Journal of Cancer Policy 34 (2022) 100364

Contents lists available at ScienceDirect

Journal of Cancer Policy

journal homepage: www.elsevier.com/locate/jcpo

FDA validation of surrogate endpoints in oncology: 2005–2022

Anushka Walia [a], Alyson Haslam [b,*], Vinay Prasad [b]

[a] School of Medicine, University of California, San Francisco, 533 Parnassus Ave., San Francisco, CA 94143, USA
[b] Department of Epidemiology and Biostatistics, University of California, San Francisco, 550 16th St., San Francisco, CA 94158, USA

ARTICLE INFO

Keywords:
Surrogate
FDA
Drug approval
Endpoint
Overall survival
Cancer

ABSTRACT

Introduction: The number of oncologic drugs approved by the US Food and Drug Administration (FDA) on the basis of surrogate endpoints is rising. However, many surrogates have not demonstrated a correlation with clinically meaningful outcomes like overall survival. We sought to investigate surrogate validation studies conducted by the FDA over the past 17 years.

Methods: We reviewed analyses of surrogate outcomes published by the FDA from 2005 to 2022. Data extracted included the number of clinical trials included in each analysis, the associations of surrogate outcomes with OS or other surrogates, and the authors' interpretation of these associations.

Results: Of the 15 surrogate analyses conducted by the FDA, only one demonstrated a strong correlation between a surrogate outcome and overall survival. 87% only included clinical trials submitted to the FDA in their analysis, and all were published from 2014 onwards.

Discussion: The vast majority of FDA analyses of surrogate outcomes did not find strong correlations between surrogates and overall survival, raising concern about the use of such outcomes as endpoints in clinical trials. As most studies were based on limited data, further research is required to assess the true validity of surrogate outcomes.

Policy summary: Drugs approved on the basis of surrogates that are not associated with clinically meaningful outcomes can cause significant harm to patients. Until surrogate outcomes have been thoroughly and robustly validated, they should be used with caution in drug approval decisions.

1. Introduction

Approximately two-thirds of cancer drugs receive US Food and Drug Administration (FDA) marketing authorization based upon improvement in a surrogate endpoint, such as tumor shrinkage or progression free survival [1]. While surrogates may allow for earlier entry of drugs into the market, they are not always correlated with clinically meaningful outcomes like overall survival (OS) [2]. Some drugs approved on the basis of surrogate endpoints alone have subsequently failed to demonstrate improvements in OS. These approvals add cost and toxity for the patient without concomitant benefit [1].

Surrogate validation studies investigate the associations between surrogate markers and OS across multiple randomized controlled trials. Simply put, they allow regulators to differentiate which surrogates to have confidence in and which are unreliable for regulatory action. In recent years, the US FDA has embraced surrogates that historically have not been used for approval, raising questions as to the process by which they examine and validate surrogates [3]. In some cases, the agency may support the choice of these surrogate approvals by leveraging surrogate validation studies they have conducted, using internal or external data. As such, we sought to examine surrogate validation studies conducted by the US FDA published between 2005 and 2022.

2. Methods

We searched Google Scholar and PubMed for FDA analyses of surrogate outcomes from January 1, 2005 through June 1, 2022. We examined all articles by Richard Pazdur, director of the FDA's Oncology Center of Excellence and Office of Hematology and Oncology Products since 2005, as well as articles containing the terms "Center for Drug Evaluation and Research"[ad] with the filter "meta-analysis." We included all studies aiming to quantify relationships between existing or potential surrogate markers and OS.

Surrogate outcomes included progression-free survival (PFS), event-

* Correspondence to: Department of Epidemiology and Biostatistics, UCSF Mission Bay Campus, Mission Hall: Global Health & Clinical Sciences Building, 550 16th St, 2nd Fl, San Francisco, CA 94158, USA.
E-mail addresses: Anushka.walia@ucsf.edu (A. Walia), Alyson.haslam@ucsf.edu (A. Haslam), Vinayak.prasad@ucsf.edu (V. Prasad).

https://doi.org/10.1016/j.jcpo.2022.100364
Received 29 July 2022; Received in revised form 20 September 2022; Accepted 21 September 2022
Available online 22 September 2022

Table 1
Analyses of surrogate outcomes by the FDA, 2005–2022.

Year	Name	Abstract (Y/N)?	Cancer type	# Trials	Agents tested in trials	Only trials submitted to FDA? (Y/N)	Trial-level or patient-level?	Outcomes compared	Association	Study authors' interpretation	Our interpretation
2022	Haddock Lobo Goulart et al.	Y	Metastatic NSCLC	13	Anti-PD-1/PD-L1 (w or w/o anti CTLA-4 antibody or platinum-based chemotherapy)	Y	Trial-level	ORR and OS PFS and OS	r = 0.78 r = 0.83	Not strong correlation	Level 1 - Medium correlation
2021	Kim et al.	Y	HR-MDS	8	Azacitidine, Decitabine, Decitabine + cedazuridine	Y	Patient-level	OS, CR vs. non-responders EFS, CR vs. non-responders OS, HI vs. non-responders EFS, HI vs. non-responders OS, PR vs. non-responders EFS, PR vs. non-responders	HR = 0.40 HR = 0.39 HR = 0.60 HR = 0.60 HR = 0.51 HR = 0.66	CR may be associated with long-term benefit in patients with HR-MDS HI may be associated with long-term benefit in patients with HR-MDS PR may be associated with long-term benefit in patients with HR-MDS	Level 2 - Hazard ratio is significant Level 2 - Hazard ratio is significant Level 2 - Hazard ratio is significant Level 2 - Hazard ratio is not significant
2021	Sheth et al.	N	Advanced melanoma	13	Pembrolizumab, Dabrafenib, Trametinib, Vemurafenib, Ipilimumab, Nivolumab, Cobimetinib, Combination Rx	Y	Trial-level	ORR and PFS ORR and OS PFS and OS	r = 0.14 r = 0.30 r = 0.27	Weak relationship	Level 1 - Low correlation
2020	Chang et al.	Y	Advanced RCC	Not stated	Immuno-oncology agent + systemic Rx, Sunitinib	Y	Patient-level	TTD and PFS TTD and OS	r = 0.80 r = 0.61	Not stated Not stated	Level 2 - Medium correlation Level 2 - Low correlation
2020	Mulkey et al.	N	Melanoma, NSCLC, RCC, HNSCC	14	Anti-PD-1/PD-L1 antibody	Y	Trial-level	iPFS and OS PFS and OS	r = 0.53 r = 0.51	Weak association	Level 1 - Low correlation
2020	Gong et al.	Y	Metastatic NSCLC	24	Immune checkpoint inhibitor, targeted therapy	Y	Patient-level	PFS, g = Q2 vs. g = Q1 OS, g = Q2 vs. g = Q1	HR = 2.3 HR = 1.8	g may be related to survival	Level 2 - Hazard ratio is significant
2019	Khozin et al.	N	Advanced NSCLC	Not based on trials, based on database of rw patients	PD-1/PD-L1 inhibitor	N/A	Patient-level	rwTTD and OS rwTTP and OS rwTTNT and OS rwPFS and OS	ρ = 0.81 ρ = 0.60 ρ = 0.60 ρ = 0.75	Moderately correlated Moderately correlated Moderately correlated Moderately correlated	Level 2 - Medium correlation Level 2 - Low correlation Level 2 - Low correlation Level 2 - Medium correlation
2019	Blumenthal et al.	N	Metastatic NSCLC	18	EGFR/ALK tyrosine kinase inhibitor, Immune checkpoint inhibitor, Chemotherapy	Y	Patient-level	TTD and PFS TTD and OS	r = 0.87 r = 0.68	Indicates an association Moderate association	Level 2 - High correlation Level 2 - Low correlation
2018	Mushti et al.	N	Melanoma, NSCLC, RCC, HNSCC	13	PD-1/PD-L1 inhibitor	Y	Trial-level	ORR and OS PFS and OS	r = 0.36 r = 0.36	Weak association	Level 1 - Low correlation
2018	Gao et al.	N	NSCLC, RCC, HNC	9	Anti-PD-1/PD-L1 antibody	Y	Trial-level	IME and OS	r = 0.82	Moderate association	Level 1 - Medium correlation
2018	Norsworthy et al.	N		8	Intensive chemotherapy	Y	Trial-level	CR and OS	r = 0.70	Moderate association	Level 1 - Low correlation

(*continued on next page*)

Table 1 (*continued*)

Year	Name	Abstract (Y/N)?	Cancer type	# Trials	Agents tested in trials	Only trials submitted to FDA? (Y/N)	Trial-level or patient-level?	Outcomes compared	Association	Study authors' interpretation	Our interpretation
			Newly diagnosed AML					EFS and OS	r = 0.93	Strong association	Level 1 - High correlation
2017	Blumenthal et al.[a]	N	Metastatic NSCLC	25	Crizotinib, Afatinib, Erlotinib, Cetuximab, Vandetanib, Gefitinib, Bevacizumab, Pemetrexed, Necitumumab, Ramucirumab, Nivolumab, Pembrolizumab, Atezolizumab Combination Rx	Y	Trial-level	6 mo ORR and OS HR 9 mo PFS milestone ratio and OS HR	r = 0.22 r = 0.44	No association	Level 1 - Low correlation
2016	Amiri-Kordestani et al.	Y	Metastatic breast cancer	13	Not stated	Y	Trial-level	CBR and PFS CBR and OS	r = 0.72 r = 0.10	Moderate association Little to no association	Level 1 - Medium correlation Level 1 - Low correlation
2015	Blumenthal et al.	N	Advanced NSCLC	14	Crizotinib, Afatinib, Erlotinib, Cetuximab, Vandetanib, Gefitinib, Bevacizumab, Pemetrexed, Combination Rx	Y	Trial-level	ORR and PFS ORR and OS PFS and OS	r = 0.94 r = 0.30 r = 0.28	Strong association No association	Level 1 - High correlation Level 1 - Low correlation
2014	Cortazar et al.	N	Breast cancer	12	Preoperative chemotherapy + surgery	N	Trial-level	PCR and EFS PCR and OS	r = 0.17 r = 0.49	Little association	Level 1 - Low correlation

Abbreviations: NSCLC, non-small cell lung cancer; HR-MDS, higher-risk myelodysplastic syndrome; RCC, renal cell carcinoma; HNSCC, head and neck squamous cell carcinoma; HNC, head and neck cancer; AML, acute myeloid leukemia; PFS, progression-free survival; OS, overall survival; ORR, overall response rate; CR, complete response; EFS, event-free survival; PR, partial remission; HI, hematologic improvement; EGFR, epidermal growth factor receptor; ALK, anaplastic lymphoma kinase; TTD, time to treatment discontinuation; iPFS, progression-free survival per iRECIST; g, tumor growth rate; rwPFS, real-world progression-free survival; rwTTP, real-world time to progression; rwTTNT, real-world time to next treatment; rwTTD, real-world time to discontinuation; CBR, clinical benefit rate; IME, intermediate response endpoint; PCR, pathological complete response.

[a] Extension of the 2015 analysis.

free survival, overall response rate, complete response, partial remission, hematologic improvement, time to treatment discontinuation, time to next treatment, time to progression, tumor growth rate, clinical benefit rate, pathological complete response, and a novel radiography-based intermediate response endpoint. For each study, we extracted the number of trials analyzed, the particular surrogates examined, their associations with OS or other surrogates, and the authors' assessment of these associations. Correlations reported as R^2 were converted to r values for ease of comparison.We also provided our interpretations of associations based on criteria published by the Institute for Quality and Efficiency in Health Care: low correlation if $r \leq 0.7$, medium correlation if $0.7 < r < 0.85$, and high correlation if $r \geq 0.85$ [4]. These criteria have been used previously in published literature. We interpreted hazard ratios as significant or not significant depending on whether 95% confidence intervals included or did not include 1. Furthermore, we categorized associations based on whether they pertained to patient-level or trial-level relationships. Level 1 (trial-level) evidence refers to correlations obtained through meta-analyses of randomized controlled trials [5]. Level 2 (patient-level) evidence is obtained from analyses performed on the level of the individual patient. While level 2 surrogacy may initially seem suitable for regulatory considerations, it is prognostic and does not address the pertinent question that faces regulators: if the product is approved based on this endpoint, are patients likely to benefit? Instead, it answers the tangential question of whether patients who achieve this endpoint do better compared to those who do not. This is a subtle yet important difference.

Correlations of $r = 0.68$ and $\rho = 0.60$ were reported as "moderate" although they fell below the accepted cutoff of 0.7 for a medium strength correlation. This study was not submitted for institutional review board approval because it involved publicly available data.

3. Results

We found 15 analyses of surrogates performed by the FDA, among which 5 (33%) were presented solely in abstract form, and all appeared after 2014. Table 1 summarizes the characteristics of each investigation and the reported associations between outcomes and surrogate markers. While one study analyzed real-world outcomes using a database of patients, the remaining 14 included 8–25 clinical trials. 13 (87%) studies included only trials submitted to the FDA in their analysis, and did not perform additional literature review.

All 15 studies quantified relationships between one or more surrogates and OS, at either the trial or patient level. The most examined surrogate was PFS, with 7 (47%) studies comparing OS to either median PFS, 9-month PFS milestone ratio, or real-world PFS. In all 6 trial-level meta-analyses, the correlation between PFS and OS was either weak or nonexistent.

The correlations between surrogate markers and OS were largely interpreted as weak to moderate. Only one study reported a strong correlation between an endpoint (event-free survival) and OS ($r = 0.93$).

4. Discussion

In our examination of surrogate validation studies performed by the US FDA, we found 15 studies from 2014 onwards, but none from 2005-2013, suggesting that the FDA's efforts to validate surrogates have occurred relatively recently. Most studies (87%) were limited by the inclusion of only clinical trials submitted to the FDA in their analysis, which can lead to false inferences around surrogate strength. In order to accurately understand the relationship between a surrogate and hard endpoint, it is vital to consider the totality of evidence and not merely trials chosen to be submitted to the US FDA, which often include the most favorable results and typically pertain solely to branded products.

Most analyses did not find strong relationships between putative surrogates and OS. In addition, several associations that were reported as "moderate" did not meet the typical $r > 0.7$ cutoff provided by the German quality and safety group, the Institute for Quality and Efficiency in Health Care, and should have been interpreted as weak.

Our results are in line with previous research on surrogate endpoints, which have generally found unknown or low associations between surrogate endpoints and overall survival. A systematic review of trial-level meta-analyses found that among 78 surrogate validation studies, only 11% reported high correlations between endpoints and survival [2]. Another study on 55 FDA cancer drug approvals based on a surrogate found that few formal analyses of surrogate strength were performed, and when they were, the majority reported low surrogate-survival correlations [1]. In addition, an umbrella review of surrogate validation studies reported that just 5 of 36 such studies considered unpublished trials, and even those only included half of eligible trials in their analysis [6], consistent with our findings that most validation studies are not based on the full set of available data.

The case of bevacizumab, which underwent accelerated approval by the FDA in 2008 for treatment of metastatic breast cancer on the basis of improvements in PFS, highlights the problem with using surrogates that have not been rigorously validated as the basis of drug approvals. After bevacizumab failed to demonstrate benefits in survival or quality of life, the FDA began to withdraw its indication for advanced breast cancer in 2010 [7]. However, the drug remained on the US market for almost a year until it was officially withdrawn, earning the sponsor billions.

Many drugs approved due to improvements in surrogate endpoints are not subsequently assessed for improvements in survival, which is problematic as their true benefits remain unknown [8]. As this trend has increasingly occurred and has been documented, the US FDA has steadily expanded the category of indications for which surrogates will be considered. Our analysis shows that when such studies are performed, they often fail to find strong correlations between surrogates and survival. In addition, some examine only level 2 or prognostic correlations and not level 1 or trial-level correlations, which are vital for regulatory and clinical consideration. Our results emphasize the need for robust surrogate validation studies and caution against the rapid expansion of unvalidated surrogates as endpoints in clinical trials.

Conflict of interest

Vinay Prasad's Disclosures. (Research funding) Arnold Ventures (Royalties) Johns Hopkins Press, Medscape, and MedPage (Honoraria) Grand Rounds/lectures from universities, medical centers, non-profits, and professional societies. (Consulting) UnitedHealthcare and OptumRX. (Other) Plenary Session podcast has Patreon backers, YouTube, and Substack. All other authors have no financial nor non-financial conflicts of interest to report.

References

[1] C. Kim, V. Prasad, Strength of validation for surrogate end points used in the US Food and Drug Administration's approval of oncology drugs, Mayo Clin. Proc. 91 (2016) 713–725, https://doi.org/10.1016/j.mayocp.2016.02.012.
[2] A. Haslam, S.P. Hey, J. Gill, V. Prasad, A systematic review of trial-level meta-analyses measuring the strength of association between surrogate end-points and overall survival in oncology, Eur. J. Cancer 106 (2019) 196–211, https://doi.org/10.1016/j.ejca.2018.11.012.
[3] E.Y. Chen, A. Haslam, V. Prasad, FDA acceptance of surrogate end points for cancer drug approval: 1992-2019, JAMA Intern. Med. 180 (2020) 912–914, https://doi.org/10.1001/jamainternmed.2020.1097.
[4] Institute for Quality and Efficiency in Health Care, Validity of surrogate endpoints in oncology: executive summary. https://www.iqwig.de/download/A10-05_Executive_Summary_v1-1_Surrogate_endpoints_in_oncology.pdf. (Accessed 9 September 2022).
[5] J. Elton, R.S. Taylor, Use of surrogate outcomes in cost-effectiveness models: A review of United Kingdom health technology assessment reports, Int. J. Technol. Assess. Health Care 25 (2009) 6–13, https://doi.org/10.1017/S0266462309090023.
[6] V. Prasad, C. Kim, M. Burotto, A. Vandross, The Strength of Association Between Surrogate End Points and Survival in Oncology: A Systematic Review of Trial-Level

Meta-analyses, JAMA Intern. Med. 175 (2015) 1389–1398, https://doi.org/10.1001/jamainternmed.2015.2829.

[7] D. Carpenter, A.S. Kesselheim, S. Joffe, Reputation and Precedent in the Bevacizumab Decision, N. Engl. J. Med. 365 (2011), https://doi.org/10.1056/NEJMp1107201.

[8] J.A. Beaver, L.J. Howie, L. Pelosof, et al., A 25-Year Experience of US Food and Drug Administration Accelerated Approval of Malignant Hematology and Oncology Drugs and Biologics: A Review, JAMA Oncol. 4 (2018) 849–856, https://doi.org/10.1001/jamaoncol.2017.5618.