends not only on formal validation studies but also on a standardized definition and unbiased ascertainment of disease progression in clinical trials. In advanced breast cancer, formal validation of PFS as a surrogate for OS has so far been unsuccessful. In advanced colorectal cancer, in contrast, current evidence indicates that PFS is a valid surrogate for OS after first-line treatment with chemotherapy. The other question is whether PFS sufficiently reflects clinical benefit to be considered a true end point in and of itself.

Key words: breast neoplasms, colorectal neoplasms, disease-free survival, end point determination, survival analysis

introduction

Over the past 10 years, significant progress has been achieved in the systemic treatment of both advanced breast cancer and advanced colorectal cancer. In breast cancer, the introduction of novel chemotherapeutic agents [1], the use of aromatase inhibitors in postmenopausal women [2, 3], of trastuzumab in patients with human epidermal growth factor type 2-overexpressing tumors [4, 5], and of bevacizumab in combination with taxanes [6–8], have improved the outlook for women with metastatic disease. In colorectal cancer, improved results have been achieved with the addition of effective chemotherapy agents to the backbone drug fluorouracil [9], as well as with the use of the mAbs bevacizumab, cetuximab, and panitumumab in various treatment lines [10–14]. Moreover, promising therapies in advanced breast cancer include pertuzumab [15], lapatinib [16], and ixabepilone [17]. In colorectal cancer, promising novel agents currently being tested include cediranib (AZD2171), dasatinib, and aflibercept, among others.

The increasing availability of active agents against cancer makes clinical development of novel therapies progressively more complex. In view of the growing number of possible drugs, combinations, and sequences to be tested, and given the opportunity to test novel therapies in different lines of treatment, the choice of end points for clinical trials is becoming a critical issue in drug development. In addition, the growing number of options is making treatment choice increasingly more complex for clinicians, who are faced with the need to distinguish between clinical trial end points and therapeutic objectives in their patients. In this paper, we provide an overview of recent achievements regarding the use of efficacy end points both in advanced breast cancer and colorectal cancer. We focus our review on the current evidence regarding the use of progression-free survival (PFS) as a surrogate for overall survival (OS) in these two frequent malignancies.

OS and other time-dependent end points

Given its objectivity and the unquestionable benefit derived by patients, OS has been historically considered the most important end point in medical oncology [18]. However, OS has also been shown to be an elusive end point, although objective and simple to measure, it has the disadvantage of requiring extended patient follow-up and of being confounded by causes of mortality unrelated to cancer. Furthermore, as
novel therapies are shown to be effective along the treatment continuum, patient survival may be influenced by the use of such therapies after participation in a given trial [19, 20]. In the latter scenario, there is frequent crossover to the agent under investigation in the trial, thus obscuring the effect of treatment on OS. As a result of these various factors, many randomized trials are grossly underpowered to detect plausible OS differences [21]. Other time-dependent end points on the basis of tumor assessment are available for use in drug development, and many of them are undergoing validation either as surrogate end points or as indicators of clinical benefit.

PFS is defined as the time elapsed between treatment initiation and tumor progression or death from any cause, with censoring of patients who are lost to follow-up [22]. Many recent randomized trials in solid tumor oncology have used PFS as the primary end point. In addition, several trials have also used time to tumor progression (TTP) as the primary end point, with TTP differing from PFS in that the events of interest are only disease progression [23, 24] and, in some studies, death due to malignancy [25]. Historically, both PFS and TTP have often been accepted as markers of clinical benefit for drug approval on the two sides of the Atlantic [26, 27]. For regulatory purposes, PFS may be preferable to TTP because it also captures fatal toxicity, while TTP is unconfounded by deaths unrelated to cancer [18].

Other time-dependent end points are also used occasionally as primary or secondary end points in randomized trials in breast or colorectal cancer. Time to treatment failure (TTF), rarely used as primary end point, considers any reason for treatment interruption as an event in a Kaplan–Meier analysis (disease progression, treatment toxicity, patient preference, or death) [22, 26]. On occasion, TTF has been misleadingly defined as the interval from study entry to disease progression or death from any cause (i.e. the definition of PFS) [25, 28]. Because TTF in its original definition is a composite end point that also includes subjective symptom assessment, it is seldom used for regulatory purposes by the Food and Drug Administration (FDA). Since the FDA must determine that approved drugs are both safe and effective, separate analyses of TTP, OS, and toxicity are required for cancer drug marketing application approval [26]. Another parameter sometimes used as a secondary end point, especially in advanced breast cancer, is duration of response (DOR), for which only responding patients constitute the denominator. The problem of defining an appropriate time to event end point in clinical trials has been made more complex because of drug holidays, whereby patients who are in stable condition can be taken off a drug having significant cumulative toxicity (e.g. oxaliplatin). The drug is restarted upon disease progression, which may lead to another period of disease stability during which the drug is discontinued again and so on in a ‘stop-and-go’ fashion. In such cases, innovative end points such as duration of disease control [29] or time to failure of strategy (TFS) [30] have been proposed as clinically meaningful end points for advanced colorectal cancer. TFS, for instance, is defined as the time to (i) addition of any agent not in the primary regimen, (ii) disease progression on full therapy or followed by no additional therapy, or (iii) death.

The majority of patients having an event in the analysis of PFS actually have documented disease progression. In other words, few patients die with no prior documentation of disease progression; hence PFS and TTP are nearly identical end points. For this reason, and because TTP and DOR are seldom used as primary end points, we will focus on PFS as the time-dependent efficacy end point of interest in the following discussion regarding surrogacy for OS in breast and colorectal cancers.

criteria for adopting a surrogate for OS

Before a surrogate end point can replace a so-called ‘true’ end point of interest, it must be formally validated, a process that has caused considerable controversy in the past two decades [31]. In the realm of medical oncology, multiple approaches toward validation have been taken in different situations. Despite the fact that PFS and TTP have often been used and informally accepted as surrogate end points for OS in oncology [26, 27], the methodological basis for establishing surrogacy is still evolving. The ongoing debate on this subject started in the literature when Prentice [32] outlined his definition and a set of criteria for surrogacy in a landmark paper published in 1989. Prentice’s definition of a surrogate end point was that of a response variable for which the test of the null hypothesis of no relationship to the treatment groups under comparison was also a valid test of the corresponding null hypothesis on the basis of the true end point. This definition put the emphasis on hypothesis testing as the paradigm behind a statistical validation. Subsequently, Freedman et al. and Buyse et al. have emphasized estimation and prediction rather than hypothesis testing for surrogate end point validation [31, 33]. Various authors have proposed meta-analytic approaches, arguing that a large body of data from individual patients were required for validation of end points [31, 34]. Buyse and Molenberghs [35] suggested that the association between the surrogate and true end points should be assessed after adjustment for the treatment effect, thus introducing the concept of individual-level surrogacy. They later proposed that a surrogate end point be assessed both at the ‘individual level’ and at the ‘trial level’ for its ability to predict the effect of treatment on the true end point, after observation of the treatment effect on the surrogate. In short, their methodology evaluates the correlation between end points (individual level) and the correlation between treatment effects on these end points (trial level), with the latter assuming greater importance in the validation process. Of note, the trial level correlation is mathematically independent of the individual level correlation (at least for normally distributed end point), which is somewhat counterintuitive but implies that a claim of surrogacy requires stronger conditions than a mere correlation between the surrogate and the true end point [36, 37].

Validated surrogates are rare in medical oncology, as the process of validation is time consuming and not always successful [18]. Fleming [34] has proposed a hierarchy of end points used to measure the outcomes of medical interventions. According to this hierarchy, level 1 describes a true end point; level 2 refers to a validated surrogate end point that is specific for a disease setting and class of interventions; level 3 is
a nonvalidated surrogate end point that is considered as 'reasonably likely to predict clinical benefit' (for a specific disease setting and class of interventions); and level 4 denotes a correlate that is a measure of biological activity but does not qualify at a higher level. Within this framework, the question to be answered in advanced breast and colorectal cancers is whether there is enough evidence to indicate that PFS may be considered a level 1 or 2 (as opposed to level 3 or 4) outcome measure.

methodological issues in the use of PFS

As stated above, PFS is an attractive end point for clinical trials because it is available earlier than OS, is less influenced than OS by competing causes of death, and is not influenced by second-line treatments. However, the use of PFS is not without problems in the context of clinical trials. In contrast to the objectivity of OS, the ascertainment of disease progression is potentially subject to measurement error and bias. The accurate determination of the time of disease progression can be problematic in individual patients, and the quality of PFS measurement can vary between centers. Moreover, PFS has been criticized because the date at which radiological evaluation confirms progression is in fact a proxy for the true progression time, which lies somewhere within the time interval between two successive assessments [38]. The consequent overestimation of median PFS may complicate comparisons across trials if patients undergoing different interventions have been subjected to different evaluation schedules. However, contrary to a commonly held view, measurement errors are unimportant in randomized trials if (and only if) such errors can be assumed to be random with respect to treatment allocation. This is the case in double-blind trials; hence in such trials, the use of PFS does not raise any serious methodological issues other than the unblinding of some patients as a result of obvious side-effects of the drugs under investigation. In open-label trials, in contrast, the ascertainment of PFS might be subject to bias if the clinician making the evaluation was aware of treatment. Such bias or systematic (treatment related) error could have serious consequences on any claims of treatment differences and would therefore make the use of PFS problematic in these trials. These caveats notwithstanding, several investigators have proposed that PFS (or TTP), as opposed to OS, is an appropriate primary end point for studies in advanced breast or colorectal tumors [19, 20, 39].

PFS may also be contrasted with objective response as a potential surrogate end point for OS. Although the FDA has used response rate as the basis for regular and accelerated approvals, the agency acknowledges that tumor responses do not necessarily equate with clinical benefit, since nonresponding patients may benefit from delay in tumor progression [26]. Although responses in individual patients correlate with OS at the individual level in both breast [20, 40] and colorectal cancer [41], treatment benefits in terms of response rates do not reliably predict treatment benefit in terms of PFS or OS [39, 41]. This issue becomes even more important with the targeted agents currently available, for which long-term benefits have sometimes been seen despite the lack of significantly improved response rates [16, 42, 43]. The validation of response, PFS, or other end points as surrogates for OS requires large datasets of patients randomly assigned to treatments and in whom both the surrogate and the true end point (OS) have been measured [44]. Since the number of patients available in most trials in metastatic breast cancer is too low to show treatment benefits on OS, a meta-analysis on the basis of individual patient data seems the only way to carry out a convincing validation. We now turn to such recently carried out meta-analyses in advanced breast and colorectal cancers.

PFS in breast cancer

In advanced breast cancer, treatment aims mainly at increasing survival and improving the quality of life [45]. However, many randomized trials in advanced breast cancer have actually been underpowered to detect plausible OS differences [21]. Indeed, OS gain has seldom been achieved in the hundreds of randomized trials conducted to date in advanced breast cancer [45]. Additionally, the assessment of quality of life is more often a secondary study objective in advanced breast cancer and measurement of quality of life provides little information beyond that obtained from other outcomes [46]. Therefore, validation of surrogate end points for OS remain an important goal in breast cancer research.

Several investigators have attempted to assess the correlation between treatment effects on OS and on potential surrogate end points in advanced breast cancer [20, 40, 44, 47]. A’Hern et al. [40] analyzed the correlation between response rates and OS and found a statistically significant relationship between these two end points by weighted linear regression. Almost two decades after the work by A’Hern et al., Hackshaw et al. [20] and Bruzzi et al. [44] also provided evidence for the association between the effects of anthracycline-based chemotherapy on response and other potential surrogate end points and OS in advanced breast cancer. Hackshaw et al. found significant linear associations between OS and various end points, such as partial and complete response, disease progression, and TTP. The proportion of the variability in the treatment effect on OS explained by the treatment effect on the surrogate end point was higher for TTP, leading the authors to conclude that this was the best surrogate marker examined [20]. All of these analyses are on the basis of data extracted from the literature; their conclusions must therefore be interpreted with caution [48–50].

Burzykowski et al. [51] conducted a meta-analysis on the basis of individual patient data from 11 randomized trials including 3953 patients and comparing an anthracyline (alone or in combination) with a taxane (alone or in combination with an anthracyline) as first-line therapy for metastatic disease. The results indicated that PFS is not a good surrogate for OS in this setting because of an only modest correlation between treatment effects on these two end points. Taking a trial-level approach, Miksad et al. [52] came to a different conclusion, after finding that the hazard ratios (HR) for PFS were significantly correlated with HR for OS in trials of anthracyclines and taxanes, albeit with only modest explained variances. As previously pointed out, individual-level data are far superior to published summary statistics as a basis for establishing surrogate [35]. Therefore, one may conclude that PFS remains a level 3 surrogate end point...
The question of whether PFS is a surrogate for OS in advanced breast cancer is increasingly pressing in view of the numerous trials showing improved PFS but no gain in OS [21, 53]. Bevacizumab, the mAb against vascular endothelial growth factor, provides an interesting recent example with regard to the issue of PFS as primary end point in a phase III trial. Bevacizumab has improved results in advanced breast cancer by significantly increasing the response rate and PFS, with the latter being the primary end point in the trial that led to FDA approval of the agent in this setting [6]. The HR for disease progression or death in the trial was 0.60, which is among the largest risk reductions seen in advanced breast cancer. In that trial, there was no significant difference in OS (HR = 0.88, \( P = 0.16 \)). Unfortunately, no data were collected regarding the use of post-trial therapy, thus precluding a quantitative assessment of the impact of crossover on OS. As a matter of fact, reporting of subsequent-line therapy is infrequent in medical oncology and so are protocol-controlled second-line therapies [54]. In some cases, a survival analysis with censoring of patients who crossed over to the experimental drug after progression indicates a survival benefit, lending further support to the hypothesis that crossovers confound any benefit of first-line treatments on OS [21]. None the less, the effect of crossover seems to vary in different situations. In the pivotal trial leading to the approval of trastuzumab in combination with chemotherapy for advanced breast cancer, patients who were randomly assigned to receive chemotherapy alone had inferior OS, despite the fact that ~70% of such patients subsequently received trastuzumab. Therefore, further research is needed to clarify the role of PFS as the true end point, rather than a (poor) surrogate end point in advanced breast cancer, in an era of effective subsequent-line therapy.

**PFS in colorectal cancer**

The debate on the appropriateness of OS as an end point in clinical trials has been ongoing in the literature on colorectal cancer for several years [19, 21, 23, 55, 56]. The increasing availability of active agents in advanced colorectal cancer has been accompanied by various attempts to define adequate end points for efficient drug development. In advanced disease, treatment intensity in first line is still a matter of controversy, and chemotherapy doublets with or without an mAb is customary in clinical practice [57]. Since effective salvage therapies currently exist and may compensate for less-active first-line therapies, OS may no longer be the most appropriate primary efficacy end point in the first-line therapy of advanced colorectal cancer [58]. As discussed previously, many recently reported trials in advanced colorectal cancer lacked power to detect a statistically significant increase in OS, even in the presence of other benefits [21]. Furthermore, continuing demonstration of OS gain in advanced colorectal cancer is likely to become increasingly rare in the near future, once the use of effective agents becomes more widespread.

Tang et al. [23] have recently conducted a literature-based analysis from 39 randomized trials of first-line chemotherapy in advanced colorectal cancer assessing the correlation between treatment effects on OS and on PFS, TTP, and response rate. These authors found a strong relationship between HR for PFS and OS indicating that PFS might be an appropriate surrogate for OS and that a novel therapy producing a 10% risk reduction for PFS would yield an estimated 5.4% risk reduction for OS. As part of an ongoing initiative toward end point validation in various disease settings, Buyse et al. obtained individual data from 3089 patients participating in 10 ‘historical’ trials comparing fluorouracil and leucovorin with either fluorouracil alone or with raltitrexed and from 1263 patients participating in three ‘validation’ trials comparing fluorouracil and leucovorin with or without irinotecan or oxaliplatin [59]. These authors found high correlation coefficients between PFS and OS and between treatment effects on PFS and on OS in the historical trials. These coefficients allowed the construction of a prediction model for the validation trials. On the basis of the observed treatment effect on PFS and on the model, the observed treatment effect on OS was compared with the predicted treatment effect on OS and was found to be within the 95% prediction intervals for the three validation trials. The authors concluded that PFS was an acceptable surrogate for OS for patients receiving first-line chemotherapies in advanced colorectal cancer. In an accompanying editorial, Yotkers [55] not only concurred with the findings by Buyse et al. but also argued that PFS is itself a more adequate end point than OS. A similar opinion has been put forward by Grothey et al. [39] who considered PFS and the percentage of patients experiencing tumor control as the most appropriate end points for trial design in advanced colorectal cancer.

Current evidence regarding the use of PFS as a surrogate for OS in advanced colorectal cancer allows the classification of PFS as a level 2 surrogate end point (i.e. a validated surrogate end point), according to the schema proposed by Fleming [34]. Such evidence applies to the setting where patients receive first-line chemotherapy. Of note, the majority of patients whose data led to this conclusion participated in trials conducted in an era when effective second-line therapies were infrequently available. Since surrogacy is specific to a disease setting and class of interventions, it remains an open question whether PFS is also a level 2 surrogate end point for OS among patients with advanced colorectal cancer treated in the era of effective second-line therapies, as well as those treated in other lines and with molecular-targeted agents. This question may never be resolved because of the availability of effective agents, which patients can receive after failing their first-line treatment regimen. However, the fact that PFS has been shown to be an acceptable surrogate for OS in absence of effective second-line therapies provides partial support to the view that PFS is a desirable end point to use in future clinical trials.

**conclusions**

This review of the literature on advanced breast and colorectal cancers indicates that PFS is commonly used as a surrogate for OS in clinical trials. Despite the fact that surrogacy has not been firmly established in breast cancer, PFS (or TTP) has
been chosen as the primary end point in the majority of recent trials and has been used as the basis for recent approval of several agents [6, 25, 26]. In colorectal cancer, PFS is a validated surrogate end point in the setting of first-line chemotherapy and can as such be more firmly recommended as a relevant end point to use in clinical trials.

In conclusion, PFS currently represents the most sensitive parameter to assess the efficacy of a new drug or combination in advanced breast cancer and advanced colorectal cancer. When coupled with a favorable toxicity profile, the demonstration of an improved PFS appears to constitute enough evidence for the superiority of a treatment both in the setting of clinical trials and for translating this information into treatment decisions for clinical practice. We believe that the question of whether PFS is a surrogate or a true end point depends not only on formal validation studies but also on a standardized definition and unbiased ascertainment of disease progression in clinical trials. Since there are no standard regulatory criteria for defining progression [24], this remains an important question for the near future. In the meantime, while it seems clear that extending survival remains the principal treatment goal in advanced cancer, the best way to achieve this goal may be the sequential use of treatments with demonstrated superiority in terms of time to disease progression as the chief indicator of therapeutic efficacy in an era of active subsequent-line therapies.

disclosure

The authors declare no conflict of interest.

references


