European Journal of Cancer 106 (2019) 196–211

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Review

A systematic review of trial-level meta-analyses measuring the strength of association between surrogate end-points and overall survival in oncology

Alyson Haslam [a,*], Spencer P. Hey [b], Jennifer Gill [a], Vinay Prasad [c,d,e,f]

[a] *Knight Cancer Institute, Oregon Health & Science University, USA*
[b] *Program on Regulation, Therapeutics, and Law (PORTAL), Brigham and Women's Hospital, Center for Bioethics, Harvard Medical School, USA*
[c] *Division of Hematology Oncology, Knight Cancer Institute, Oregon Health & Science University, USA*
[d] *Department of Public Health and Preventive Medicine, Oregon Health & Science University, USA*
[e] *Senior Scholar in the Center for Health Care Ethics, Oregon Health & Science University, USA*
[f] *Division of General Medicine, Department of Medicine, Oregon Health & Science University, USA*

Received 26 October 2018; accepted 1 November 2018
Available online 5 December 2018

KEYWORDS
Surrogate outcome;
Oncology;
Validation study

Abstract ***Background:*** Surrogates are frequently used in cancer medicine as the end-point of clinical trials and as the basis of United States Food and Drug Administration approvals, but they do not always represent outcomes that are important for patients. We aim to build upon previous umbrella reviews of surrogate validation studies by identifying and examining all meta-analyses of randomised controlled trials that evaluate the strength of correlation between overall survival (OS) and surrogate markers.

Methods: Google Scholar and PubMed were searched by two independent reviewers for all eligible meta-analyses of randomised controlled trials examining the correlation between a surrogate end-point and OS in medical oncology. Included studies were trial-level (level-1) meta-analyses of randomised controlled trials in cancer. Data abstracted include date of publication, tumour type, setting, trial set, number of studies included in the analysis, dates of included publications, correlation coefficients and method to determine the correlation coefficient.

Results: Seventy-eight articles met the inclusion criteria and reported correlations in 89 settings. Eleven (12%) of these validation studies found only high correlation(s), while nine (10%) settings showed a moderate-only correlation. Thirty-four (38%) reported only low correlation(s). Thirty-five (39%) reported correlations of different strengths, depending on surrogate marker used and test of correlation.

* *Corresponding author*: Oregon Health & Science University 3181 SW Sam Jackson Park Road Portland, Oregon, 97239, USA. Tel.: +503 494 3159; Fax: +503 494 3257.
E-mail address: haslama@ohsu.edu (A. Haslam).

https://doi.org/10.1016/j.ejca.2018.11.012

Conclusions: In this large, umbrella analysis of surrogate validation studies, we found most surrogates in oncology had low or modest correlation with OS, which suggests that caution should be used when making conclusions based on surrogate markers.

1. Background

Surrogates are frequently used in cancer medicine as the end-point of clinical trials and as the basis of United States Food and Drug Administration approvals. Approximately two-thirds of all cancer drugs are approved based on improvements in surrogates, such as response rate or progression-free survival (PFS) [1]. Downing *et al*. found that among all pivotal trials supporting United States Food and Drug Administration approval between 2005 and 2012, 84% used surrogate end-points [2], and Kay *et al*. found that approximately two-thirds of randomised trials in cancer medicine rely on surrogate end-points [3].

Surrogate end-points provide the opportunity to speed the time it takes for cancer drugs to reach market or for clinical trial results to accrue. However, they also carry the potential downside of providing misleading inferences regarding the benefit of therapy. For instance, bevacizumab was approved in metastatic breast cancer on the basis of improvement in PFS (a surrogate) and later withdrawn from market when multiple randomised trials failed to confirm a survival benefit [4].

One method to validate surrogate end-points is to identify their correlation with clinically meaningful end-points, such as overall survival (OS). Measuring surrogate validity requires the use of meta-analyses of randomised controlled trials, calculating the correlation between the change in the surrogate end-point and change in OS across a series of clinical trials. These analyses, often called meta-regression or surrogate validation studies, use prior randomised trials in a specific disease setting (tumour type/line of therapy/setting of therapy) to characterise the value of the surrogate end-point in that setting. There is growing interest in performing such analyses.

Prior studies have examined surrogate validation studies in cancer medicine (i.e. umbrella reviews) [5,6]. Since these publications, there have been additional meta-analyses examining the correlation between OS and surrogate makers. Updating prior umbrella reviews could help to refine knowledge about the strength of association between surrogate markers and OS across a wider range of cancer types. It could also deepen our understanding of the specific nuances of the correlation between surrogate markers and OS (e.g. newer versus older treatments or setting of treatment, etc.). Furthermore, treatment options often change over time, potentially modifying the association between surrogate markers and more clinically meaningful end-points.

In this review, we aim at building upon previous umbrella reviews of surrogate validation studies [5,6] by identifying and examining all meta-analyses of randomised controlled trials that evaluate the strength of correlation between OS and surrogate markers. In addition to presenting findings from all years and in all settings, we will also characterise findings from more recent years and in specific settings of treatment.

2. Methods

We sought to assemble an updated and comprehensive collection of all trial-level surrogate validation studies in oncology. A surrogate validation study seeks to assess the correlation between a surrogate end-point (e.g. response rate or PFS) and an end-point that is inherently valuable to patients, such as OS or health-related quality of life. Trial-level validation is also referred to as level 1 validation and examines the extent of changes in the surrogate end-point across arms of clinical trials and how they predict changes in clinically meaningful end-points [7]. We limited our analysis to correlations between surrogates and survival, as other groups have examined correlations between surrogates and health related quality of life [8].

2.1. Study selection

Eligible studies were all meta-analyses of randomised controlled trials examining the association between a surrogate end-point and OS in medical oncology. There were no restrictions on the setting of medical therapy, including metastatic, neoadjuvant, adjuvant, or locally advanced. Immunotherapy correlations were included as a separate category, as it is increasingly appreciated that the relationships between surrogates and survival may be unique for this class of agents [9].

Only trial-level (level-1) analyses were included because this is the highest level of evidence for establishing the validity of a surrogate end-point [10,11] and was the standard used in a prior umbrella review [5]. We included studies that sought to compare any surrogate end-point (response rate, disease-free survival, etc.) against OS. Studies that only reported individual-level (level 2) correlations were excluded. We also excluded investigations containing non-randomised data. Studies

that did not use each trial as unique point (i.e. studies with multiple locations using each location as a unique point) were excluded. These criteria were adopted from a prior investigation [5].

2.2. Search strategy

We searched both Google Scholar and PubMed for relevant studies. Search terms used for MEDLINE were ('regression' or 'correlation') and 'surrogate' and 'end point' [or endpoint] and 'overall survival' and ('oncology' or 'cancer'). Search terms used for Google Scholar were ('regression' or 'correlation') and 'surrogate' and 'end point' [or endpoint] and 'overall survival' and 'trial level'. Additionally, we supplemented our results with studies included in three prior analyses [5,6,12]. Searches were performed on 25th January 2018, but an analysis that was published after that date provided additional studies to add to the present analysis [12]. Each article was screened by two reviewers (A.H. and J.G.).

2.3. Extracted data and statistical analysis

Data abstracted from each of the meta-analyses included date of publication, tumour type, setting, trial set, number of studies included in the analysis, dates of included publications, correlation coefficients and method to determine the correlation coefficient. In describing the trial set, we defined a convenience sample as a data source that was readily available to the study investigator in the absence of a systematic review.

The strength of association between surrogate endpoints and OS was graded using criteria proposed by the Institute of Quality and Efficiency in Health Care [13] and adapted by Prasad *et al.* [5]: low-strength correlation ($r \leq 0.70$), medium-strength correlation ($r > 0.7$ to $r < 0.85$), and high-strength correlation ($r \geq 0.85$). Descriptive statistics (number and percentage) were ascertained for the following: general characteristics of the studies included, studies and correlations in each setting, correlations in individual trial-level analyses and correlations in each treatment setting since 2015. Descriptive statistics were calculated in Microsoft Excel. All data for this analysis are presented in Table 1.

2.4. AERO graphing

An AERO graph was created to map the temporal data for trial-level surrogates in oncology [14]. For this analysis, we graphed date of surrogate study publication versus indication (i.e. tumour type and setting). The nodes were colour-coded according to the strength of correlation between surrogate outcome and OS. Studies that found a high level of correlation were coloured green, studies that found a medium level of correlation were coloured yellow, and studies that found a low level of correlation were coloured red. The size of the node corresponded to the number of trials included in each meta-analysis.

3. Results

Our search strategy yielded 227 results (Fig. 1). A recursive, manual search based on references in prior reviews identified an additional 12 studies. Seventy-eight articles met the inclusion criteria and reported 89 trial-level or level-1 correlations for one or multiple cancer settings. Data abstracted from each article are presented in Table 1.

Most studies reported correlations in the metastatic setting ($n = 59$; 76%). Ten articles (13%) reported correlations in the adjuvant setting, six (8%) in the neoadjuvant, three studies (4%) reported correlations in the locally advanced, and three articles (4%) examined immunotherapy agents. We grouped these separately as the relationship between surrogates and survival may be unique among this class of drugs [9]. Thirty-seven (47%) articles were published in 2015 or thereafter.

Twenty-three (29%) studies drew correlations from published articles and abstracts; 21 (27%) studies used published articles only; 16 (21%) used articles, meeting abstracts, and searched trial registries; seven (9%) used articles and registries; three (4%) did not provide or clearly provide data sources; five (6%) used a convenience sample; and three (4%) used 'other' data sources (e.g. trials). One study was published in 2000; none were published between 2001 and 2004 and during 2010; 2–5 studies were published per year during years 2005–2009, 2011–2012, and 2018; 11 per year were published during years 2013 and 2016, while nine were published each year for years 2014 and 2015; and 13 were published in 2017.

In total, correlations were reported in 89 settings, as some articles presented correlations in multiple settings. A map of the evolution of evidence for trial-level surrogates (AERO diagram) is shown in Fig. 2. Most of these correlations were in the metastatic setting ($n = 64$), where 16 reported on colorectal cancer, nine reported on non–small-cell lung cancer (NSCLC), and eight reported on breast cancer, among other cancer types.

Eleven (12%) of these validation studies found only high correlation(s), while nine (10%) settings showed a moderate-only correlation (see Table 1). Thirty-four (38%) reported only low correlation(s). Thirty-five (39%) reported correlations of different strengths, depending on surrogate marker used and test of correlation. When considering the treatment setting (all years), only six (9%) of the 65 studies reporting in the metastatic setting were high correlation only; eight (12%) were moderate strength only; 23 (35%) were low strength only; and 28 (43%) had correlations of different strengths.

Table 1
Summary of characteristics and correlations in trial-level surrogate meta-analyses examining the correlation between surrogate end-points and overall survival in oncology.

Primary author name	Tumour	Trial set	Number of trials examined	Years included	1st Trial-level correlation	1st Correlation description	2nd Trial-level correlation	2nd Correlation description	3rd Trial−Level correlation	3rd Correlation description	Surrogacy strength
Adjuvant setting											
Ng *et al.*, 2008 [21]	Breast	Convenience sample	126	Published 1966−2006	0.62	DFS					Low
Buyse *et al.*, 2008 [22]	Colorectal	Not well described	10	Not well described	0.94	DFS, early disease, copula, logHR	0.82	DFS, advanced disease, copula, logHR			High (early) Medium (advanced)
Sargent *et al.*, 2005 [23]	Colon	Convenience sample	18	Accrual period 1977−1999	0.78	DFS, bivariate survival model	0.94	DFS, weighted, Spearman			Medium (linear) High (weighted)
Sargent *et al.*, 2011 [24]	Colon	6 Phase III adjuvant colon clinical trials not included in prior ACCENT analyses	6	Accrued period 1997−2002	0.76	DFS, 2-year DFS and 5-year OS, weighted	0.61	DFS, copula 2-year DFS on 5-year OS, copula			Low (copula) Medium (weighted)
Oba *et al.*, 2013 [25]	Gastric	Articles and registries	14	Close to accrual in 2004	0.98	DFS, Spearman, logHR					High
Harshman *et al.*, 2017 [26]	Renal cell	Articles, abstracts and registries	11	Published 1966−2016	0.69	DFS	0.66	PFS (logHR)			Low
Suciu *et al.*, 2017 [27]	Melanoma	Articles	13	Published 1996−2008	0.49	RFS (13 trials)	0.96	RFS (11 trials, excluding the trials using vaccine)	0.95	RFS (excluding ECOG 2696 trial)	Low (13 trials) High (fewer)
Michiels *et al.*, 2009 [28]	Head and neck	Articles and abstracts	9	Accrual period 1965−2000	0.84	LRC, weighted, HR	0.93	EFS, weighted, HR			High (EFS) Medium (LRC)
Mauguen *et al.*, 2013 [29]	Lung (operable and locally advanced)	Articles, abstracts and registries	17	Accrual period 1988−2003	0.96	DFS, bivariate survival model					High
Petrelli *et al.*, 2017 [30]	Pancreatic	Articles and registries	12	Published before February 2017	0.66	DFS, weighted, HR					Low
Locally advanced											
Michiels *et al.*, 2009 [28]	Head and neck	Articles and abstracts	50	Accrual period 1965−2000	0.72	LRC	0.86	EFS			Medium (LRC) High (EFS)
Mauguen *et al.*, 2013 [29]	Lung (operable and locally advanced), concurrent	Articles, abstracts and registries	15	Accrual period 1984−2001	0.98	PFS, weighted, logHR					High
Mauguen *et al.*, 2013 [29]	Lung (operable and advanced), sequential	Articles, abstracts and registries	8	Accrual period 1984−1999	0.98	PFS, weighted, logHR					High
Chen *et al.*, 2015 [94]	Nasopharyngeal	Articles, abstracts and registries	21	Published before 2015	0.95	PFS, weighted linear regression (logHR)	0.94	Failure-free survival, weighted linear regression (logHR)			High
Metastatic											
Moriwaki *et al.*, 2016 [31]	Biliary tract (advanced)	Articles and abstracts	17	Published through February 2015	0.81	PFS, weighted linear regression	0.54	RR, weighted linear regression	0.58	DCR, weighted linear regression	Medium (PFS) Low (RR and DCR)

(*continued on next page*)

Table 1 (*continued*)

Primary author name	Tumour	Trial set	Number of trials examined	Years included	1st Trial-level correlation	1st Correlation description	2nd Trial-level correlation	2nd Correlation description	3rd Trial−Level correlation	3rd Correlation description	Surrogacy strength
Bruzzi *et al.*, 2005 [32]	Breast	Articles	10	Published 1991 −2001	0.32	RR, logOR	0.45	RR, delta			Low
Hackshaw, 2005 [33]	Breast, advanced	Articles	46 for RR 26 for PFS	Published 1966 −2005	0.58	Tumour response and disease progression, weighted linear regression (logHR)	0.35	pCR, weighted linear regression (logHR)	0.75	TTP, weighted linear regression (logHR)	Low (TR and pCR) Medium (TTP)
Burzykowski *et al.*, 2008 [34]	Breast	Convenience sample	11	Published 1999 −2008	0.69/0.48	PFS HR (Spearman correlation/ copula) (end-point versus treatment effects)	0.68/0.49	TTP (Spearman/ copula)			Low
Miksad, 2008 [35]	Breast, anthracycline	Articles and abstracts	16	Published 1987 −2005	0.59	HR (log) PFS					Low
Miksad, 2008 [35]	Breast, taxane	Articles and abstracts	15	Published 1996 −2006	0.70	HR (log) PFS					Low
Sherrill *et al.*, 2008 [36]	Breast	Articles	67	Published 1994 −2007	0.55	TTP (HR), weighted					Low
Wilkerson and Fojo, 2009 [37]	Breast	Articles	14	Published 2000 −2008	0.55	PFS (delta)	0.88	PFS (HR)			Low (delta) High (HR)
Beauchemin *et al.*, 2014 [38]	Breast	Articles, abstracts and registries	144	Published 1990 −2010	0.43	PFS/TTP (unweighted Spearman)	0.52	PFS (Spearman)	0.31	TTP (Spearman)	Low
Petrelli and Barni, 2014 [39]	Breast	Articles	21	2000−2012	0.81	PFS	0.61	RR			High (PFS) Low (RR)
Adunlin *et al.*, 2015 [40]	Breast (anthracyclines, taxanes, or targeted therapies)	Articles, abstracts and registries	72	Published 1990 −2015	0.46/0.56	PFS, HR univariate Spearman/ weighted multivariate	0.52/0.66	PFS delta univariate Spearman/ weighted multivariate			Low
Michiels *et al.*, 2016 [41]	Breast (HER2-positive)	Articles	9	Published 1992 −2008	0.71	PFS, logHR, weighted linear regression	0.81	PFS, weighted			Medium
Liu *et al.*, 2016 [42]	Breast (2nd or 3rd line)	Articles and abstracts	24	Published 1999 −2014	0.78	PFS/TTP, Spearman's rank	0.54	ORR, Spearman's rank	0.57	PFS/TTP, HR Spearman's rank	Low (ORR and PFS/TTP, HR) Medium (PFS/ TTP, Spearman)
Buyse *et al.*, 2000 [43]	Colorectal, advanced, first line	Convenience sample	25	Data collected 1990−1996	0.62	Tumour response, log HR, linear					Low
Johnson *et al.*, 2006 [44]	Colorectal	Articles and abstracts	146	Before 1992 −2005	0.57	TTP (delta) weighted	0.32	RR (delta) weighted			Low
Buyse *et al.*, 2007 [45]	Colorectal (advanced), first line	Not well described	13	Published 1981 −1990	0.99	PFS, log (HR) (treatment effects)	0.82	PFS, Spearman correlation			High (PFS, logHR) Medium (PFS, Spearman)
Tang *et al.*, 2007 [46]	Colorectal (metastatic), first line	Articles, abstracts and registries	39	Published 1990 −2005	0.59/0.39	RR (rho/delta) (across/within)	0.24/0.52	TTP (rho/delta)	0.79/0.74	PFS (rho/delta)	Low (RR, TTP) Medium (PFS)

Wilkerson and Fojo, 2009 [37]	Colorectal	Articles	23	Published 1993–2009	0.80	PFS (delta)	0.72	PFS (HR)			Medium
Chirila *et al.*, 2012 [47]	Colorectal	Articles and abstracts	62 PFS/TTP 35 PFS	1991–2009	0.87	PFS/TTP, simple correlation	0.69	PFS/TTP, regression	0.77	PFS, regression	Low (PFS/TTP) Medium (PFS) High (PFS/TTP)
Giessen *et al.*, 2013 [48]	Colorectal, chemo only, first line	Articles and abstracts	50	Published 2000–2012	0.87	PFS, weighted linear					High
Giessen *et al.*, 2013 [48]	Colorectal, monoclonal antibodies, first line	Articles and abstracts	19	Published 2000–2012	0.47	PFS, weighted linear					Low
Petrelli and Barni, 2013 [49]	Colorectal	Articles, abstracts and registries	34	Published before 2012	0.64	Median PFS, Spearman	0.88	PPS, Spearman	0.59	PFS delta, Spearman	Low (PFS) High (PPS)
Sidhu, 2013 [50]	Colorectal	Articles and abstracts	24	Published 2000–2011	0.86	PFS, weighted linear regression, logHR	0.62	RR			High (PFS) Medium (RR)
Giessen *et al.*, 2015 [51]	Colorectal, second line	Articles and abstracts	23	Published 2000–2013	0.73	PFS, weighted Pearson's	0.58	RR, weighted	0.74	DCR, weighted Pearson's	Medium (PFS and DCR) Low (RR)
Shi *et al.*, 2014 [52]	Colorectal, first line	Convenience sample	22	Published 1997–2006	0.68	PFS, copula logHR	0.73	PFS, weighted logHR			Low (copula) Medium (weighted)
Ciani *et al.*, 2015 [53]	Colorectal (advanced)	Articles and registries	101	Published 2003–2013	0.75/0.58	PFS Spearman/ weighted linear, HR	0.80/0.81	TTP Spearman/ weighted linear, HR			Low (PFS, weighted) Medium (others)
Colloca *et al.*, 2016 [54]	Colorectal	Articles and abstracts	11	Published 2000–2014	0.84	PFS, linear, delta	0.76	RR, linear, delta	0.75	DCR, linear, delta	Medium
Montagnani *et al.*, 2016 [55]	Colorectal (advanced)	Articles and abstracts	12	Published 2003–2015	0.82	PFS, weighted logHR	0.78	PFS, unweighted logHR			Medium
Cremolini *et al.*, 2017 [56]	Colorectal	Articles and abstracts	21 (PFS) 20 (RR)	Published before October 2015	0.73	PFS (HR), weighted Pearson's	0.17	RR, weighted Pearson's	0.66/0.86	PFS (HR) for anti-angiogenic/ 'other' drugs, weighted Pearson's	Low (RR) Medium (PFS, Pearson's) High (PFS, 'other' drugs)
Zhu *et al.*, 2017 [57]	Diffuse large B-cell lymphoma	Articles and abstracts	108	Published 1993–2013	0.90/0.73	PFS, 6-month PFS (2-year OS/5-year OS), weighted linear regression					High (2 years) Medium (5 years)
Kawakami *et al.*, 2013 [58]	Gastric	Articles and abstracts	43	Published 1990–2011	0.50	PFS, Spearman	0.73	PPS, Spearman	0.55	PFS delta, Spearman	Low (PFS) Medium (PPS)
Paoletti *et al.*, 2013 [59]	Gastric, advanced/ recurrent	Articles	20	Accrual period through 2006	0.78	PFS, linear regression (logHR)					Medium
Shitara *et al.*, 2014 [60]	Gastric, advanced, second line	Articles and abstracts	10	Published 2002–2013	0.36	PFS/TTP, HR Spearman's rank					Low
Lee, 2016 [61]	Hepatocellular	Articles	9	Published 2008–2015	0.73/0.64	TTP, Spearman/ weighted linear regression					Medium (Spearman) Low (weighted)
Flaherty *et al.*, 2014 [62]	Metastatic melanoma	Articles, abstracts and registries	12	Published 2006–2013	0.89	PFS, weighted Pearson's (logHR)					High

(*continued on next page*)

Table 1 (*continued*)

Primary author name	Tumour	Trial set	Number of trials examined	Years included	1st Trial-level correlation	1st Correlation description	2nd Trial-level correlation	2nd Correlation description	3rd Trial−Level correlation	3rd Correlation description	Surrogacy strength
Cartier *et al.*, 2015 [63]	Multiple myeloma	Articles and registries	21	Published 2002−2013	0.80	PFS (logHR)	0.82	PFS (HR)			Medium
Rotolo *et al.*, 2017 [64]	Nasopharyngeal	Articles, abstracts and registries	20	Accrual period 1988−2010	0.97	PFS, bivariate survival model (logHR)	0.98	Distant metastasis-free survival			High
Zhu *et al.*, 2017 [57]	Non-Hodgkin's lymphoma, mantle cell lymphoma	Articles and abstracts	108	Published 1993−2013	0.83	PFS, 6-month PFS (2-year OS), weighted linear regression					Medium
Lee *et al.*, 2011 [65]	Non-Hodgkin's lymphoma (aggressive)	Articles, abstracts and registries	17	Published 1990−2009	0.58/0.41	pCR (3-year OS) rho for aggressive/indolent	0.5/0.21	pCR (5-year OS) rho for aggressive/indolent	0.81	EFS/PFS (5-year OS) for aggressive, linear regression	Low (pCR) Medium (EFS/PFS)
Johnson *et al.*, 2006 [44]	NSCLC	Articles and abstracts	191	Before 1992−2005	0.44	TTP (delta) weighted	0.40	RR, (delta) weighted			Low
Hotta *et al.*, 2009 [66]	NSCLC (advanced)	Articles, abstracts and registries	54	Published 1994−2006	0.57	TTP, weighted					Low
Hotta *et al.*, 2011 [67]	NSCLC (advanced)	Articles and abstracts	70	Published 1991−2010	0.51	PFS and median survival time, weighted linear regression					Low
Petrelli and Barni, 2013 [68]	NSCLC, maintenance	Articles, abstracts and registries	10	Published before 2012	0.37	PFS, Spearman	0.75	PPS, Spearman	0.64	PFS delta, Spearman	Low (PFS) Medium (PPS)
Hotta *et al.*, 2013 [69]	NSCLC	Articles, abstracts and registries	34	Published 1991−2010	0.38	PFS, weighted linear regression					Low
Aboshi *et al.*, 2014 [70]	NSCLC	Articles	65	Published 2003−2012	0.69	PFS, Spearman	0.84	PPS, Spearman			Low (PFS) Medium (PPS)
Hotta *et al.*, 2015 [71]	NSCLC, phase 3 trials	Articles, abstracts and registries	18		0.48	PFS-HR, weighted linear regression	0.32	RR, weighted linear regression			Low
Suzuki *et al.*, 2015 [72]	NSCLC, second line and further	Articles	32	Published 2000−2012	0.38	PFS, Spearman, weighted	0.42	PFS (HR), Spearman, weighted			Low
Nakashima *et al.*, 2016 [73]	NSCLC	Articles	44	Published 2005−2015	0.47	PFS, (logHR; delta PFS: 0.40; simple r = 0.48)	0.59	RR, (logOR; simple r = 0.48)	0.42	DCR (logOR; simple r = 0.47)	Low
Wilkerson and Fojo, 2009 [37]	Ovarian	Articles	15	Published 1983−2008	0.77	PFS (delta)	0.85	PFS (HR)			Medium (delta) High (HR)
Colloca and Venturino, 2017 [74]	Ovarian (advanced)	Articles	38	1990−2016	0.71	PFS (logHR)	0.55	PFS, Spearman	0.41	RR, Spearman	Medium (PFS, linear) Low (PFS and RR, Spearman)
Siddiqui *et al.*, 2017 [75]	Ovarian, second line	Articles and abstracts	31	Published 2000−2015	0.82	RR, weighted linear regression	0.73	DCR, weighted linear regression			Medium (RR) Low (DCR)
Petrelli *et al.*, 2015 [76]	Pancreatic	Articles, abstracts and registries	18	Published 2002−2013	0.78	PFS/TTP					Medium

Hamada *et al.*, 2016 [77]	Pancreatic (advanced)	Articles and registries	50	Published 1995–2015	0.92	PFS (HR) weighted	0.55	RR (OR) weighted	0.81	DCR (OR) weighted	High (PFS) Medium (DCR) Low (RR)
Colloca *et al.*, 2016 [78]	Pancreatic (advanced) gemcitabine-based treatment versus gemcitabine alone	Articles	22 (chemo) + 15 (targeted therapy)	Published 1997–2014	0.67/0.80	PFS, Spearman's/ linear (logHR)	0.29/0.38	RR, Spearman's/ linear (logHR)	0.62/0.75	DCR, Spearman's/ linear (logHR)	Low (PFS, Spearman; RR) Medium (PFS, linear; DCR)
Makris *et al.*, 2017 [79]	Pancreatic (metastatic, locally advanced, or unresectable)	Articles and abstracts	24	Published 2000–2015	0.85/0.83/0.86	DFS, weighted/ fixed effects/ random effects (logHR)	0.71/0.74/0.74	DCR, weighted/ fixed effects/ random effects (logHR)	0.27/0.52/0.45	RR, weighted/ fixed effects/ random effects (logHR)	Medium (DFS and DCR) Low (RR)
Xie *et al.*, 2017 [80]	Prostate (localised)	Not well described	24 (DFS); 19 (MFS)	Accrual 1987–2011	0.85	DFS, weighted (logHR)	0.96	metastasis-free survival, weighted (logHR)			High
Delea *et al.*, 2012 [81]	Renal cell	Articles and abstracts	31	Published 1997–2010	0.80	PFS/TTP, weighted Pearson correlation (logHR)	0.54	PFS/TTP delta, weighted Pearson correlation	0.78	RR, weighted Pearson's (log RR)	Low (PFS/TTP, delta) Medium (RR, PSF/TTP, logHR)
Petrelli and Barni., 2013 [82]	Renal cell	Articles	6	Published before 2011	0.87	PFS (not delta)	0.36/0.26	PFS, delta, Spearman's /linear	0.49/0.52	Delta RR, Spearman's /linear	Low (RR, PFS, delta) High (PFS, not delta)
Bria *et al.*, 2015 [83]	Renal cell, targeted	Articles and abstracts	10	Published before 2013	0.85	PFS, Pearson's					High
Johnson *et al.*, 2015 [84]	Renal cell	Articles and registries	11	Published 1946–2013	0.70	PFS, weighted errors in variable regression	0.66	PFS, ordinary linear regression			Low
Hotta *et al.*, 2009 [85]	SCLC	Articles, abstracts and registries	48	Published 1990–2008	0.58	RR and median survival time, multivariable adjusted regression					Low
Foster *et al.*, 2015 [86]	SCLC	Articles	10	Accrual 1982–2007	0.93	PFS (platinum/ etoposide treatment combination; copula); weighted 0.91	0.90	PFS, (all 10 trials; Copula); 0.88 weighted			High
Roviello *et al.*, 2017 [87]	Solid tumours (non-SCLC, melanoma, SCLC, renal, prostate)	Articles and registries	17	Published before March 2017	0.69	RR (solid tumours), weighted, (logHR)	0.03	RR (NSCLC)			Low
Amir *et al.*, 2012 [88]	Solid tumours with treatments approved by the United States Food and Drug Administration in the 10 years prior	FDA-approved therapies ≥12 months SPP	7	Published 2002–2012	0.38	PFS, weighted Pearson's					Low

(*continued on next page*)

Table 1 (*continued*)

Primary author name	Tumour	Trial set	Number of trials examined	Years included	1st Trial-level correlation	1st Correlation description	2nd Trial-level correlation	2nd Correlation description	3rd Trial−Level correlation	3rd Correlation description	Surrogacy strength
Amir *et al.*, 2012 [88]	Solid tumours with treatments approved by the United States Food and Drug Administration in the 10 years prior	FDA-approved therapies <12 months SPP	19	Published 2002−2012	0.64	PFS, weighted Pearson's					Low
Neoadjuvant											
Cortazar, 2014 [89]	Breast	Articles	12	Published 1990−2011	0.49	pCR, (HR) weighted linear regression					Low
Berruti *et al.*, 2014 [32]	Breast	Articles and abstracts	29	Published before March 2013	0.30	pCR, logOR(pCR) on logHR(OS); weighted					Low
Kataoka *et al.*, 2017 [90]	Oesophageal	Articles and abstracts	10	Published 1990−2014	0.53	PFS (HR), weighted linear regression					Low
Petrelli *et al.*, 2017 [91]	Gastro-oesophageal	Articles	16	Published before August 2016	0.52	DFS/PFS (logHR)	0.41	pCR (logOR; 15 trials)			Low
Michiels *et al.*, 2009 [28]	Head and neck	Articles and abstracts	30	Accrual period 1965−2000	0.53	LRC	0.79	EFS			Medium (EFS) Low (DCR)
Petrelli *et al.*, 2017 [30]	Rectal	Articles	9 DFS/13 pCR	Published before August 2015	0.64	PFS, delta Spearman (3-year DFS and 5-year OS)	0.20	pCR, delta Spearman			Low
Immunotherapy											
Bria *et al.*, 2015 [83]	Renal cell, targeted	Articles and abstracts	10	Published before 2013	0.84	PFS, Pearson's					Medium
Mushti *et al.*, 2018 [92]	Not specific	Immunotherapy trials of anti-PD1/PD-L1 67 agents submitted to the FDA	13	2014−2016	0.13	PFS	0.13	RR			low
Abdel-Rahman, 2018 [93]	Renal cell	Articles	4	Published through 2017	0.40	RR, Spearman	0.39	PFS, Spearman			Low
Abdel-Rahman, 2018 [93]	Urothelial	Articles	9	Published through 2017	0.12	RR, Spearman					Low

DCR = disease control rate; DFS = disease-free survival; ECOG = Eastern Cooperative Oncology Group; EFS = event-free survival; FDA = Food and Drug Administration; HR = hazard ratio; LRC = locoregional control; MFS = metastasis-free survival; NSCLC = non−small-cell lung cancer; OR = odds ratio; ORR = overall response rate; OS = overall survival; PFS = progression-free survival; PPS = post-progression survival; pCR = partial complete response; RFS = relapse-free survival; RR = response rate; SCLC = small-cell lung cancer; SPP = survival post-progression; TTP = time to progression.

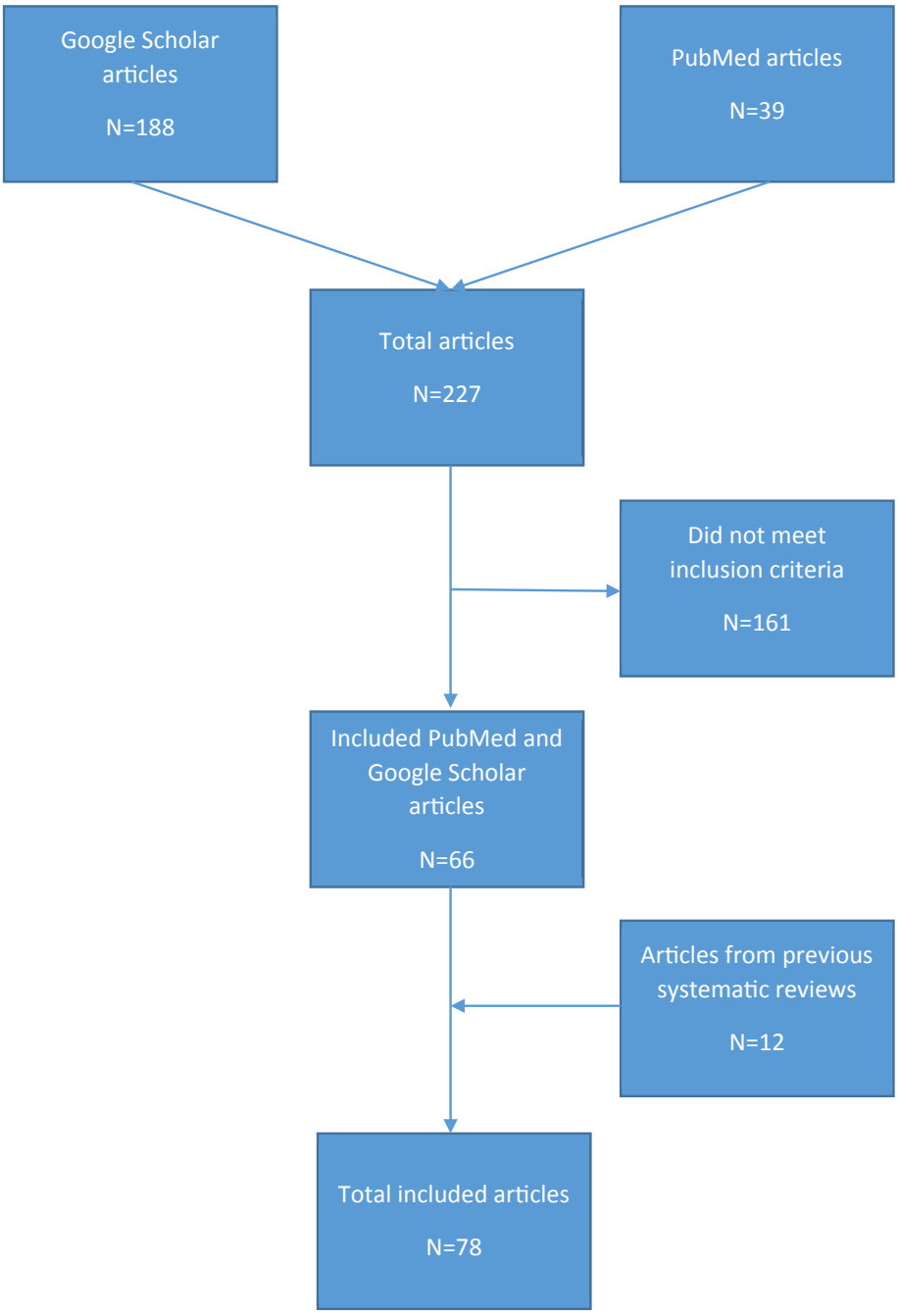

Fig. 1. Consort flow diagram of the process to determine which studies were included for the strength of correlation analysis.

For the ten correlations in the adjuvant setting, two (20%) were high only, none were moderate only, three (30%) were low only, and five (50%) reported correlations of different strengths. In the four studies reporting in the locally advanced setting, three had high correlations and one had different strengths of correlation. Most correlations in the neoadjuvant setting were low ($n = 5$; 83%) and only one (17%) had correlations of different strengths. Three correlations in the immunotherapy setting were low, while one was moderate strength.

When restricting to the 37 settings from articles published in 2015 or later, the percentages were similar: four studies (11%) had high correlation only, seven (19%) had moderate correlation, 12 (32%) had low correlation only, and 14 (38%) had correlations of different strengths.

There were 193 individual trial-level correlations reported, as some reported different methods to calculate correlation and used different surrogate markers (Fig. 3). Of these, 34 (18%) were high correlation, 55 (28%) were moderate correlation, and 104 (54%) had

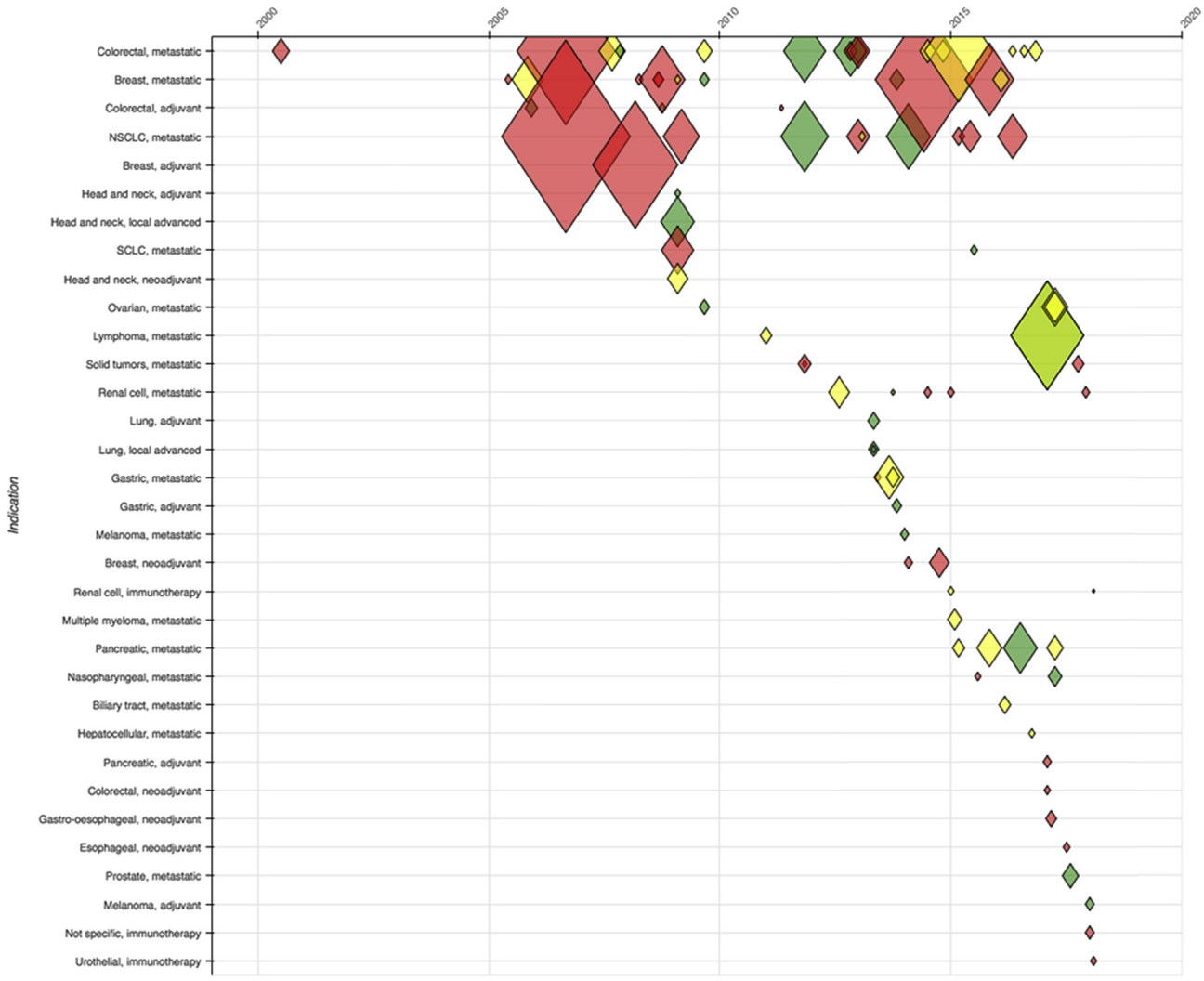

Fig. 2. Accumulating Evidence and Research Organization (AERO) graph of 81 trial-level surrogate validation studies. Each node in the figure corresponds to a single correlation analysis (some publications included multiple analyses), arranged by publication date on the *x*-axis and cancer/setting on the *y*-axis. The node size corresponds to the number of trials included in each meta-analysis, and the node colour represents the strength of correlation between a surrogate outcome and overall survival: green corresponds to high level of correlation, yellow to medium, and red to a low level of correlation. In studies that used multiple methods of determining correlation, the strongest correlation is used for colour classification. Abbreviations: NSCLC = non–small-cell lung cancer; SCLC = small-cell lung cancer.

low correlation (Fig. 2). When looking at the most common surrogate markers, regardless of cancer type, correlation was low in 48% (n = 40/83) of PFS individual trial-level analyses, moderate for 33% (n = 27/83) analyses, and high for 19% (n = 16/83) analyses. Conversely, for the 32 individual trial-level analyses that used response rate, 29 (91%) were low correlation, three (9%) were moderate, and none were high.

4. Discussion

4.1. Principle findings

In this, the largest and most updated umbrella analysis of all published correlation studies, we find that about 82% of the correlations between surrogate markers and OS in oncology are low or moderate in strength, and only about 18% are strong. The largest prior study found 36 articles that assessed the correlation between surrogate markers and OS, but the present study found 78 articles that assessed this correlation. Another very recently published study was comprehensive but also included studies that were not limited to trial-level data or randomised trials, which are not considered as strong of evidence [12]. In this comprehensive analysis, we find that most surrogates in oncology only weakly predict survival.

A notable proportion of cancer drugs are approved on surrogate markers [1,2], and the proportion is even higher among cancer drugs approved with the

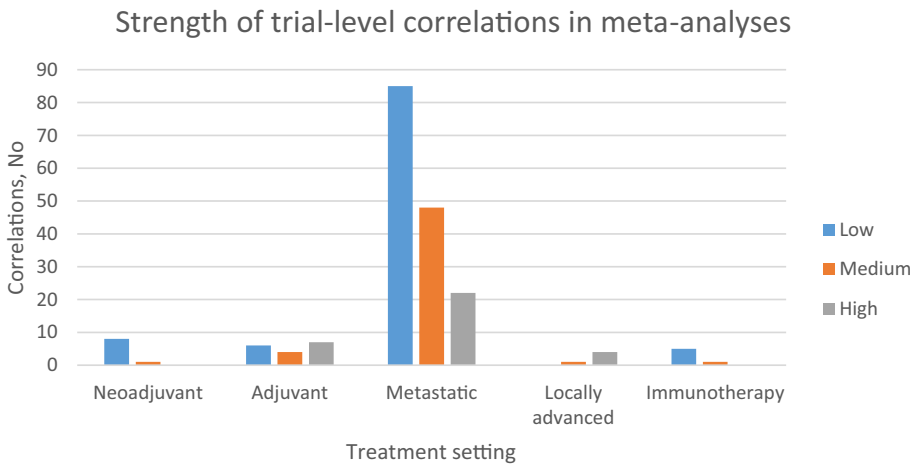

Fig. 3. Distributions of correlations by treatment setting. Strength of trial-level correlation was scored according to a modification to surrogate criteria proposed by the Institute of Quality and Efficiency in Health Care: low correlation ($r \leq 0.7$), medium-strength correlation ($r > 0.7$ to <0.85), and high correlation ($r \geq 0.85$).

breakthrough designation, where 96% relied on a surrogate end-point [15]. Thus, the findings of our article suggest that large clinical trial agendas in oncology may be using end-points that only weakly predict improved survival.

Weak surrogate end-points may be particularly concerning considering that many drugs that are approved are not later assessed for OS. For a number of these drugs, OS is not shown, either because post-market studies fail to report these findings or because the drug may not extend survival [1,16,17]. The use of weak surrogates is also concerning because common indications may be more likely to be affected by the approval of drugs based on weak surrogates than other less common indications. For example, in the present analysis, almost one-third of studies in the metastatic setting that had low-only correlations were in studies on breast cancer, whereas only about 17% of all included studies were on the topic of breast cancer.

It is notable that immunotherapeutic drugs appear to have low correlation between surrogate markers and OS. Consider pembrolizumab that, as of 8th August 2018, has 12 approvals, seven of which are based on a surrogate end-point [18]. While the studies included in this meta-analysis represent a small segment of the uses of immunology drugs, when examined, these drugs had largely poor correlations between surrogate markers and OS.

The major strength of our study is that it looks at these correlations over a wide variety of settings and cancer types. While the number of different settings and cancer types is large and making direct comparisons is not very meaningful, the distribution of low, medium and high correlations across the different settings has been consistent with prior reports.

As can be seen in the AERO graph, the majority of indications have only one study, and the correlation is usually small. For many cancers, this is what we would expect, because there is not a high volume of research activity to do a large trial-level meta-analyses. By contrast, NSCLC has many large correlation studies with discordant results. This may be explained by the fact that there are many NSCLC lung cancer trials in diverse settings. However, given that the datasets for these studies should overlap (at least in part), this may be a setting where we have enough evidence to come to a reasonably definitive conclusion about the correlation between surrogate markers and survival: namely, that they are poor trial-level surrogates and should not be used for approval.

4.2. Strengths and limitations

The major strength of our study is that it looks at these correlations over a wide variety of settings and cancer types. While the number of different settings and cancer types is large and making direct comparisons is not very meaningful, the distribution of low, medium, and high correlations across the different settings has been consistent with prior reports. Another strength is that this analysis uses only trial-level correlations and meta-analyses from randomised controlled trials, both considered highest quality evidence.

There are at least four limitations to this study. First, because of different methodology used to determine correlation in each of the studies, the results may not always be directly comparable, so we presented the raw correlation coefficients reported. Different methods or analysis plans could influence the strength of association. For example, when multiple methods to determine correlation were presented, weighted correlation coefficients were almost always higher than non-weighted models. For models that used the delta hazard ratios, the correlation coefficient was often, but not always,

lower than the other models. The concern here is that multiple analysis plans may be pursued and strongest results reported [19].

The second limitation is that different researchers may disagree on what correlation is strong, moderate or weak. We used a scaling system adapted from the German Institute for Quality and Efficiency in Health Care group, but others have proposed more stringent thresholds. For example, some authors have proposed a threshold of 0.9 for determining high correlation [13]. We report all the raw data in the article to let readers interpret these coefficients in other ways.

The third limitation is that of crossover, which may modify the relationship between the surrogate and survival [20]. While we were unable to consider studies with and without crossover separately—as these were not separated in the primary publications—we nevertheless feel it is unlikely that crossover would have impacted the results of this analysis, as many studies were conducted before the widespread use of crossover in study design. Moreover, more recent meta-analyses would arguably include more crossover studies, yet the correlations seen here were not remarkably different than in prior articles.

Finally, by using multiple individual trial-level correlations, the number of correlations may be inflated for each strength-level category. However, this inflation is unlikely to affect the relative proportion of articles classified as weak, moderate or strong. We have presented all reported associations to allow the reader to conclude which correlation is most relevant (e.g. delta versus hazard ratio).

5. Conclusion

In this large, umbrella analysis of surrogate validation studies, we found most surrogates in oncology had low or modest correlation with OS. Given the widespread use of surrogates in regulatory decisions and clinical practice guidance, our results provide a cautionary note.

Funding/support

Dr. Prasad reports receiving royalties from his book Ending Medical Reversal, that his work is funded by the Laura and John Arnold Foundation, that he has received honoraria for Grand Rounds/lectures from several universities, medical centres, non-profit groups, professional societies and that he is a writer for Medscape. Dr. Hey receives funding from the Laura and John Arnold Foundation. Dr. Haslam and Ms. Gill have no funding to report.

Author contribution

A.H. helped in design of study, data abstraction, data analysis and drafting of article; S.P.H. contributed to data analysis and drafting of article; J.G. helped in data abstraction and drafting of article; and V.P. assisted in design of study and drafting of article. All authors had access to the data. A.H. is the guarantor.

Conflict of interest disclosures

The authors have declared no conflicts of interest.

References

[1] 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 Int Med 2015;175(12):1992–4.

[2] Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005-2012. J Am Med Assoc 2014;311(4):368–77.

[3] Kay A, Higgins J, Day A, Meyer R, Booth C. Randomized controlled trials in the era of molecular oncology: methodology, biomarkers, and end points. Ann Oncol 2011;23(6):1646–51.

[4] Carpenter D, Kesselheim AS, Joffe S. Reputation and precedent in the bevacizumab decision. N Engl J Med 2011;365(2):e3.

[5] 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 Int Med 2015;175(8):1389–98.

[6] Sherrill B, Kaye JA, Sandin R, Cappelleri JC, Chen C. Review of meta-analyses evaluating surrogate endpoints for overall survival in oncology. Onco Targets Ther 2012;5:287.

[7] Little RF, McShane LM, Freidlin B. Myeloma minimal residual disease and surrogacy. JAMA Oncol 2017;3(8). 1136-1136.

[8] Kovic B, Guyatt G, Brundage M, Thabane L, Bhatnagar N, Xie F. Association between progression-free survival and health-related quality of life in oncology: a systematic review protocol. BMJ Open 2016;6(9):e012909.

[9] Korn EL, Freidlin B. Interim ffutility monitoring assessing immune therapies with a potentially delayed treatment effect. J Clin Oncol 2018:2444–9. JCO. 2018.2077. 7144.

[10] Lassere MN. The Biomarker-Surrogacy Evaluation Schema: a review of the biomarker-surrogate literature and a proposal for a criterion-based, quantitative, multidimensional hierarchical levels of evidence schema for evaluating the status of biomarkers as surrogate endpoints. Stat Methods Med Res 2008;17(3):303–40.

[11] Elston J, Taylor RS. Use of surrogate outcomes in cost-effectiveness models: a review of United Kingdom health technology assessment reports. Int J Technol Assess Health Care 2009;25(1):6–13.

[12] Hernandez-Villafuerte K, Fischer A, Latimer N. Challenges and methodologies in using progression free survival as a surrogate for overall survival in oncology. Int J Technol Assess Health Care 2018;24(3):300–16.

[13] Nasser M, Sawicki P, Institute for quality and efficiency in health care. Validity of surrogate endpoints in oncology: executive summary. 2011. https://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0065194/.

[14] Hey SP, Heilig CM, Weijer C. Accumulating Evidence and Research Organization (AERO) model: a new tool for representing, analyzing, and planning a translational research program. Trials 2013;14(1):159.

[15] Puthumana J, Wallach JD, Ross JS. Clinical trial evidence supporting FDA approval of drugs granted breakthrough therapy designation. J Am Med Assoc 2018;320(3):301–3.

[16] Zettler M, Nabhan C. Fulfillment of postmarketing requirements to the FDA for therapies granted oncology indications between 2011 and 2016. JAMA Oncol 2018;4(7):993–4.
[17] Beaver JA, Howie LJ, Pelosof L, 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 2018.
[18] US Food and Drug Administration. KEYTRUDA (pembrolizumab) label. In: Services US department of health and human services. MD: Silver Springs; 2018.
[19] Ioannidis JP. The importance of potential studies that have not existed and registration of observational data sets. J Am Med Assoc 2012;308(6):575–6.
[20] Hashim M, Pfeiffer BM, Bartsch R, Postma M, Heeg B. Do surrogate endpoints better correlate with overall survival in studies that did not allow for crossover or reported balanced postprogression treatments? An application in advanced non–small cell lung cancer. Value Health 2018;21(1):9–17.
[21] Ng R, Pond G, Tang P, MacIntosh P, Siu L, Chen E. Correlation of changes between 2-year disease-free survival and 5-year overall survival in adjuvant breast cancer trials from 1966 to 2006. Ann Oncol 2007;19(3):481–6.
[22] Buyse M, Burzykowski T, Michiels S, Carroll K. Individual- and trial-level surrogacy in colorectal cancer. Stat Methods Med Res 2008;17(5):467–75.
[23] Sargent DJ, Wieand HS, Haller DG, et al. Disease-free survival versus overall survival as a primary end point for adjuvant colon cancer studies: individual patient data from 20,898 patients on 18 randomized trials. J Clin Oncol 2005;23(34):8664–70.
[24] Sargent D, Shi Q, Yothers G, et al. Two or three year disease-free survival (DFS) as a primary end-point in stage III adjuvant colon cancer trials with fluoropyrimidines with or without oxaliplatin or irinotecan: data from 12,676 patients from MOSAIC, X-ACT, PETACC-3, C-06, C-07 and C89803. Eur J Canc 2011;47(7):990–6.
[25] Oba K, Paoletti X, Alberts S, et al. Disease-free survival as a surrogate for overall survival in adjuvant trials of gastric cancer: a meta-analysis. J Natl Cancer Inst 2013;105(21):1600–7.
[26] Harshman LC, Xie W, Moreira RB, et al. Evaluation of disease-free survival as an intermediate metric of overall survival in patients with localized renal cell carcinoma: a trial-level meta-analysis. Cancer 2018;124(5):925–33.
[27] Suciu S, Eggermont AM, Lorigan P, et al. Relapse-free survival as a surrogate for overall survival in the evaluation of stage II–III melanoma adjuvant therapy. JNCI (J Natl Cancer Inst) 2018;110(1).
[28] Michiels S, Le Maître A, Buyse M, et al. Surrogate endpoints for overall survival in locally advanced head and neck cancer: meta-analyses of individual patient data. Lancet Oncol 2009;10(4):341–50.
[29] Mauguen A, Pignon J-P, Burdett S, et al. Surrogate endpoints for overall survival in chemotherapy and radiotherapy trials in operable and locally advanced lung cancer: a re-analysis of meta-analyses of individual patients' data. Lancet Oncol 2013;14(7):619–26.
[30] Petrelli F, Borgonovo K, Cabiddu M, Ghilardi M, Lonati V, Barni S. Pathologic complete response and disease-free survival are not surrogate endpoints for 5-year survival in rectal cancer: an analysis of 22 randomized trials. J Gastrointest Oncol 2017;8(1):39.
[31] Moriwaki T, Yamamoto Y, Gosho M, et al. Correlations of survival with progression-free survival, response rate, and disease control rate in advanced biliary tract cancer: a meta-analysis of randomised trials of first-line chemotherapy. Br J Cancer 2016;114(8):881–8.
[32] Berruti A, Amoroso V, Gallo F, et al. Pathologic complete response as a potential surrogate for the clinical outcome in patients with breast cancer after neoadjuvant therapy: a meta-regression of 29 randomized prospective studies. J Clin Oncol 2014;32(34):3883–91.
[33] Hackshaw A, Knight A, Barrett-Lee P, Leonard R. Surrogate markers and survival in women receiving first-line combination anthracycline chemotherapy for advanced breast cancer. Br J Cancer 2005;93(11):1215–21.
[34] Burzykowski T, Buyse M, Piccart-Gebhart MJ, et al. Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. J Clin Oncol 2008;20(26):1987–92.
[35] Miksad RA, Zietemann V, Gothe R, et al. Progression-free survival as a surrogate endpoint in advanced breast cancer. Int J Technol Assess Health Care 2008;24(4):371–83.
[36] Sherrill B, Amonkar M, Wu Y, et al. Relationship between effects on time-to-disease progression and overall survival in studies of metastatic breast cancer. Br J Cancer 2008;99(10):1572–8.
[37] Wilkerson J, Fojo T. Progression-free survival is simply a measure of a drug's effect while administered and is not a surrogate for overall survival. Cancer J 2009;15(5):379–85.
[38] Beauchemin C, Cooper D, Lapierre M-È, Yelle L, Lachaine J. Progression-free survival as a potential surrogate for overall survival in metastatic breast cancer. Onco Targets Ther 2014;7:1101.
[39] Petrelli F, Barni S. Surrogate endpoints in metastatic breast cancer treated with targeted therapies: an analysis of the first-line phase III trials. Med Oncol 2014;31(1):776.
[40] Adunlin G, Cyrus JW, Dranitsaris G. Correlation between progression-free survival and overall survival in metastatic breast cancer patients receiving anthracyclines, taxanes, or targeted therapies: a trial-level meta-analysis. Breast Cancer Res Treat 2015;154(3):591–608.
[41] Michiels S, Pugliano L, Marguet S, et al. Progression-free survival as surrogate end point for overall survival in clinical trials of HER2-targeted agents in HER2-positive metastatic breast cancer. Ann Oncol 2016;27(6):1029–34.
[42] Liu L, Chen F, Zhao J, Yu H. Correlation between overall survival and other endpoints in metastatic breast cancer with second- or third-line chemotherapy: literature-based analysis of 24 randomized trials. Bulletin du cancer 2016;103(4):336–44.
[43] Buyse M, Thirion P, Carlson RW, Burzykowski T, Molenberghs G, Piedbois P. Relation between tumour response to first-line chemotherapy and survival in advanced colorectal cancer: a meta-analysis. Lancet 2000;356(9227):373–8.
[44] Johnson KR, Ringland C, Stokes BJ, et al. Response rate or time to progression as predictors of survival in trials of metastatic colorectal cancer or non-small-cell lung cancer: a meta-analysis. Lancet Oncol 2006;7(9):741–6.
[45] Buyse M, Burzykowski T, Carroll K, et al. Progression-free survival is a surrogate for survival in advanced colorectal cancer. J Clin Oncol 2007;25(33):5218–24.
[46] Tang PA, Bentzen SM, Chen EX, Siu LL. Surrogate end points for median overall survival in metastatic colorectal cancer: literature-based analysis from 39 randomized controlled trials of first-line chemotherapy. J Clin Oncol 2007;25(29):4562–8.
[47] Chirila C, Odom D, Devercelli G, et al. Meta-analysis of the association between progression-free survival and overall survival in metastatic colorectal cancer. Int J Colorectal Dis 2012;27(5):623–34.
[48] Giessen C, Laubender RP, Ankerst DP, et al. Progression-free survival as a surrogate endpoint for median overall survival in metastatic colorectal cancer: literature-based analysis from 50 randomized first-line trials. Clin Cancer Res 2013;19(1):225–35.
[49] Petrelli F, Barni S. Correlation of progression-free and post-progression survival with overall survival in advanced colorectal cancer. Ann Oncol 2013;24(1):186–92.
[50] Sidhu R, Rong A, Dahlberg S. Evaluation of progression-free survival as a surrogate endpoint for survival in chemotherapy and targeted agent metastatic colorectal cancer trials. Clin Cancer Res 2013;19(5):969–76.

[51] Giessen C, Laubender RP, Ankerst DP, et al. Surrogate endpoints in second-line treatment for mCRC: a systematic literature-based analysis from 23 randomised trials. Acta Oncologica 2015;54(2):187–93.
[52] Shi Q, De Gramont A, Grothey A, et al. Individual patient data analysis of progression-free survival versus overall survival as a first-line end point for metastatic colorectal cancer in modern randomized trials: findings from the analysis and research in cancers of the digestive system database. J Clin Oncol 2015;33(1):22.
[53] Ciani O, Buyse M, Garside R, et al. Meta-analyses of randomized controlled trials show suboptimal validity of surrogate outcomes for overall survival in advanced colorectal cancer. J Clini Epidemiol 2015;68(7):833–42.
[54] Colloca G, Venturino A, Guarneri D. Analysis of clinical end points of randomised trials including bevacizumab and chemotherapy versus chemotherapy as first-line treatment of metastatic colorectal cancer. Clin Oncol 2016;28(10):e155–64.
[55] Montagnani F, Di Leonardo G, Pino MS, et al. Progression-free survival as a surrogate end-point in advanced colorectal cancer treated with antiangiogenic therapies. Anticancer Res 2016;36(8):4259–65.
[56] Cremolini C, Antoniotti C, Pietrantonio F, et al. Surrogate endpoints in second-line trials of targeted agents in metastatic colorectal cancer: a literature-based systematic review and meta-analysis. Cancer Res Treat 2017;49(3):834.
[57] Zhu R, Lu D, Chu Y-W, et al. Assessment of correlation between early and late efficacy endpoints to identify potential surrogacy relationships in non-hodgkin lymphoma: a literature-based meta-analysis of 108 phase II and phase III studies. AAPS J 2017:1–13.
[58] Kawakami H, Okamoto I, Hayashi H, Taguri M, Morita S, Nakagawa K. Postprogression survival for first-line chemotherapy in patients with advanced gastric cancer. Eur J Cancer (Oxford, England 1990 2013;49(14):3003–9.
[59] Paoletti X, Oba K, Bang Y-J, et al. Progression-free survival as a surrogate for overall survival in advanced/recurrent gastric cancer trials: a meta-analysis. J Natl Cancer Inst 2013;105(21):1667–70.
[60] Shitara K, Matsuo K, Muro K, Doi T, Ohtsu A. Correlation between overall survival and other endpoints in clinical trials of second-line chemotherapy for patients with advanced gastric cancer. Gastric Cancer 2014;17(2):362–70.
[61] Lee D-W, Jang M-J, Lee K-H, et al. TTP as a surrogate endpoint in advanced hepatocellular carcinoma treated with molecular targeted therapy: meta-analysis of randomised controlled trials. Brit J Cancer 2016.
[62] Flaherty KT, Hennig M, Lee SJ, et al. Surrogate endpoints for overall survival in metastatic melanoma: a meta-analysis of randomised controlled trials. Lancet Oncol 2014;15(3):297–304.
[63] Cartier S, Zhang B, Rosen VM, et al. Relationship between treatment effects on progression-free survival and overall survival in multiple myeloma: a systematic review and meta-analysis of published clinical trial data. Oncol Res Treat 2015;38(3):88–94.
[64] Rotolo F, Pignon J-P, Bourhis J, et al. Surrogate end points for overall survival in loco-regionally advanced nasopharyngeal carcinoma: an individual patient data meta-analysis. J Natl Cancer Inst 2017;109(4).
[65] Lee L, Wang L, Crump M. Identification of potential surrogate end points in randomized clinical trials of aggressive and indolent non-Hodgkin's lymphoma: correlation of complete response, time-to-event and overall survival end points. Ann Oncol 2011;22(6):1392–403.
[66] Hotta K, Fujiwara Y, Matsuo K, et al. Time to progression as a surrogate marker for overall survival in patients with advanced non-small cell lung cancer. J Thorac Oncol 2009;4(3):311–7.
[67] Hotta K, Kiura K, Fujiwara Y, et al. Role of survival post-progression in phase III trials of systemic chemotherapy in advanced non-small-cell lung cancer: a systematic review. PLoS One 2011;6(11):e26646.
[68] Petrelli F, Barni S. Is overall survival still the primary endpoint in maintenance non-small cell lung cancer studies? An analysis of phase III randomised trials. Trans Lung Cancer Res 2013;2(1):6.
[69] Hotta K, Suzuki E, Di Maio M, et al. Progression-free survival and overall survival in phase III trials of molecular-targeted agents in advanced non-small-cell lung cancer. Lung cancer (Amsterdam, Netherlands) 2013;79(1):20–6.
[70] Aboshi M, Kaneko M, Narukawa M. Factors affecting the association between overall survival and progression-free survival in clinical trials of first-line treatment for patients with advanced non-small cell lung cancer. J Cancer Res Clin Oncol 2014;140(5):839–48.
[71] Hotta K, Kato Y, Leighl N, et al. Magnitude of the benefit of progression-free survival as a potential surrogate marker in phase 3 trials assessing targeted agents in molecularly selected patients with advanced non-small cell lung cancer: systematic review. PLoS One 2015;10(3), e0121211.
[72] Suzuki H, Hirashima T, Okamoto N, et al. Relationship between progression-free survival and overall survival in patients with advanced non-small cell lung cancer treated with anticancer agents after first-line treatment failure. Asia-Pacific J Clin Oncol. 2015;11(2):121–8.
[73] Nakashima K, Horita N, Nagai K, et al. Progression-free survival, response rate, and disease control rate as predictors of overall survival in phase III randomized controlled trials evaluating the first-line chemotherapy for advanced, locally advanced, and recurrent non–small cell lung carcinoma. J Thorac Oncol 2016;11(9):1574–85.
[74] Colloca G, Venturino A. Trial-level analysis of progression-free survival and response rate as end points of trials of first-line chemotherapy in advanced ovarian cancer. Med Oncol 2017;34(5):87.
[75] Siddiqui MK, Tyczynski J, Pahwa A, Fernandes AW. Objective response rate is a possible surrogate endpoint for survival in patients with advanced, recurrent ovarian cancer. Gynecol Oncol 2017.
[76] Petrelli F, Pietrantonio F, Cremolini C, et al. Early tumour shrinkage as a prognostic factor and surrogate end-point in colorectal cancer: a systematic review and pooled-analysis. Eur J Cancer 2015;51(7):800–7.
[77] Hamada T, Nakai Y, Isayama H, et al. Progression-free survival as a surrogate for overall survival in first-line chemotherapy for advanced pancreatic cancer. Eur J Cancer 2016;65:11–20.
[78] Colloca G, Venturino A, Guarneri D. Analysis of response-related and time-to-event endpoints in randomized trials of gemcitabine-based treatment versus gemcitabine alone as first-line treatment of patients with advanced pancreatic cancer. Clin Colorectal Cancer 2016;15(3):264–76.
[79] Makris EA, MacBarb R, Harvey DJ, Poultsides GA. Surrogate end points for overall survival in metastatic, locally advanced, or unresectable pancreatic cancer: a systematic review and meta-analysis of 24 randomized controlled trials. Ann Surg Oncol 2017:1–8.
[80] Xie W, Regan M, Buyse M, et al. Metastasis-free survival is a strong surrogate of overall survival in localized prostate cancer. J Clin Oncol 2017;35(27):3097–114.
[81] Delea T, Khuu A, Heng DY, Haas T, Soulieres D. Association between treatment effects on disease progression end points and overall survival in clinical studies of patients with metastatic renal cell carcinoma. Br J Cancer 2012;107(7):1059–68.
[82] Petrelli F, Barni S. Surrogate end points and postprogression survival in renal cell carcinoma: an analysis of first-line trials with targeted therapies. Clin Genitourin Cancer 2013;11(4):385–9.
[83] Bria E, Massari F, Maines F, et al. Progression-free survival as primary endpoint in randomized clinical trials of targeted agents for advanced renal cell carcinoma. Correlation with overall survival, benchmarking and power analysis. Crit Rev Oncol Hematol 2015;93(1):50–9.

[84] Johnson K, Liauw W, Lassere M. Evaluating surrogacy metrics and investigating approval decisions of progression-free survival (PFS) in metastatic renal cell cancer: a systematic review. Ann Oncol 2015;26(3):485–96.
[85] Hotta K, Kiura K, Fujiwara Y, et al. Association between incremental gains in the objective response rate and survival improvement in phase III trials of first-line chemotherapy for extensive disease small-cell lung cancer. Ann Oncol 2009;20(5):829–34.
[86] Foster NR, Renfro LA, Schild SE, et al. Multitrial evaluation of progression-free survival as a surrogate end point for overall survival in first-line extensive-stage small-cell lung cancer. J Thorac Oncol 2015;10(7):1099–106.
[87] Roviello G, Andre F, Venturini S, et al. Response rate as a potential surrogate for survival and efficacy in patients treated with novel immune checkpoint inhibitors: a meta-regression of randomised prospective studies. Eur J Cancer 2017;86:257–65.
[88] Amir E, Seruga B, Kwong R, Tannock IF, Ocaña A. Poor correlation between progression-free and overall survival in modern clinical trials: are composite endpoints the answer? Eur J Cancer 2012;48(3):385–8.
[89] Cortazar P, Zhang L, Untch M, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. Lancet 2014;384(9938):164–72.
[90] Kataoka K, Nakamura K, Mizusawa J, et al. Surrogacy of progression-free survival (PFS) for overall survival (OS) in esophageal cancer trials with preoperative therapy: literature-based meta-analysis. Eur J Surg Oncol (EJSO). 2017;43(10):1956–61.
[91] Petrelli F, Tomasello G, Barni S. Surrogate end-points for overall survival in 22 neoadjuvant trials of gastro-oesophageal cancers. Eur J Cancer 2017;76:8–16.
[92] Mushti SL, Mulkey F, Sridhara R. Evaluation of overall response rate and progression-free survival as potential surrogate endpoints for overall survival in immunotherapy trials. Clin Cancer Res 2018:2268–75. clincanres. 1902.2017.
[93] Abdel-Rahman O. Surrogate end points for overall survival in trials of PD-(L) 1 inhibitors for urinary cancers: a systematic review. Immunotherapy 2018;10(2):139–48.
[94] Chen YP, Sun Y, Chen L, et al. Surrogate endpoints for overall survival in combined chemotherapy and radiotherapy trials in nasopharyngeal carcinoma: meta-analysis of randomised controlled trials. Radiother Oncol 2015;116(2):157–66.