Progression-free survival in oncology

caveat emptor!

Bergmann, Troels K; Christensen, Mette Marie H; Henriksen, Daniel; Haastrup, Maija Bruun; Damkier, Per

Published in:
Basic & Clinical Pharmacology & Toxicology

DOI:
10.1111/bcpt.13168

Publication date:
2019

Document version:
Accepted manuscript

Citation for pulished version (APA):
Bergmann, T. K., Christensen, M. M. H., Henriksen, D., Haastrup, M. B., & Damkier, P. (2019). Progression-free survival in oncology: caveat emptor! *Basic & Clinical Pharmacology & Toxicology*, *124*(3), 240-244. https://doi.org/10.1111/bcpt.13168

Article Type: Annotations & Reflections

Progression-free survival (PFS) in oncology: caveat emptor!

Short title: Progression-free survival: caveat emptor!

Troels K. Bergmann (1,2,3), Mette Marie H. Christensen (2,3,4), Daniel Henriksen (2,4), Maija Bruun Haastrup (1,2), Per Damkier (2,4)

1) Clinical Pharmacology and Pharmacy, Department of Public Health, University of Southern Denmark, Odense, Denmark
2) Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, Odense, Denmark
3) Hospital Pharmacy, Hospital of South West Denmark, Esbjerg, Denmark
4) Department of Clinical Research, University of Southern Denmark, Odense, Denmark

Keywords:

Progression-free survival, overall survival, surrogate endpoint, oncology, quality of life, cancer drugs

Correspondence:

Troels K. Bergmann

Clinical Biochemistry and Pharmacology

Odense University Hospital

J. B. Winsløws Vej 19, 5000 Odense

Denmark

Ph. +45 65503305

Mail: tbergmann@health.sdu.dk

Disclosure

All authors have nothing to disclose.

Abstract

Overall survival (OS) is the undisputed gold standard efficacy endpoint in cancer drug trials. It is with growing concern that we observe how progression-free survival (PFS) gains ground as surrogate endpoint in its place. PFS has appeal because it is resource-efficient, but it has severe shortcomings. Our concern is that uncritical use of PFS will harm the evidence-based evaluation of cancer drugs when considering them for standard use in publicly financed health care systems. PFS is only valid as a surrogate endpoint for OS if it correlates strongly with OS and if the cancer drug being investigated has the same effect on PFS and OS such that effects on one predicts effects on the other. The latter might be less obvious than the former but is no less critical. Research indicates that in a majority of cases, correlation between surrogate endpoints and OS is of medium strength or lower. PFS is therefore unreliable as a surrogate for OS. We do not find it justified to use PFS as surrogate for OS without first having assessed its validity. Stakeholders who take part in evaluating cancer drugs considered for standard use in a health care system must be particularly vigilant about this issue to minimize the risk of introducing cancer drugs that have an unacceptable cost-risk-benefit profile.

Annotations & Reflections

Progression-free survival (PFS) is increasingly used as a primary endpoint in oncology trials. This continuous incremental adoption of a surrogate for overall survival (OS) has important implications for patients, clinical oncologists, hospital owners, regulatory authorities, politicians and societal stakeholders. With this paper, we argue that surrogate endpoints – especially PFS - should be used with great care - if at all - within oncology. We discuss the following questions: 1) Is progression-free survival (PFS) valid as a surrogate for overall survival (OS) in oncology 2) Is PFS a relevant endpoint in its own right?

Background

We are of the opinion, and find it self-evident, that the only endpoints in oncology relevant to the patients are OS and quality of life (QoL). However, OS (and QoL) are increasingly being substituted by PFS in pivotal phase III trials supporting marketing authorization [1]. An American study from 2015

found that during the period from 2008 to 2012, 36 of 54 (67%) cancer drugs were approved by the United States Food and Drug Administration (US FDA) based on a surrogate endpoint. After a median subsequent follow-up of 4.4 years, improved OS had been demonstrated for an additional 5 drugs. Of the remaining 31 drugs, 18 failed to demonstrate effect on OS, and 13 drugs continued to have unknown survival effects [2]. A similar study from 2017 found that in the period of 2009 to 2013, the European Medicines Agency (EMA) approved 48 cancer drugs for 68 indications. Of these, 24 (35%) showed significantly improved OS with a median effect of 2.7 months. Improved QoL was demonstrated for 10%. After a median follow-up of 5.4 years, only 51% had shown significant improvement in OS or QoL [3].

PFS was developed as a descriptive biomarker of anti-neoplastic efficacy in scientific trials testing new cancer drugs. Typically, this involved imaging techniques, e.g. computerized tomography scans (CT) and a set of standardized evaluation criteria based on uni-, or bi-dimensional tumour measurements [4,5]. This approach has been useful in the trial and error approach to testing new compounds and new combinations of cytotoxic agents in oncology. RECIST (Response Evaluation Criteria in Solid Tumors) is an example of a set of criteria which has gained almost universal use [6]. However, it is important to note that tumour response, i.e. the shrinkage of a tumour, is not sufficient to demonstrate a clinically relevant effect on OS or QoL. In fact, the question of substitution in this regard is not simple. The word surrogate is from Latin and simply means substitute, "*to put in the place of another*" [7]. The term "surrogate endpoint" has been defined as: "a biomarker that is intended to substitute for a clinical endpoint. A surrogate endpoint is expected to predict clinical benefit (or harm or lack of benefit or harm) based on epidemiologic, therapeutic, pathophysiologic or other scientific evidence" [8]. The surrogate marker must be strongly correlated to the endpoint it substitutes – but even that does not suffice. It is also necessary that the intervention, i.e. the drug under investigation, has the same effect on the surrogate as on the endpoint, meaning that the surrogate will predict the effect on the endpoint of interest (fig. 1). Therefore, surrogacy can strictly only be established within the framework of a disease-specific intervention.

Is PFS a valid substitute for OS?

To address this question, it is useful to look for explanations in situations where a statistically significant effect is detected for either PFS or OS, but not the other. There are several possible explanations: 1) When treating small or very early solid tumours, it is conceivable that early tumour

regression and a corresponding prolonged PFS does not translate into postponed death. 2) Trials can be powered sufficiently to pick up an effect on PFS but not OS. This is possible because PFS invariably has more and earlier events compared to OS. This means that follow-up studies with continued observation will pick up the signal on OS. 3) PFS can be false positive. This may happen by chance, but the risk of a spurious result is increased in cases where the experimental arm and the control arm of a trial have different evaluation cadences. Evaluation bias, attrition bias and other biases may also play a role. 4) Theoretically, initial tumour remission might cause a change in the tumour phenotype towards a more aggressive type such that initial PFS improvement is related to worse outcome with decreased or unchanged OS. 5) Cross-over design, where patients from the experimental arm and/or the control arm may receive (or are automatically offered) the other treatment, is a widespread practice. This will attenuate any OS effect without diluting the effect on PFS. However, it cannot reasonably be assumed that post trial treatments cause a similar effect, as it is unlikely that the trial participants will be treated systematically differently in subsequent therapy lines. 6) Immune modulators have been associated with initial tumour swelling to a degree that the disease would meet the criteria for disease progression based on RECIST [9]. In such situations, PFS is clearly invalid as a substitute for OS. RECIST was modified in 2009 to include immune-related response criteria. Several updates and branched versions now exist with iRECIST from 2017 being the most recent addition.

Strong correlation between PFS and OS is required for surrogacy. A systematic review of trial-level meta-analyses in oncology found 36 articles with a total of 65 specific correlations between a surrogate endpoint and OS. Fifty-two percent of the correlations were of low strength ($r^2 \leq 0.49$). Approximately 25% were of medium strength (r^2, between 0.49 and 0.72). Twenty-three percent were highly correlated ($r^2 \geq 0.72$) with survival. The results showed that the conclusions varied considerably among cancer types [10]. There is no consensus on which factors that can be expected to impact on the relationship between PFS and OS. Proposed factors include: treatment line, year of the clinical trial, type of therapy (combination/mono), treatment type, region (global or regional), tumour type, sample size, cross-over and targeted therapies [11]. Estimations of OS and PFS for treatment-specific interventions may be complicated by sequential treatment lines.

To summarize, PFS is not ubiquitously valid as a surrogate for OS. PFS and OS may be well correlated, but having data that demonstrate that the intervention has the same effect on PFS and OS at the individual level happens very rarely. Consequently, the validity of PFS as a substitute for OS must be assessed on a case-by-case basis keeping the intervention and the disease-specific context in mind.

Is PFS valid as an endpoint in its own right?

What is the value of a drug that improves PFS but has no effect on OS?

PFS may be a relevant endpoint if it is a surrogate for QoL. The requirements are the same as discussed above, but in case of PFS the issue of what comprises a clinically meaningful effect size is even more critical. Two new studies consistently report poor correlation between PFS and QoL within cancer trials. Among 30 qualifying trials, no association between global QoL scores and median difference in PFS was found as regression slope between differences was 0.12 (95% CI -0.27-0.52)[12]. In a broader setting evaluating QoL in 147 clinical phase III cancer trials, the overall trial-level correlation between QoL and PFS was very weak with r squared value of 0.12[13]. Despite these global findings, examples of dubious reporting practices are common. Recently, a statistically significant improvement in PFS of about 9 *days* – which actually met prespecified study target (!), was reported in a study of *ramucirumab* in the treatment of metastatic gastric and gastroesophageal cancer. There was no impact on OS and QoL was not reported [14]. We severely doubt that any patient holds an interest in a PFS of 9 days with no data on QoL, and neither should clinicians or regulatory authorities. On a related note, the Marketing Authorization Holder (MAH) withdrew its EMA application for adjuvant *sunitinib* treatment in renal cell carcinoma. The primary outcome reported was PFS but the MAH position was challenged in a negative opinion from EMA Committee for Medicinal Products for Human use (CHMP). Sunitinib remains approved by the FDA for this very indication supported by the same data [15]. Most recently, a MAH-sponsored study on patient-reported QoL in this setting suggested a detrimental effect on QoL: "*Patients on sunitinib did report increased symptoms and reduced HRQoL…*" [16]. Notably, the authors applied the following narrative interpretation to their findings: "...these changes were generally not clinically meaningful, apart from appetite loss and diarrhoea, and were expected in the context of known sunitinib effects". Consequently, in the USA - an expensive drug on the market that is approved based on prolongation of PFS with documented lack of effect on OS and documented lack of beneficial effect on QoL. It appears that the argument here is that as long as there is no clinically meaningful *detrimental* effect on QoL, we are good. Similar conclusions were made in a – MAH-sponsored – study of niraparib for recurrent ovarian cancer: "…are able to maintain QoL during their treatment

compared with placebo."[17]. We are of the opinion that such not-so-subtle attempts to lower the clinical and regulatory bar of the evidence necessary to support meaningful patient benefits are unjustifiable and unacceptable.

In situations where it is unknown whether the requirements have been met, it is not justifiable to use PFS on this pretense. EMA does not require evaluation of quality of life data for cancer drugs to be authorized for the market [18], but these data are of obvious pertinence for all other stakeholders. Many validated health-related QoL assessment tools exist [19]. In situations where QoL data are available, it is unnecessary and counter-productive to use PFS in its place. In situations where QoL data are unavailable, PFS can be considered as a surrogate only after specific assessment of its validity in the case at hand. Failure to do so increases the risk of drawing wrong conclusions with regard to the true value of the drug.

Discussion

OS and QoL are the crucial endpoints in oncology trials of cancer drugs. PFS is fast and resource-efficient compared to OS (and QoL) but its use has pitfalls because the assumption that it adequately reflects true value to the patients, health-care professionals and other stake-holders may be wrong. The EMA and the US FDA approve cancer drugs based on an overall risk-benefit assessment of the available evidence on efficacy and safety. The most recent EMA guideline on the evaluation of cancer drugs states that: "*Prolonged PFS is considered to be of benefit to the patient*" [18]. EMA requires that if PFS is selected as a primary endpoint, OS should be the secondary endpoint, and vice versa. The guideline adds that in situations where there is a large effect on PFS, or if there is a long-expected survival after progression, OS is not necessarily required for approval of the drug. This means that while newly approved cancer drugs inherently are non-inferior or efficacious relative to PFS, it cannot be concluded that they necessarily meet the patients', health-care professionals' or societal stakeholders' reasonable expectations of clinically meaningful improvement in OS and /or QoL. In Denmark, The Danish Medicines Council [20] assesses new hospital medicines and approves their introduction as acceptable standard of care based on the balance between efficacy, safety and – recently introduced – price relative to added treatment value. Many other countries in Europe have similar agencies, e.g. The National Institute for Health and Care Excellence (NICE) in the United Kingdom [21]. It is the task of these agencies to balance all the information available to ensure a rational allocation of finite financial resources within their respective health-care systems. We are particularly worried that the increasing use of PFS may lower the bar for drug approval leading to

less effective drugs with high marginal costs which ultimately undermine the health economies and deprive other patients of proper and timely care.

In conclusion, using PFS without due consideration to its validity as a surrogate is not justifiable. We urge health care professionals, regulatory authorities and other stakeholders to exercise caution when drawing conclusions based solely on PFS in the assessment of the clinical value of cancer drugs and their place in the treatment hierarchy. We should always seek out and critically appraise the validity of the surrogate and demand that OS and QoL data are collected and presented as soon as possible allowing for decisions based on PFS to be amended if necessary.

References

[1] Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable? J Clin Oncol 2012;30:1030–3. doi:10.1200/JCO.2011.38.7571.

[2] Kim C, Prasad V. Cancer Drugs Approved on the Basis of a Surrogate End Point and Subsequent Overall Survival: An Analysis of 5 Years of US Food and Drug Administration Approvals. JAMA Intern Med 2015;175:1992–4. doi:10.1001/jamainternmed.2015.5868.

[3] Davis C, Naci H, Gurpinar E, Poplavska E, Pinto A, Aggarwal A. Availability of evidence of benefits on overall survival and quality of life of cancer drugs approved by European Medicines Agency: retrospective cohort study of drug approvals 2009-13. BMJ 2017:j4530. doi:10.1136/bmj.j4530.

[4] World Health Organization. WHO handbook for reporting results of cancer treatment. Geneva: Geneva : World Health Organization; 1979.

[5] Villaruz LC, Socinski MA. The Clinical Viewpoint: Definitions, Limitations of RECIST, Practical Considerations of Measurement. Clin Cancer Res 2013;19:2629–36. doi:10.1158/1078-0432.CCR-12-2935.

[6] The official site of the RECIST working group n.d. http://recist.eortc.org/ (accessed August 20, 2018).

[7] Definition of SURROGATE n.d. https://www.merriam-webster.com/dictionary/surrogate (accessed September 6, 2018).

[8] Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 2001;69:89–95. doi:10.1067/mcp.2001.113989.

[9] Seymour L, Bogaerts J, Perrone A, Ford R, Schwartz LH, Mandrekar S, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. The Lancet Oncology 2017;18:e143–52. doi:10.1016/S1470-2045(17)30074-8.

[10] Prasad V, Kim C, Burotto M, Vandross A. The Strength of Association Between Surrogate End Points and Survival in Oncology: A Systematic Review of Trial-Level Meta-analyses. JAMA Intern Med 2015;175:1389–98. doi:10.1001/jamainternmed.2015.2829.

[11] Extrapolation from Progression-Free Survival to Overall Survival in Oncology n.d. https://www.ohe.org/publications/extrapolation-progression-free-survival-overall-survival-oncology (accessed July 24, 2018).

[12] Kovic B, Jin X, Kennedy SA, Hylands M, Pędziwiatr M, Kuriyama A, et al. Evaluating

Progression-Free Survival as a Surrogate Outcome for Health-Related Quality of Life in Oncology: A Systematic Review and Quantitative Analysis. JAMA Intern Med 2018. doi:10.1001/jamainternmed.2018.4710.
[13] Hwang TJ, Gyawali B. Association between progression-free survival and patients' quality of life in cancer clinical trials. Int J Cancer 2018. doi:10.1002/ijc.31957.
[14] Fuchs CS, Shitara K, Di Bartolomeo M, Lonardi S, Al-Batran S-E, Van Cutsem E, et al. RAINFALL: A randomized, double-blind, placebo-controlled phase III study of cisplatin (Cis) plus capecitabine (Cape) or 5FU with or without ramucirumab (RAM) as first-line therapy in patients with metastatic gastric or gastroesophageal junction (G-GEJ) adenocarcinoma. JCO 2018;36:5–5. doi:10.1200/JCO.2018.36.4_suppl.5.
[15] Withdrawal of Application for a Change to the EU Marketing Authorisation for Sunitinib | ESMO n.d. https://www.esmo.org/Oncology-News/Withdrawal-of-Application-for-a-Change-to-the-EU-Marketing-Authorisation-for-Sunitinib (accessed September 14, 2018).
[16] Staehler M, Motzer RJ, George DJ, Pandha HS, Donskov F, Escudier B, et al. Adjuvant sunitinib in patients with high-risk renal cell carcinoma: safety, therapy management, and patient-reported outcomes in the S-TRAC trial. Ann Oncol n.d. doi:10.1093/annonc/mdy329.
[17] Oza AM, Matulonis UA, Malander S, Hudgens S, Sehouli J, Del Campo JM, et al. Quality of life in patients with recurrent ovarian cancer treated with niraparib versus placebo (ENGOT-OV16/NOVA): results from a double-blind, phase 3, randomised controlled trial. Lancet Oncol 2018;19:1117–25. doi:10.1016/S1470-2045(18)30333-4.
[18] European Medicines Agency - Clinical efficacy and safety - Evaluation of anticancer medicinal products in man n.d. http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_001122.jsp&mid=WC0b01ac0580034cf3 (accessed August 22, 2018).
[19] Soni MK, Cella D. Quality of life and symptom measures in oncology: an overview. Am J Manag Care 2002;8:S560-573.
[20] Medicinrådet - Medicinrådet - In English n.d. http://www.medicinraadet.dk/om-os/in-english (accessed August 21, 2018).
[21] NICE | The National Institute for Health and Care Excellence n.d. https://www.nice.org.uk/ (accessed August 21, 2018).

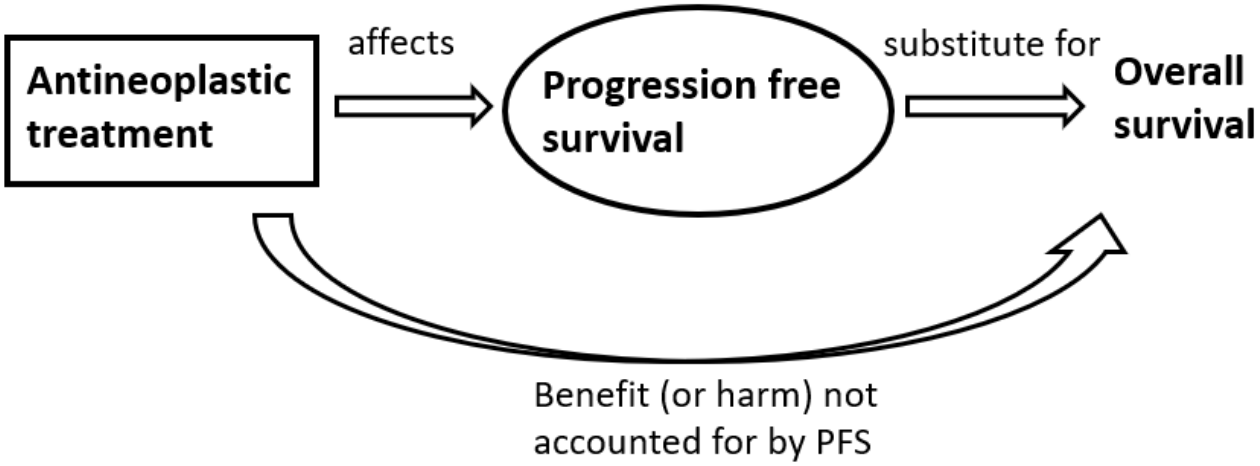

Figure 1. Progression-free survival is valid as a substitute for overall survival if the cancer drug being tested affects both PFS and OS similarly such that PFS and OS are correlated, and PFS effects predict OS effects on an individual level. Effects of the intervention not accounted for by PFS are indicated by the arrow at the bottom. The figure is modified from Atkinson et al. [8].