Progression-free survival as a primary endpoint in clinical trials of metastatic colorectal cancer

S. Gill BScPharm MD MPH,* S. Berry BSc MD MHSc,† J. Biagi MD,‡ C. Butts MD,§ M. Buyse ScD,|| E. Chen MD,## D. Jonker MD,** C. Mărginean MS(Biostatistics) MD,†† B. Samson MD,‡‡ J. Stewart RN BScN MHS,§§ M. Thirlwell MDCM,|||| R. Wong MD,### and J.A. Maroun MD***

ABSTRACT

In recent years, significant advances have been made in the management of metastatic colorectal cancer. Traditionally, an improvement in overall survival has been considered the “gold standard”—the most convincing measure of efficacy. However, overall survival requires larger patient numbers and longer follow-up and may often be confounded by other factors, including subsequent therapies and crossover. Given the number of active therapies for potential investigation, demand for rapid evaluation and early availability of new therapies is growing. Progression-free survival is regarded as an important measure of treatment benefit and, compared with overall survival, can be evaluated earlier, with fewer patients and no confounding by subsequent lines of therapy. The present paper reviews the advantages, limitations, and relevance of progression-free survival as a primary endpoint in randomized trials of metastatic colorectal cancer.

KEY WORDS
Progression-free survival, metastatic colorectal cancer, surrogate endpoints, targeted therapies, randomized clinical trials

1. INTRODUCTION

Colorectal cancer is the third most common malignancy worldwide and the second cause of cancer death in both sexes in the developed world 1. One quarter of patients with colorectal cancer have liver metastases at the time of diagnosis, and among those with initially localized disease, 25%–30% will present with metastases in the following 2–3 years 2,3. Without treatment, the life expectancy for patients with metastatic colorectal cancer (mCRC) ranges from 5 months to 9 months 2,4.

2. CURRENT MANAGEMENT OF mCRC

Since the year 2000, significant progress has been made in the treatment of mCRC. Clinically meaningful advances have been achieved with the introduction of new chemotherapeutic agents (particularly irinotecan and oxaliplatin, but also 5-fluorouracil 5,6) and the targeted monoclonal antibodies (bevacizumab, cetuximab, and panitumumab) in various therapeutic schedules 7–9. In addition, novel and promising targeted agents such as aflibercept are currently being tested. Prolonged durations of therapy and the dose-limiting neurotoxicity of oxaliplatin have prompted a shift from the conventional paradigm of treatment until progression or toxicity, to chemotherapy with planned interruptions. Curative metastasectomy after successful downsizing with neoadjuvant chemotherapy has revolutionized treatment objectives to include curative intent in stage IV disease.

3. OVERALL SURVIVAL AS A PRIMARY ENDPOINT IN ONCOLOGY TRIALS

In cancer drug trials, an improvement in overall survival (os), defined as the time from randomization to death from any cause, is considered the most convincing measure of drug efficacy and clinical benefit. Easily measured, unambiguous, and objective, os is a variable that is not subject to the biases associated with endpoints requiring clinical judgment. However, the measurement of os requires large patient numbers and prolonged follow-up, and it may be confounded by the use of effective subsequent-line therapies, by crossover within randomized trials from control to the investigational drug, and by mortality unrelated to cancer. Those factors obscure the effect of treatment on os and have led to randomized trial designs that lack the statistical power to detect plausible differences in os 10.
In view of the growing number of active drugs from various therapeutic classes (cytotoxic, cytostatic) and of drug combinations and sequences to be tested, and because of increasing demands for rapid evaluation and early availability of efficacious therapies in metastatic disease, the selection of auxiliary (“surrogate”) endpoints for OS has become a critical issue in drug development for regulatory marketing approval. Progression-free survival (PFS), defined as the time from randomization until first evidence of objective tumour progression or death from any cause, with censoring of patients who are lost to follow-up, is regarded as a sensitive indicator of treatment activity. Progression events are informative regarding the effect of treatment on the disease process and, if generally not confounded by subsequent lines of therapy and, compared with OS, can be assessed earlier.

In the present paper, the potential advantages and limitations, statistical efficiency, and clinical relevance of PFS as a primary endpoint in randomized trials of mCRC are discussed. That discussion, together with post hoc statistical analyses of data from older studies (when possible), may lead to the implementation by regulatory agencies of PFS as a surrogate for OS in this clinical context.

4. CRITERIA FOR SURROGACY

Before a surrogate endpoint can replace a true endpoint of interest, it must be formally validated. Prentice 12, in a seminal paper, outlined his criteria and definition for surrogacy, according to which a surrogate endpoint is a response variable for which the test of the null hypothesis (no relationship to the true endpoint) is a response variable for which the test of the null hypothesis (no relationship to the true endpoint) is generally not confounded by measurement error and bias. First, the determination of the time at which disease progresses is an overestimation of median OS (as measured by the hazard ratios [HRS] comparing treatment with controls within each trial) was compared with the predicted treatment effect on OS based on the observed treatment effect on PFS. All observed HRS fell within the 95% prediction limits, including close agreement in validation trials using irinotecan and, to a lesser extent, oxaliplatin. The authors concluded that PFS is a valid measure of surrogacy for OS in the first-line treatment of mCRC with cytotoxic agents. The “surrogate threshold effect,” defined as the minimum treatment effect on PFS that would predict a statistically significant treatment effect on OS, corresponded to a PFS HR of 0.86. Thus, achieving that threshold effect on PFS would be predictive of a benefit in terms of OS.

In a literature-based surrogate endpoint analysis of 39 randomized mCRC trials containing 87 treatment arms, Tang et al. 17 reported correlation coefficients of 0.79 (95% confidence interval: 0.65 to 0.87) between PFS and OS, and 0.74 (95% confidence interval: 0.48 to 0.88) between the within-trial differences in PFS and OS. Poorer correlation (0.59) was observed between response rate (RR) and OS, and between within-trial changes in RR and in OS (0.39). In another report, RR was not predictive of the treatment effect on OS, thus failing to meet criteria for surrogacy.

The use of PFS as an endpoint is also associated with a significant lead-time advantage (the amount of lead time that would be gained by using PFS rather than OS) of almost 1 year for a median OS of 1.5 years. Consequently, PFS has emerged as the primary endpoint of choice in recent first-line phase III trials in mCRC (Table 1). In addition, the regulatory agencies have considered PFS to be an acceptable trial endpoint for accelerated or regular approval in first-line treatment of this disease 31.

Compared with OS, PFS is an attractive endpoint in clinical trials because it is available earlier, is less influenced by competing causes of death, and is unaffected by crossover and subsequent lines of therapy. However, PFS is associated with measurement error and bias. First, the determination of the time at which disease progresses is an approximation that lies within the time interval between 2 successive, protocol-specified radiologic evaluations. The particular scheduling may lead to an overestimation of median PFS and complicates comparisons across trials if patients undergoing different interventions have been subject to different evaluation schedules. Moreover, the evaluation of progression based on radiographic scans is subjective. In spite of efforts (such as the Response Evaluation Criteria in Solid Tumors) to standardize the evaluation of progressive disease discrepancies persist in investigator interpretation of radiographic assessments. In non-blinded studies, a blinded independent central radiology committee review is proposed as a strategy to control bias. However, it was recently suggested such blinding does not

5. PFS AS A SURROGATE ENDPOINT IN mCRC

Based on the criteria for surrogacy already described, Buyse et al. put forth a method for the statistical validation of PFS as a surrogate for OS in mCRC. Using data from historical phase III trials, those authors demonstrated a high rank correlation coefficient between PFS and OS (0.82), and a high correlation coefficient between treatment effects on PFS and on OS (0.99, or 0.74 after one very influential study was excluded from the analysis). The observed treatment effect on OS [as measured by the hazard ratios (HRS) comparing treatment with controls within each trial] was compared with the predicted treatment effect on OS based on the observed treatment effect on PFS. All observed HRS fell within the 95% prediction limits, including close agreement in validation trials using irinotecan and, to a lesser extent, oxaliplatin. The authors concluded that PFS is a valid measure of surrogacy for OS in the first-line treatment of mCRC with cytotoxic agents. The “surrogate threshold effect,” defined as the minimum treatment effect on PFS that would predict a statistically significant treatment effect on OS, corresponded to a PFS HR of 0.86. Thus, achieving that threshold effect on PFS would be predictive of a benefit in terms of OS.

In a literature-based surrogate endpoint analysis of 39 randomized mCRC trials containing 87 treatment arms, Tang et al. 17 reported correlation coefficients of 0.79 (95% confidence interval: 0.65 to 0.87) between PFS and OS, and 0.74 (95% confidence interval: 0.48 to 0.88) between the within-trial differences in PFS and OS. Poorer correlation (0.59) was observed between response rate (RR) and OS, and between within-trial changes in RR and in OS (0.39). In another report, RR was not predictive of the treatment effect on OS, thus failing to meet criteria for surrogacy.

The use of PFS as an endpoint is also associated with a significant lead-time advantage (the amount of lead time that would be gained by using PFS rather than OS) of almost 1 year for a median OS of 1.5 years. Consequently, PFS has emerged as the primary endpoint of choice in recent first-line phase III trials in mCRC (Table 1). In addition, the regulatory agencies have considered PFS to be an acceptable trial endpoint for accelerated or regular approval in first-line treatment of this disease 31.

Compared with OS, PFS is an attractive endpoint in clinical trials because it is available earlier, is less influenced by competing causes of death, and is unaffected by crossover and subsequent lines of therapy. However, PFS is associated with measurement error and bias. First, the determination of the time at which disease progresses is an approximation that lies within the time interval between 2 successive, protocol-specified radiologic evaluations. The particular scheduling may lead to an overestimation of median PFS and complicates comparisons across trials if patients undergoing different interventions have been subject to different evaluation schedules. Moreover, the evaluation of progression based on radiographic scans is subjective. In spite of efforts (such as the Response Evaluation Criteria in Solid Tumors) to standardize the evaluation of progressive disease discrepancies persist in investigator interpretation of radiographic assessments. In non-blinded studies, a blinded independent central radiology committee review is proposed as a strategy to control bias. However, it was recently suggested such blinding does not
TABLE 1  Endpoints in first-line phase III trials in metastatic colorectal cancer

<table>
<thead>
<tr>
<th>Reference</th>
<th>Phase</th>
<th>Treatment arms</th>
<th>Primary</th>
<th>Secondary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goldberg et al., 2004</td>
<td>III</td>
<td>IFL VS. FOLFOX VS. IROX</td>
<td>TTP</td>
<td>RR, OS</td>
</tr>
<tr>
<td>Hurwitz et al., 2004</td>
<td>III</td>
<td>IFL ± bevacizumab</td>
<td>OS</td>
<td>RR, response duration, PFS</td>
</tr>
<tr>
<td>Tournigand et al., 2004</td>
<td>III</td>
<td>FOLFOX→FOLFI vs. reverse sequence</td>
<td>OS</td>
<td>RR, PFS</td>
</tr>
<tr>
<td>(non-inferiority)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tournigand et al., 2006</td>
<td>III</td>
<td>FOLFOX7–LV5FU2 VS. FOLFOX4</td>
<td>DDC</td>
<td>RR, R0, OS</td>
</tr>
<tr>
<td>Falcone et al., 2007</td>
<td>III</td>
<td>FOLFOXIRI VS. FOLFI</td>
<td>RR</td>
<td>PFS, OS</td>
</tr>
<tr>
<td>Nordlinger et al., 2008</td>
<td>III</td>
<td>Liver resection ± perioperative FOLFOX</td>
<td>PFS</td>
<td>OS</td>
</tr>
<tr>
<td>Saltz et al., 2008</td>
<td>III</td>
<td>XELOX/FOLFOX ± bevacizumab</td>
<td>PFS</td>
<td>RR, OS</td>
</tr>
<tr>
<td>Cunningham et al., 2009</td>
<td>III</td>
<td>OX–5FU, OX–5FU–LV, 5FU, 5FU–LV</td>
<td>2-Year survival</td>
<td>PFS, TTF, toxicity</td>
</tr>
<tr>
<td>Hecht et al., 2009</td>
<td>III</td>
<td>5FU–IR/OX–bevacizumab ± panitumumab</td>
<td>PFS</td>
<td>RR, OS</td>
</tr>
<tr>
<td>Tol et al., 2009</td>
<td>III</td>
<td>CAPOX–bevacizumab ± cetuximab</td>
<td>PFS</td>
<td>RR, OS</td>
</tr>
<tr>
<td>Van Cutsem et al., 2009</td>
<td>III</td>
<td>FOLFI ± cetuximab</td>
<td>PFS</td>
<td>RR, R0, OS</td>
</tr>
<tr>
<td>Fischer von Weikersthal et al.,</td>
<td>III</td>
<td>MIROX VS. FUFIRI</td>
<td>PFS</td>
<td>RR, OS, R0, toxicity</td>
</tr>
<tr>
<td>2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tebbutt et al., 2010</td>
<td>III</td>
<td>C–CB, C–CBM</td>
<td>PFS</td>
<td>OS, RR, toxicity, QOL</td>
</tr>
</tbody>
</table>

± = with or without; IFL = irinotecan, 5-fluorouracil (5FU), leucovorin; RR = response rate; PFS = progression-free survival; FOLFOX = 5FU, leucovorin, oxaliplatin; IROX = irinotecan, oxaliplatin; TTP = time to progression; OS = overall survival; FOLFI = 5FU, leucovorin, oxaliplatin, high-dose oxaliplatin, followed by 5FU and leucovorin maintenance; FOLFOX4 = oxaliplatin, leucovorin, 5FU; DDC = duration of disease control; R0 = R0 resection rate; FOLFOXIRI = 5FU, leucovorin, oxaliplatin, irinotecan, XELOX = capecitabine, oxaliplatin; OX = oxaliplatin; LV = leucovorin; TTP = time to treatment failure; IR = irinotecan; CAPOX = capecitabine, oxaliplatin; C = capecitabine; CB = capecitabine, bevacizumab; CBM = capecitabine, bevacizumab, mitomycin; QOL = quality of life; MIROX = irinotecan, oxaliplatin; FUFIRI = irinotecan, 5FU, folinic acid.

In contrast to blinded trials, the ascertainment of PFS in open-label studies is more problematic for a number of reasons beyond those contributed by clinician interpretation and by biases introduced when the clinician making the evaluation is aware of the treatment. For instance, a different patient evaluation schedule for each treatment arm can lead to differential measurement error and a highly biased estimate of treatment effect. Properly designed trials should specify the same evaluation schedule for all treatment arms. Additionally, because progression can be captured at non-scheduled patient evaluation times (for example, when the patients develop new symptoms), further increases in PFS assessment error may occur. Lastly, the clinical assessment of progression without radiologic confirmation by the Response Evaluation Criteria in Solid Tumors can also distort PFS data.

6. TARGETED THERAPIES AND OBJECTIVE RESPONSE RATES

Although regulatory agencies have occasionally accepted RRs as the basis for regular and accelerated drug approvals, RRs are recognized not always to translate into clinical benefit, and as mentioned...
earlier, RRs may not reliably predict treatment benefit in terms of PFS and OS in mCRC. However, patients classified as nonresponders according to traditional criteria may benefit from a therapeutically-induced delay in tumour progression—that is, disease stabilization. This issue becomes more important with the advent of biologic ("targeted") therapies in mCRC, which may be better characterized by a cytostatic rather than a cytotoxic mechanism of action. Stable disease rather than objective tumor shrinkage has been proposed as a relevant clinical endpoint by some investigators. It would be of interest to explore associations between PFS and OS in mCRC patients with confirmed stable disease. A trial design called "randomized discontinuation" has been suggested as a method to investigate agents with predominantly cytostatic activity.

7. USING PFS AS AN ENDPOINT FOR HEALTH CANADA MARKETING AUTHORIZATION

For the purposes of regulatory approval of a new agent, PFS should be ascertained in randomized double-blind trials. For first-registration submissions of first-line therapeutic agents for mCRC, PFS may be used as the primary endpoint provided that confirmatory trials with the statistical capacity for OS analysis are ongoing. Claims involving PFS should be substantiated with other endpoints with which PFS correlates and should be consistent with the pharmacology and biologic plausibility of the treatment. For subsequent submissions for first-line therapy, PFS would likely be unsuitable if the therapy in question for the same indication had already demonstrated an improved OS.

For market authorization of a treatment that is associated with symptom-free or health-related quality-of-life measures, PFS may also be the primary endpoint, provided that treatment toxicity is acceptable. Efforts should be made to minimize missing or unevaluable radiologic images. Biomarkers that are predictive of clinical response or toxicity, or both, should be used to better identify patients who could benefit from treatment with minimal adverse drug reactions. The regulatory approach should be conducted on a case-by-case basis, with selective approval (small effect in broad indication) or conditional approval, and requires restrictions in the label and a commitment by sponsors to provide confirmatory data.

8. CONCLUSIONS

The advances in the treatment of mCRC witnessed since the year 2000 are encouraging. A number of therapeutic options now exist that have increased OS from 5–9 months in untreated patients to nearly 3 years in treated patients. The growing diversity of treatment regimens, the introduction of new regimens with treatment interruptions and modifications to reduce long-term interruptions (particularly oxaliplatin-induced neurotoxicity), and curative metastectomies have increased the difficulty in evaluating survival data from patients undergoing first- and possibly second-line therapies. As illustrated in Table 1, PFS has indeed emerged as a primary endpoint in several randomized first-line studies in mCRC. Despite the limitations of statistical evaluation from older clinical trials data, recent analyses have demonstrated that PFS may be an acceptable surrogate for OS in mCRC.

Finally, proper validation of PFS as a surrogate for OS requires the application of, and adherence to, established guidelines and recommendations. In the Canadian Workshop on the Use of PFS in Clinical Trials in mCRC, the following items were viewed as most important when considering PFS as an endpoint:

- Crossover or contamination from other sources should be minimized.
- Timing of response assessments should be consistent between study arms within a trial.
- Clinically relevant absolute gains in PFS (for example, ≥2 months) should be achieved.
- Relevant IRS (that is, 20%–30% improvement) should be obtained.
- Changes in PFS should be corroborated by changes in one or more other endpoints (for example, RR, time to progression).
- Sufficient clinical data should be available at some point during drug development to allow assessment of OS, provided that data can be ethically obtained.

9. CONFLICT OF INTEREST DISCLOSURES

MB holds stock in idddi. JAM has received research funding from Roche and Sanofi–Aventis, and has received honoraria for participation on boards and as a speaker for Roche, Novartis, Amgen, and Sanofi–Aventis. The remaining authors have no conflicts of interest pertinent to the topic of this manuscript to declare.

10. REFERENCES


35. Saad ED, Katz A, Hoff PM, Buyse M. Progression-free survival as surrogate and as true end point: insights from the breast and colorectal cancer literature. *Ann Oncol* 2010;21:7–12.


**Correspondence to:** Sharlene Gill, BC Cancer Agency, 600 West 10th Avenue, Vancouver, British Columbia V5Z 4E6.

*E-mail:* sgill@bccancer.bc.ca

---

* University of British Columbia, Division of Medical Oncology, and BC Cancer Agency, Vancouver, BC.
† Sunnybrook Odette Cancer Centre and University of Toronto, Toronto, ON.
‡ Department of Oncology, Queen’s University, Kingston, ON.
§ Cross Cancer Institute and University of Alberta, Edmonton, AB.
∥ International Drug Development Institute, Louvain-la-Neuve, and I-BioStat, Hasselt University, Diepenbeek, Belgium.
# Department of Medical Oncology and Hematology, Princess Margaret Hospital, Toronto, ON.
** Ottawa Hospital Cancer Centre, Ottawa, ON.
†† Ottawa Hospital Research Institute, Ottawa, ON.
‡‡ Centre intégré de cancer de la Montérégie de l’Hôpital Charles-Lemoyne, Greenfield Park, QC.
$$ University Health Network, Princess Margaret Hospital, Toronto, ON.
|| McGill University, McGill University Health Centre, and Montreal General Hospital, Montreal, QC.
||| CancerCare Manitoba, University of Manitoba, Winnipeg, MB.
*** Ottawa Hospital Cancer Centre and University of Ottawa, Ottawa, ON.