REVIEW

Immune-related adverse events as potential surrogates of immune checkpoint inhibitors' efficacy: a systematic review and meta-analysis of randomized studies

V. Amoroso[1*†], F. Gallo[2†], A. Alberti[1], D. Paloschi[1], W. Ferrari Bravo[1], A. Esposito[1], D. Cosentini[1], S. Grisanti[1], R. Pedersini[1], F. Petrelli[3] & A. Berruti[1]

[1]Department of Medical and Surgical Specialties, Radiological Sciences and Public Health, Medical Oncology Unit, University of Brescia, Azienda Socio Sanitaria Territoriale (ASST) Spedali Civili, Brescia; [2]Clinical Epidemiology Unit, IRCCS, Ospedale Policlinico San Martino, Genova; [3]Medical Sciences Department, Oncology Unit, Azienda Socio Sanitaria Territoriale (ASST) Bergamo Ovest, Treviglio, Bergamo, Italy

Available online xxx

Background: Immune-related adverse events (irAEs) are frequently reported during immune checkpoint inhibitor (ICI) therapy and are associated with long-term outcomes. It is unknown if the irAE occurrence is a valid surrogate of ICIs' efficacy.
Methods: We identified articles reporting the results of randomized trials of experimental ICI therapy in solid tumors with a systematic search. The control arms could be placebo, cytotoxic/targeted therapy, or ICI therapy. We extracted the hazard ratios for overall survival (OS) with the number of OS events per arm and the number and percentages of overall and specific irAEs of grade 1-2 and grade 3-4 per arm. We estimated the treatment effect on the potential surrogate outcome with the odds ratio of the irAE rate between the experimental and the control arm. The statistical analysis consisted of weighted linear regression on a logarithmic scale between treatment effects on irAE rate and treatment effects on OS.
Results: Sixty-two randomized trials were included for a total of 79 contrasts and 42 247 patients. The analyses found no significant association between the treatment effects for overall grade 1-2 or grade 3-4 irAE rates or specific (skin, gastrointestinal, endocrine) irAE rates. In the non-small-cell lung cancer (NSCLC) trial subset, we observed a negative association between treatment effects on overall grade 1-2 irAEs and treatment effects on OS in studies with patients selected for programmed death-ligand 1 expression ($R^2 = 0.55$; 95% confidence interval 0.20-0.95; $R = -0.69$). In the melanoma trial subset, a negative association was shown between treatment effects on gastrointestinal grade 3-4 irAEs and treatment effects on OS in trials without an ICI-based control arm ($R^2 = 0.77$; 95% confidence interval 0.24-0.99; $R = -0.89$).
Conclusions: We found low-strength correlations between the ICI therapy effects on overall or specific irAE rates and the treatment effects on OS in several cancer types.
Key words: immune checkpoint inhibitors, immune-related adverse events, surrogate endpoints, trial-based meta-analysis

INTRODUCTION

Immune checkpoint inhibitors (ICIs), which include anti-cytotoxic T-lymphocyte antigen 4 (CTLA-4), anti-programmed cell death protein 1 (PD-1), and anti-programmed death-ligand 1 (PD-L1) monoclonal antibodies, represent significant progress in oncology, particularly in the treatment of solid tumors. These agents block the ability of malignancies to escape immune surveillance, thus reactivating the T-cell immune response against cancer. The activation of the immune system can also induce an immune-mediated inflammation of different organs and tissues, however, especially the skin, endocrine system, gastrointestinal (GI) tract, and lungs.[1-4]

The immune-related adverse events (irAEs), secondary to ICI therapy, significantly impact treatment delivery. Most irAEs tend to be mild and self-limiting, but potentially life-threatening events can occur in a few severe cases (with toxicities of grade 3 or 4).[3] Hence, either the European Society of Medical Oncology (ESMO) or the American

*Correspondence to: Dr Vito Amoroso, Oncologia Medica, ASST Spedali Civili di Brescia, p.le Spedali Civili 1, 20123, Brescia, Italy. Tel: +39-030-3995260
E-mail: vitoamoroso@alice.it (V. Amoroso).

†V.A. and F.G. contributed equally to this work and are co-primary authors.

Society of Clinical Oncology (ASCO) has issued guidelines for managing the various irAEs.[5,6]

Several studies evaluating different ICIs have shown an association between the development of irAEs and long-term survival, particularly in patients with melanoma,[7] non-small-cell lung cancer (NSCLC),[8,9] and urothelial cancers.[10] A recent exploratory pooled analysis of 2503 patients from three randomized trials in non-squamous NSCLC (IMpower130, IMpower132, and IMpower150)[11] showed that patients with irAEs (excluding rash, perceived as the least specific irAE) had longer overall survival (OS) compared with those without irAEs in the atezolizumab-containing arm [hazard ratio (HR) = 0.75; 95% confidence interval (CI) 0.65-0.87] but not in the control arm (HR = 0.90; 95% CI 0.71-1.12). NSCLC patients with irAEs of higher grade (grade 3 or 4) had inferior outcomes compared with those with lower grade (grade 1 or 2) irAEs.[11]

The previously mentioned individual-level relationship does not imply that the irAE occurrence might be a valid surrogate of immunotherapy efficacy. A surrogate endpoint to be validated has to correlate with the OS, and the effect of the intervention on the surrogate should predict the effect on the long-term clinical outcome.[12,13] A meta-analysis of randomized, controlled trials represents an essential tool for validating surrogate endpoints.[12,14,15]

Currently, no candidate surrogate for long-term survival has been validated in ICI trials. Traditional intermediate endpoints, such as progression-free survival and objective response rate, failed to demonstrate either individual-level or trial-level association with OS in studies evaluating ICI therapy.[16-19]

This study aimed to assess whether the treatment effects on irAEs can predict the treatment effects on the OS at the trial level, supporting the hypothesis of irAE occurrence as a surrogate of ICI treatment efficacy. The validation of this surrogacy would establish that the irAEs lie in the causal pathway of the ICI therapy effect on OS and thus capture most of the intervention's impact on the clinical outcome. The secondary aims were to explore the heterogeneity of the trial-level associations according to specific trial characteristics: the study population (selected for tumor biomarkers, such as PD-L1 expression), the treatment type administered in the experimental arm [anti-CTLA-4 versus anti-programmed cell death protein 1 (anti-PD-1)/PD-L1 versus a combination of anti-CTLA4 and anti-PD-1/PD-L1 antibodies], the type of control arm (an ICI therapy versus a cytotoxic/targeted therapy/placebo control arm), and the treatment line (first line versus second or later line). For this purpose, we carried out a systematic literature search and a trial-based meta-analysis of randomized, controlled trials of experimental ICI therapy in patients with advanced solid tumors.

METHODS

We planned a meta-analysis with aggregate data (AD) derived from randomized trials assessing the efficacy of ICI therapy and retrieved by a systematic search, according to the Reporting of Surrogate Endpoint Evaluation using Meta-analyses (ReSEEM) guidelines and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[20,21] The primary objective was the trial-level validation of the overall irAE rate as a potential surrogate endpoint of the effect on OS of the ICI therapy in advanced-stage solid tumors (i.e. evaluating whether the size of the treatment effect on irAEs allows one to predict the size of the effect on the final outcome). The primary analysis was planned in specific tumor types for which a significant number of randomized trials are available [i.e. NSCLC, melanoma, renal cell carcinoma, small-cell lung cancer (SCLC)].

Systematic search

We identified original articles reporting the results of randomized trials of ICI therapy in solid tumors from PubMed, EMBASE, and Cochrane Library database search from inception to 31 March 2021. Keywords included the following terms: 'immunotherapy', 'immune checkpoint inhibitor therapy', 'anti-PD-1', 'anti-PD-L1', 'anti-CTLA-4', 'atezolizumab', 'avelumab', 'durvalumab', 'ipilimumab', 'nivolumab', 'pembrolizumab', 'tremelimumab', 'randomized OR randomised', 'phase 2 OR phase 3 trial'. Proceedings of the ASCO Annual Meeting, American Association for Cancer Research (AACR) Annual Meeting, and the ESMO Meeting were hand-searched for the years 2016 through 2020. We examined references of clinical trials, review articles, and editorials for other relevant articles. Commentaries, reviews, preclinical studies, observational studies, and articles not written in English were not eligible for the present study. We reported the detailed electronic search strategy and the PRISMA flow diagram in the Supplementary Material, available at https://doi.org/10.1016/j.esmoop.2023.100787.

Eligibility criteria

All citations were independently screened by three authors (DP, WFB, and VA). Eligible studies should have an ICI-based regimen as experimental arm (either ICI monotherapy or a combination of two different ICIs or an ICI plus a cytotoxic or targeted therapy). The control arm could be a cytotoxic/targeted therapy, placebo, or ICI-based therapy. We included randomized phase II or III trials that reported OS data and irAE rates according to the Common Terminology Criteria for Adverse Events (CTCAE). We excluded trials evaluating different forms of immunotherapy (i.e. anti-tumor vaccines). We included the most recent report of eligible randomized trials in the case of duplicate publications.

Data extraction

We recorded the following information from each selected trial: authors' names, journal and year of publication, study design, and duration of follow-up. We initially defined the experimental and control arms for each included trial. We extracted the following data from each treatment arm: the

number of patients randomly assigned to treatment and analyzed per arm, the HRs for OS with the number of OS events per arm, and the overall number and percentages of the irAEs per arm of any grade and grade 1-2 or 3-4. Furthermore, we extracted irAE rates for each specific organ (i.e. GI tract, skin, and endocrine system). In trials not reporting overall irAE rates, we summed up GI (diarrhea, colitis), skin (rash, pruritus, dermatitis, severe skin reaction), endocrine (hypothyroidism, hyperthyroidism, hypophysitis, adrenal insufficiency), and pulmonary (pneumonitis) toxicities per each arm.

Endpoints

We considered OS as the long-term clinical outcome of interest. The HR of death between the experimental and the control arm was used as the treatment effect estimate on the clinical outcome. We estimated the treatment effect on the potential surrogate outcome with the odds ratio (OR) of irAEs between the experimental and the control arm. The irAEs were collected as reported by the investigators according to CTCAE in each included study.

In addition, we carried out a sensitivity analysis using the attributable risk instead of the OR to estimate the treatment effects on the candidate surrogate endpoint, the irAE rate. Finally, a validation analysis was carried out, excluding trials that needed missing data imputation.

Statistical analysis

The statistical analysis consisted of weighted linear regression on a logarithmic scale between treatment effects on the OS and treatment effects on the irAE rates. Treatment effects on the clinical outcome (expressed as HRs) and on the surrogate outcome (expressed as ORs) were extracted from each trial or estimated when not reported using the survival curves (for HR) and the proportion of irAEs (for OR). Each study was weighted according to the total number of events (deaths) in patients included in each trial. When events were not explicitly reported, or the irAE event number was beyond the total number of patients, they were estimated using available information, as detailed in the Supplementary Material, available at https://doi.org/10.1016/j.esmoop.2023.100787. For trials with multiple arms, a contrast with the control arm was studied for each experimental arm (that is, the treatment effect was estimated for each experimental arm versus the control arm). As multiple contrasts from the same trial create 'clustered data,' each contrast coming from a trial with more than two arms was given a lower weight, obtained by equally dividing the number of events in the control arm among the experimental arms.

The coefficient of determination (R^2) was used to assess the goodness of fit for each model and quantify the surrogacy level. The R^2 threshold to identify a valid surrogate was set at 0.65 or higher according to Ciani et al.[22] We estimated 95% CIs of the R^2 values by bootstrap techniques.

We scored the strength of trial-level correlation according to the criteria proposed by the Institute of Quality and Efficiency in Health Care,[23] and adapted by Prasad et al.[24]: low-strength correlation ($R \leq 0.70$), medium-strength correlation ($R > 0.7$ to $R < 0.85$), and high-strength correlation ($R \geq 0.85$).

Finally, an interaction analysis was used to test whether the surrogacy (that is, the slope of the weighted regression line) was different according to the following factors: type of experimental agent, selection of study population for PD-L1 expression or not, type of comparison, and treatment line. A further subgroup analysis was carried out only for significant results, dividing studies according to the factors mentioned above. This latter analysis was not pre-planned. P values <0.05 were selected as significant, and data were acquired and analyzed in R v4.1.1 software environment.[25]

RESULTS

Trials included in the analysis

We identified 3150 citations with a systematic literature search. From a total of 98 randomized trials identified after level 1 screening, we included 62 studies (79 contrasts and 42 247 patients) (Supplementary Figure S1, available at https://doi.org/10.1016/j.esmoop.2023.100787 and Table 1). We selected 6 randomized trials (8 contrasts) assessing ICI therapy in patients with upper GI cancer (esophageal, gastric, or gastroesophageal junction cancer),[26-31] 6 trials (8 contrasts) in patients with renal cell carcinoma,[32-37] 12 trials (17 contrasts) in patients with melanoma,[38-49] 16 trials (18 contrasts) in NSCLC,[50-64] 5 trials (7 contrasts) in SCLC,[65-69] and 6 trials (9 contrasts) in urothelial cancers.[70-75] We included 11 trials (12 contrasts) in patients with other primary tumor sites [breast cancer (2 trials);[76,77] glioblastoma (1 trial);[78] mesothelioma (2 trials);[79,80] prostate cancer (2 trials);[81,82] hepatocellular carcinoma (1 trial);[83] head and neck cancer (3 studies and 4 contrasts)].[84-86]

The median follow-up of the included studies ranged from 6.5 to 51 months, with a median of 19.0 months [interquartile range (IQR), 11.8-27.5 months]. In 20 out of 79 contrasts (25.3%), the study population was selected for expressing PD-L1 in at least 1% of tumor or immune cells. We provided treatments' details in Supplementary Table S1, available at https://doi.org/10.1016/j.esmoop.2023.100787.

Relationship between treatment effects on irAEs and treatment effects on the clinical outcome (OS)

We estimated the relationship between treatment effects on irAEs versus treatment effects on OS and estimated regression equations based on data from all included studies (Table 1).

NSCLC

The analysis, regressing the log(HR) for OS on the log(OR) for irAEs, using the complete set of data, demonstrated no association between the two effects for either the overall (any grade, grade 1-2, grade 3-4) irAE rates or the specific organ (skin, GI, endocrine) irAE rates (Figure 1A-C;

Table 1. Trials included in the analysis

| Contrast	Tumor type	Experimental arm	Control arm	No. of patients	Median follow-up time (months)	OR (95% CI)	HR (95% CI)
Kojima et al.[26]	Upper GI	Pembrolizumab	Paclitaxel, docetaxel, or irinotecan	628	7.0	3.77 (2.27-6.26)	0.89 (0.75-1.05)
Chau et al.[27]	Upper GI	Nivolumab and chemotherapy	Chemotherapy	645	n.a.	4.97 (3.47-7.11)	0.74 (0.58-0.96)
Chau et al.[27]	Upper GI	Nivolumab and ipilimumab	Chemotherapy	649	n.a.	17.4 (11.7-25.87)	0.78 (0.62-0.98)
Chen et al.[28]	Upper GI	Nivolumab	Placebo	493	27.3	8.69 (1.15-65.89)	0.62 (0.51-0.76)
Shitara et al.[29]	Upper GI	Pembrolizumab	Paclitaxel	395	8.5	2.73 (1.6-4.66)	0.82 (0.66-1.03)
Shitara et al.[30]	Upper GI	Pembrolizumab	Cisplatin and fluorouracil	506	29.4	3.2 (1.83-5.58)	0.91 (0.74-1.1)
Shitara et al.[30]	Upper GI	Pembrolizumab, cisplatin, and fluorouracil	Cisplatin and fluorouracil	507	29.4	3.74 (2.16-6.49)	0.85 (0.7-1.03)
Janjigian et al.[31]	Upper GI	Nivolumab, oxaliplatin, and capecitabine	Oxaliplatin and capecitabine	1581	12.1	5.5 (4.39-6.9)	0.8 (0.68-0.94)
Ferris et al.[84]	Head and neck	Nivolumab	MTX, docetaxel, or cetuximab	361	>24.2	1.52 (0.93-2.48)	0.68 (0.54-0.86)
Cohen et al.[85]	Head and neck	Pembrolizumab	MTX, docetaxel, or cetuximab	495	7.3	2.53 (1.56-4.12)	0.8 (0.65-0.98)
Siu et al.[86]	Head and neck	Durvalumab and tremelimumab	Durvalumab	194	6.5	2.92 (1.06-7.99)	0.99 (0.69-1.43)
Siu et al.[86]	Head and neck	Durvalumab and tremelimumab	Tremelimumab	192	7.6	32.29 (1.94-538.89)	0.72 (0.51-1.03)
Motzer et al.[32]	RCC	Nivolumab 2 mg/kg	Nivolumab 0.3 mg/kg	114	24.0	1.45 (0.69-3.07)	0.8 (0.6-1.1)
Motzer et al.[32]	RCC	Nivolumab 10 mg/kg	Nivolumab 0.3 mg/kg	114	24.0	2.88 (1.34-6.2)	0.9 (0.6-1.2)
Motzer et al.[32]	RCC	Nivolumab 10 mg/kg	Nivolumab 2 mg/kg	108	24.0	1.98 (0.92-4.3)	n.a.
Motzer et al.[33]	RCC	Nivolumab and ipilimumab	Sunitinib	1096	32.4	0.88 (0.65-1.21)	0.71 (0.59-0.86)
Rini et al.[34]	RCC	Atezolizumab and bevacizumab	Sunitinib	915	24.0	1.44 (1.11-1.87)	0.93 (0.76-1.14)
Powles et al.[35]	RCC	Pembrolizumab and axitinib	Sunitinib	861	30.6	1.85 (1.41-2.44)	0.53 (0.38-0.74)
Motzer et al.[36]	RCC	Lenvatinib and pembrolizumab	Sunitinib	712	26.6	3.47 (2.53-4.75)	0.66 (0.49-0.88)
Choueiri et al.[37]	RCC	Nivolumab and cabozantinib	Sunitinib	651	18.1	9.11 (6.02-13.79)	0.6 (0.44-0.82)
Hodi et al.[38]	Melanoma	Ipilimumab and gp100	gp100	539	19.1	2.98 (1.96-4.53)	0.68 (0.55-0.85)
Hodi et al.[38]	Melanoma	Ipilimumab	gp100	273	22.5	3.36 (2.02-5.58)	0.66 (0.51-0.87)
Maio et al.[39]	Melanoma	Ipilimumab and dacarbazine	Dacarbazine	502	9.9	5.64 (3.8-8.35)	0.69 (0.57-0.84)
Wolchok et al.[40]	Melanoma	Ipilimumab 10 mg/kg	Ipilimumab 0.3 mg/kg	145	9.5	6.64 (3.2-13.8)	0.77 (0.52-1.13)
Wolchok et al.[40]	Melanoma	Ipilimumab 10 mg/kg	Ipilimumab 3 mg/kg	144	9.7	1.29 (0.64-2.62)	0.88 (0.59-1.29)
Wolchok et al.[40]	Melanoma	Ipilimumab 3 mg/kg	Ipilimumab 0.3 mg/kg	145	8.5	5.13 (2.51-10.5)	0.88 (0.6-1.28)
Long et al.[41]	Melanoma	Pembrolizumab	Ipilimumab	834	45.9	1.58 (1.09-2.27)	0.73 (0.61-0.89)
Hamid et al.[42]	Melanoma	Pembrolizumab 2 mg/kg	Investigator-choice chemotherapy	359	28.0	12.27 (3.68-40.92)	0.86 (0.67-1.1)
Hamid et al.[42]	Melanoma	Pembrolizumab 10 mg/kg	Investigator-choice chemotherapy	360	28.0	15.09 (4.56-49.93)	0.74 (0.57-0.96)
Hodi et al.[43]	Melanoma	Nivolumab and ipilimumab	Ipilimumab	142	24.5	1.84 (0.7-4.8)	0.74 (0.43-1.26)
Wolchok et al.[44]	Melanoma	Nivolumab	Ipilimumab	631	>36.0	0.58 (0.42-0.82)	0.63 (0.48-0.81)
Wolchok et al.[44]	Melanoma	Nivolumab and ipilimumab	Ipilimumab	629	>36.0	2.59 (1.7-3.95)	0.55 (0.42-0.72)
Ribas et al.[45]	Melanoma	Tremelimumab	Temozolomide or dacarbazine	655	n.a.	n.a.	0.88 (NA-NA)
Weber et al.[46]	Melanoma	Nivolumab → ipilimumab	Ipilimumab → nivolumab	138	17.2	n.a.	0.48 (0.29-0.8)
Ascierto et al.[47]	Melanoma	Ipilimumab 10 mg/kg	Ipilimumab 3 mg/kg	727	12.8	2.4 (1.76-3.28)	0.84 (0.7-0.99)
Larkin et al.[48]	Melanoma	Nivolumab	Investigator's choice chemotherapy	405	24.0	5.92 (3.58-9.79)	0.95 (0.73-1.24)
Ascierto et al.[49]	Melanoma	Nivolumab	Ipilimumab	906	51.0	n.a.	0.87 (0.66-1.14)
Horn et al.[50]	NSCLC	Nivolumab	Docetaxel	272	n.a.	0.87 (0.51-1.5)	0.62 (0.47-0.8)
Brahmer et al.[53]	NSCLC	Pembrolizumab	Platinum-doublet chemotherapy	305	19.1	9.05 (4.12-19.87)	0.63 (0.46-0.88)
Carbone et al.[51]	NSCLC	Nivolumab	Platinum-based chemotherapy	541	13.5	2.6 (1.78-3.79)	1.07 (0.86-1.33)
Borghaei et al.[52]	NSCLC	Pembrolizumab, carboplatin, and pemetrexed	Carboplatin and pemetrexed	123	23.9	2.22 (0.82-6.03)	0.56 (0.32-0.95)
Horn et al.[50]	NSCLC	Nivolumab	Docetaxel	582	n.a.	1.28 (0.91-1.79)	0.75 (0.63-0.91)
Herbst et al.[54]	NSCLC	Pembrolizumab 2 mg/kg	Docetaxel	688	13.1	5.76 (3.04-10.9)	0.71 (0.58-0.88)
Herbst et al.[54]	NSCLC	Pembrolizumab 10 mg/kg	Docetaxel	689	13.1	6.58 (3.49-12.38)	0.61 (0.49-0.75)
Gandhi et al.[55]	NSCLC	Pembrolizumab, pemetrexed, and platinum-based drug	Pemetrexed and platinum-based drug	616	10.5	2.09 (1.28-3.4)	0.49 (0.38-0.64)
Socinski et al.[56]	NSCLC	Atezolizumab, bevacizumab, carboplatin, and paclitaxel	Bevacizumab, carboplatin, and paclitaxel	692	15.4	4.19 (3.06-5.75)	0.78 (0.64-0.96)
Jotte et al.[57]	NSCLC	Atezolizumab, carboplatin, and nab-paclitaxel	Carboplatin and nab-paclitaxel	683	25.8	4.42 (3-6.49)	0.88 (0.73-1.05)

Continued

Table 1. Continued

Contrast	Tumor type	Experimental arm	Control arm	No. of patients	Median follow-up time (months)	OR (95% CI)	HR (95% CI)
Paz-Ares et al.[58]	NSCLC	Pembrolizumab, carboplatin, and paclitaxel or nab-paclitaxel	Carboplatin and paclitaxel or nab-paclitaxel	559	7.8	4.44 (2.69-7.31)	0.64 (0.49-0.85)
Herbst et al.[59]	NSCLC	Atezolizumab	Platinum-based chemotherapy	572	17.5	3.35 (2.24-5)	0.83 (0.65-1.07)
Rizvi et al.[60]	NSCLC	Durvalumab	Platinum-based chemotherapy	325	30.2	4.44 (2.32-8.49)	0.76 (0.56-1.02)
Rizvi et al.[60]	NSCLC	Durvalumab and tremelimumab	Platinum-based chemotherapy	325	30.2	11.18 (6.03-20.76)	0.85 (0.61-1.17)
West et al.[61]	NSCLC	Atezolizumab, carboplatin, and nab-paclitaxel	Carboplatin and nab-paclitaxel	679	18.8	3.21 (2.13-4.83)	0.79 (0.64-0.98)
Paz-Ares et al.[62]	NSCLC	Nivolumab, ipilimumab, and platinum-based chemotherapy	Platinum-based chemotherapy	719	13.2	18.85 (11.46-31)	0.66 (0.55-0.8)
Mok et al.[63]	NSCLC	Pembrolizumab	Platinum-based chemotherapy	1274	16.7	5 (3.52-7.12)	0.81 (0.71-0.93)
Govindan et al.[64]	NSCLC	Ipilimumab, carboplatin, and paclitaxel	Carboplatin and paclitaxel	749	12.1	2.01 (1.49-2.7)	0.91 (0.77-1.07)
Liu et al.[65]	SCLC	Atezolizumab, carboplatin, and etoposide	Carboplatin and etoposide	403	22.9	2.18 (1.42-3.35)	0.76 (0.6-0.95)
Reck et al.[66]	SCLC	Ipilimumab (concurrent), carboplatin, and paclitaxel	Carboplatin and paclitaxel	88	>11.1	n.a.	0.95 (0.59-1.54)
Reck et al.[66]	SCLC	Ipilimumab (phased), carboplatin, and paclitaxel	Carboplatin and paclitaxel	87	>11.1	0.37 (0.15-0.89)	0.75 (0.46-1.23)
Rudin et al.[67]	SCLC	Pembrolizumab, etoposide, and platinum	Etoposide and platinum	453	21.6	2.85 (1.68-4.83)	0.8 (0.64-0.98)
Goldman et al.[68]	SCLC	Durvalumab, tremelimumab, etoposide, and platinum	Etoposide and platinum	537	25.1	20.89 (9.47-46.1)	0.82 (0.68-1)
Goldman et al.[68]	SCLC	Durvalumab, etoposide, and platinum	Etoposide and platinum	537	25.1	9.25 (4.12-20.77)	0.75 (0.62-0.91)
Reck et al.[69]	SCLC	Ipilimumab, etoposide, and platinum	Etoposide and platinum	954	10.3	3.51 (2.68-4.6)	0.94 (0.81-1.09)
Fradet et al.[70]	Urothelial	Pembrolizumab	Taxane or vinflunine	542	27.7	3.4 (1.91-6.06)	0.7 (0.57-0.85)
Powles et al.[71]	Urothelial	Atezolizumab	Taxane or vinflunine	234	17.3	2.16 (1.21-3.84)	0.87 (0.63-1.21)
Powles et al.[72]	Urothelial	Durvalumab	Gemcitabine and platinum	690	41.2	11.43 (4.87-26.82)	0.99 (0.83-1.17)
Powles et al.[72]	Urothelial	Durvalumab and tremelimumab	Gemcitabine and platinum	686	41.2	30.13 (13.04-69.6)	0.85 (0.72-1.02)
Galsky et al.[73]	Urothelial	Atezolizumab and platinum-based chemotherapy	Platinum-based chemotherapy	851	11.8	1.9 (1.44-2.51)	0.83 (0.69-1)
Galsky et al.[73]	Urothelial	Atezolizumab	Platinum-based chemotherapy	762	11.8	1.12 (0.83-1.52)	1.02 (0.83-1.24)
Powles et al.[74]	Urothelial	Maintenance avelumab	Best supportive care	700	>19.0	28.26 (11.34-70.42)	0.69 (0.56-0.86)
Powles et al.[75]	Urothelial	Pembrolizumab, gemcitabine and platinum	Gemcitabine and platinum	703	31.7	11.82 (5.61-24.91)	0.86 (0.72-1.02)
Powles et al.[75]	Urothelial	Pembrolizumab	Gemcitabine and platinum	659	31.7	11.01 (5.18-23.39)	0.92 (0.77-1.11)
Schmid et al.[76]	Breast	Atezolizumab and nab-paclitaxel	Nab-paclitaxel	902	18.0	1.9 (1.46-2.48)	0.86 (0.72-1.02)
Winer et al.[77]	Breast	Pembrolizumab	Single-drug chemotherapy	622	31.4	6.37 (2.95-13.73)	0.97 (0.82-1.15)
Reardon et al.[78]	Glioblastoma	Nivolumab	Bevacizumab	369	9.5	3.13 (1.89-5.18)	1.04 (0.83-1.3)
Finn et al.[83]	HCC	Pembrolizumab	Placebo	413	12.2	2.5 (1.26-4.97)	0.78 (0.61-1)
Scherpereel et al.[79]	Mesothelioma	Nivolumab and ipilimumab	Nivolumab	125	10.4	0.46 (0.31-0.69)	0.92 (0.76-1.12)
Maio et al.[80]	Mesothelioma	Tremelimumab	Placebo	571	n.a.	n.a.	n.a.
Kwon et al.[81]	Prostate	Ipilimumab	Placebo	799	9.6	6.23 (4.55-8.54)	0.85 (0.72-1)
Beer et al.[82]	Prostate	Ipilimumab	Placebo	602	n.a.	8.55 (5.81-12.59)	1.11 (0.88-1.39)

CI, confidence interval; GI, gastrointestinal; HCC, hepatocellular carcinoma; HR, hazard ratio for overall survival (experimental versus control arm); MTX, methotrexate; n.a./NA, not available; NSCLC, non-small-cell lung cancer; OR, odds ratio for the overall any grade immune-related adverse event (experimental versus control arm); RCC, renal cell carcinoma; SCLC, small-cell lung cancer.

Supplementary Figure S2, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel A to I). Sensitivity and validation analysis gave similar results. The regression coefficients obtained from the main, sensitivity, and validation analyses were similar according to toxicity grade (Supplementary Figure S3, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel A).

The trials included in the analyses were classified as reported in Supplementary Table S1, available at https://doi.org/10.1016/j.esmoop.2023.100787. Regarding interaction analysis, the surrogacy (that is, the slope of the weighted regression line) was significantly different according to the study population selected for PD-L1 expression or not, treatment line, and type of experimental regimen factors (*P*
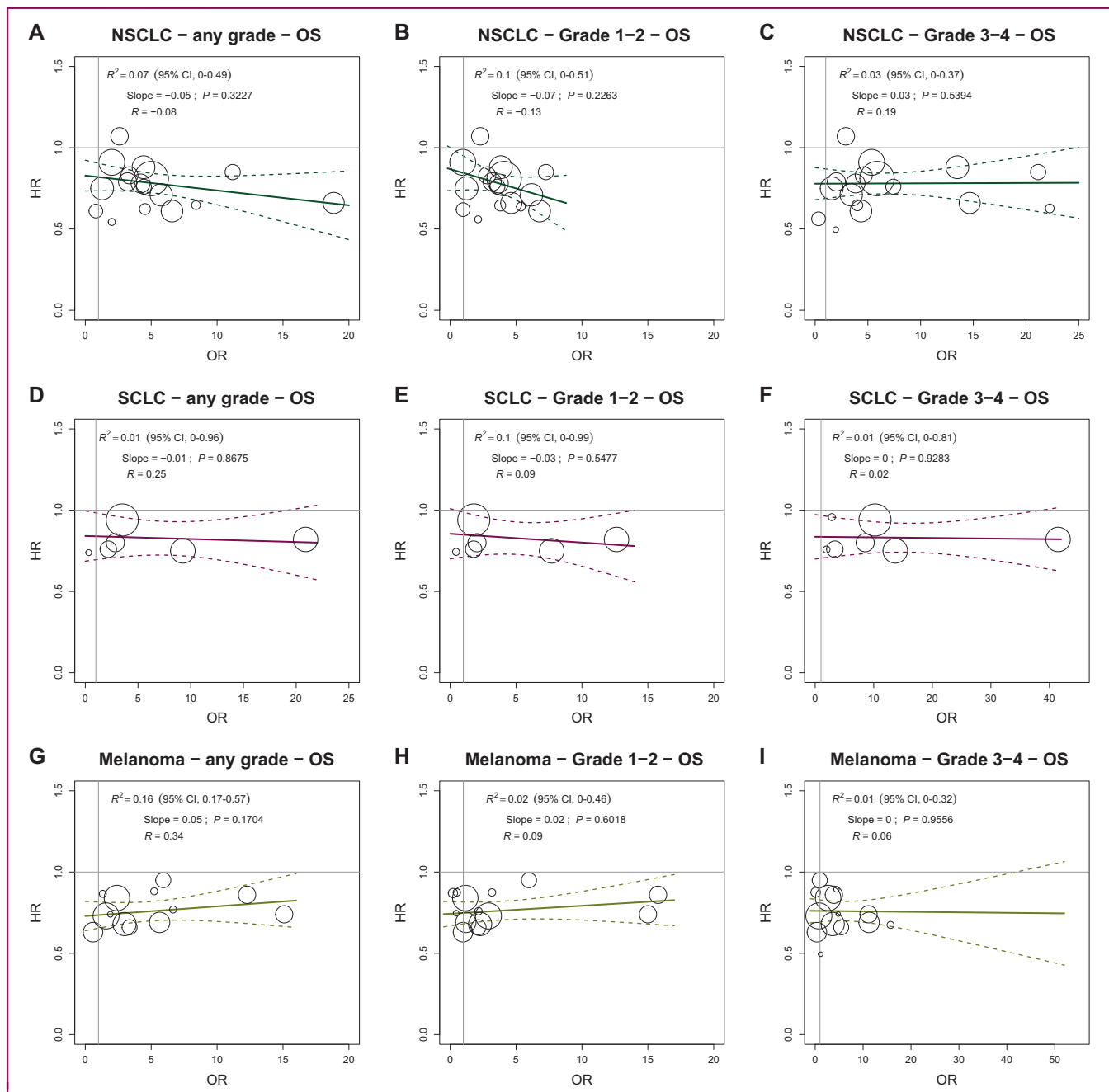
Figure 1. Treatment effects on immune-related adverse events versus treatment effects on overall survival in specific tumor sites according to toxicity grade.
Treatment effects are expressed as ORs for irAEs and HRs for OS. Every circle represents a comparison of the experimental group versus the control group, with the size of the circles representing the weight of the comparison, proportional to the number of events (deaths). The straight lines represent the weighted regressions, which show the effects on OS predicted by the observed effects on irAEs.
CI, confidence interval; GI, gastrointestinal; HCC, hepatocellular carcinoma; HR, hazard ratio; irAEs, immune-related adverse events; NSCLC, non-small-cell lung cancer; OR, odds ratio; OS, overall survival; SCLC, small-cell lung cancer.

value for interaction: 0.0496, 0.0027, and 0.0004, respectively). Concerning the selection of study population for the PD-L1 factor, a negative association between treatment effects on the overall grade 1-2 irAE rate and treatment effects on OS was observed in studies with the population selected for PD-L1 expression (that means: as irAE OR of experimental versus control arm increases, HR for OS of experimental versus control arm is predicted to decrease). The R^2 value of the weighted regression line was 0.55 (95% CI 0.20 - 0.95; P value = 0.03), indicating that 55% of the

variability among treatment effects on the OS is explained by the effects observed on the overall irAE rate of grade 1-2. Moreover, a negative low-strength correlation was observed in this trial subset ($R = -0.69$) (Figure 2B). In contrast, a positive association (that means: as irAE OR of experimental versus control arm increases, HR for OS of experimental versus control arm is predicted to increase) was observed between treatment effects on the endocrine grade 3-4 irAEs and treatment effects on OS in trials evaluating anti-PD1/PD-L1 antibody as experimental arm
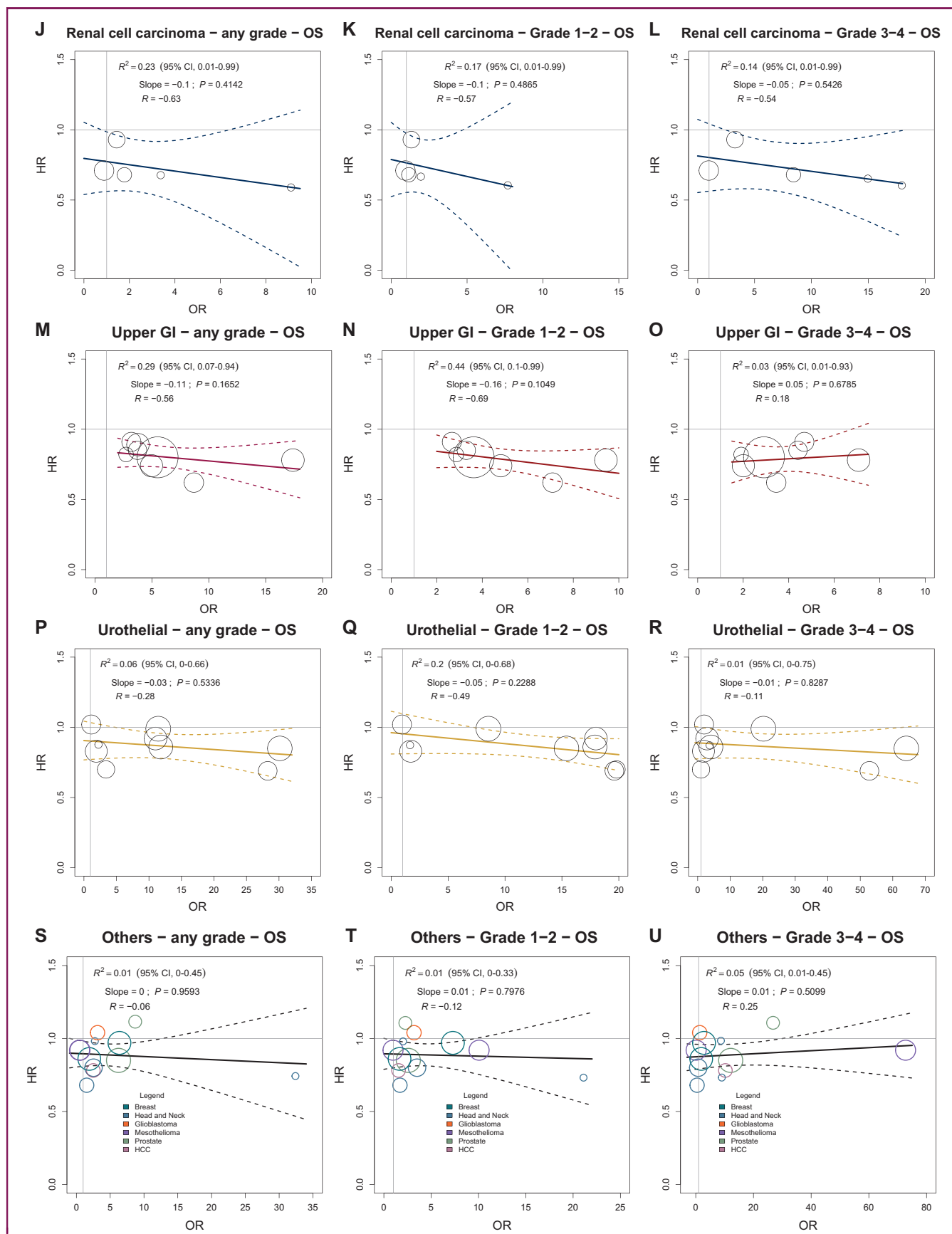
Figure 1. Continued.

($R^2 = 0.72$; 95% CI 0.10-0.98; P value $= 0.02$). A medium-strength correlation was estimated ($R = 0.76$) (Figure 2E). Only two trials that studied the combination of anti-PD1/PD-L1 and anti-CTLA4 antibodies as experimental arm were available, however, for this subgroup analysis (Figure 2F). Moreover, due to the large estimated R^2 CI values, all the subgroup analyses must be considered preliminary and warrant further validation.

SCLC

The analysis, regressing the log(HR) for OS on the log(OR) for irAEs using the complete set of data, did not demonstrate any association between the two effects for both the overall (any grade, grade 1-2, grade 3-4) irAE rates and the specific organ (skin, GI, endocrine) irAE rates (Figure 1D-F; Supplementary Figure S4, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel A to I).

Sensitivity and validation analysis gave similar results for irAEs of grade 1-2 and 3-4. The regression coefficients obtained from the main, sensitivity, and validation analyses were similar. As regards the irAEs of any grade, validation analysis results were in line with those obtained in the main analysis. In contrast, sensitivity analysis gave discordant results, but the large 95% CIs suggest a low precision in the regression coefficient estimation (Supplementary Figure S3, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel B).

Concerning the interaction analysis, the surrogacy was significantly different according to the type of experimental agent. In all the comparisons, however, at least one trial subgroup (i.e. the trial subset with an anti-CTLA-4 antibody as experimental arm) included only two randomized trials. Therefore, this finding deserves further validation (Supplementary Figure S5, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel A to L).

Melanoma

A total of 10 out of 17 contrasts (58.8%) compared two different ICI-based regimens, whereas 7 contrasts (41.2%) had a control arm that was not ICI-based (chemotherapy or gp100 peptide vaccine). Anti-CTLA-4, anti-PD-1/PD-L1, and a combination of anti-PD-1/PD-L1 and anti-CTLA-4 agents were employed as experimental arms in eight (47.0%), eight (47.0%), and one contrast (5.9%), respectively (Supplementary Table S1, available at https://doi.org/10.1016/j.esmoop.2023.100787).

The analysis, regressing the log(HR) for OS on the log(OR) for irAE using the complete set of data, demonstrated no association between the two effects for either the overall (any grade, grade 1-2, grade 3-4) irAE rates or the specific organ (skin, GI, endocrine) irAE rates (Figure 1G-I; Supplementary Figure S6, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel A to I). Sensitivity and validation analysis gave similar results for the irAEs of grade 3-4. Regarding the irAEs of any grade, validation analysis showed opposite results compared with those from the primary analysis. We estimated, however, a large 95% CI for

the regression coefficient in the validation analysis (Supplementary Figure S3, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel C).

Regarding interaction analysis, the association between treatment effects was significantly different according to the type of comparison and experimental ICI agent (P value for interaction: 0.0365, 0.0114, 0.0082, 0.0490, 0.0347, and 0.0082, for overall grade 3-4, skin any grade and grade 1-2, GI any grade, grade 1-2 and grade 3-4 irAE rate, respectively).

The regression lines, estimated separately for the type of control arm, indicated a weak association between treatment effects on the overall grade 3-4 irAE rate and treatment effects on OS in studies with a non-ICI-based control arm ($R^2 = 0.44$, 95% CI 0.17-0.99; slope $= -0.11$; P value $= 0.15$) (Figure 3A). We observed a significant negative association between treatment effects on the GI grade 3-4 irAE rate and treatment effects on OS in studies with a non-ICI-based control arm. The R^2 value of the weighted regression line was 0.77 (95% CI 0.24-0.99; P value $= 0.02$), indicating that 77% of the variability among treatment effects on OS was explained by the observed effects on the GI grade 3-4 irAE rate. In this case, we estimated a high-strength correlation that reached the cut-off level for surrogacy ($R = -0.89$) (Figure 3C). Due to the large estimated R^2 CI values, however, these findings must be considered preliminary and warrant further validation.

In the trial subset with an anti-CTLA-4 single agent as experimental arm, we observed negative low-strength correlations between treatment effects on the skin any grade and grade 1-2 irAEs, the GI any grade and grade 1-2 irAE rate, and treatment effects on OS ($R = -0.47$, -0.47, -0.63, and -0.52, respectively) (Figure 3E, G, I, and K). These correlations did not reach the cut-off level for claiming surrogacy despite the significant interaction of the relationship among treatment effects according to the type of experimental ICI agent.

Renal cell carcinoma

The analysis, regressing the log(HR) for OS on the log(OR) for irAE using the complete set of data, did not demonstrate any association between the two effects for either the overall (any grade, grade 1-2, grade 3-4) irAE rates or the specific organ (skin, GI, endocrine) irAE rates (Figure 1J-L; Supplementary Figure S7, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel A to I). Sensitivity and validation analyses gave similar results (Supplementary Figure S3, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel D). Due to missing values for trials with treatment line ≥2 and comparison between two ICI-based regimens, we did not carry out the interaction analysis.

Upper GI

The analysis, regressing the log(HR) for OS on the log(OR) for irAEs using the complete set of data, did not demonstrate any significant association between the two effects
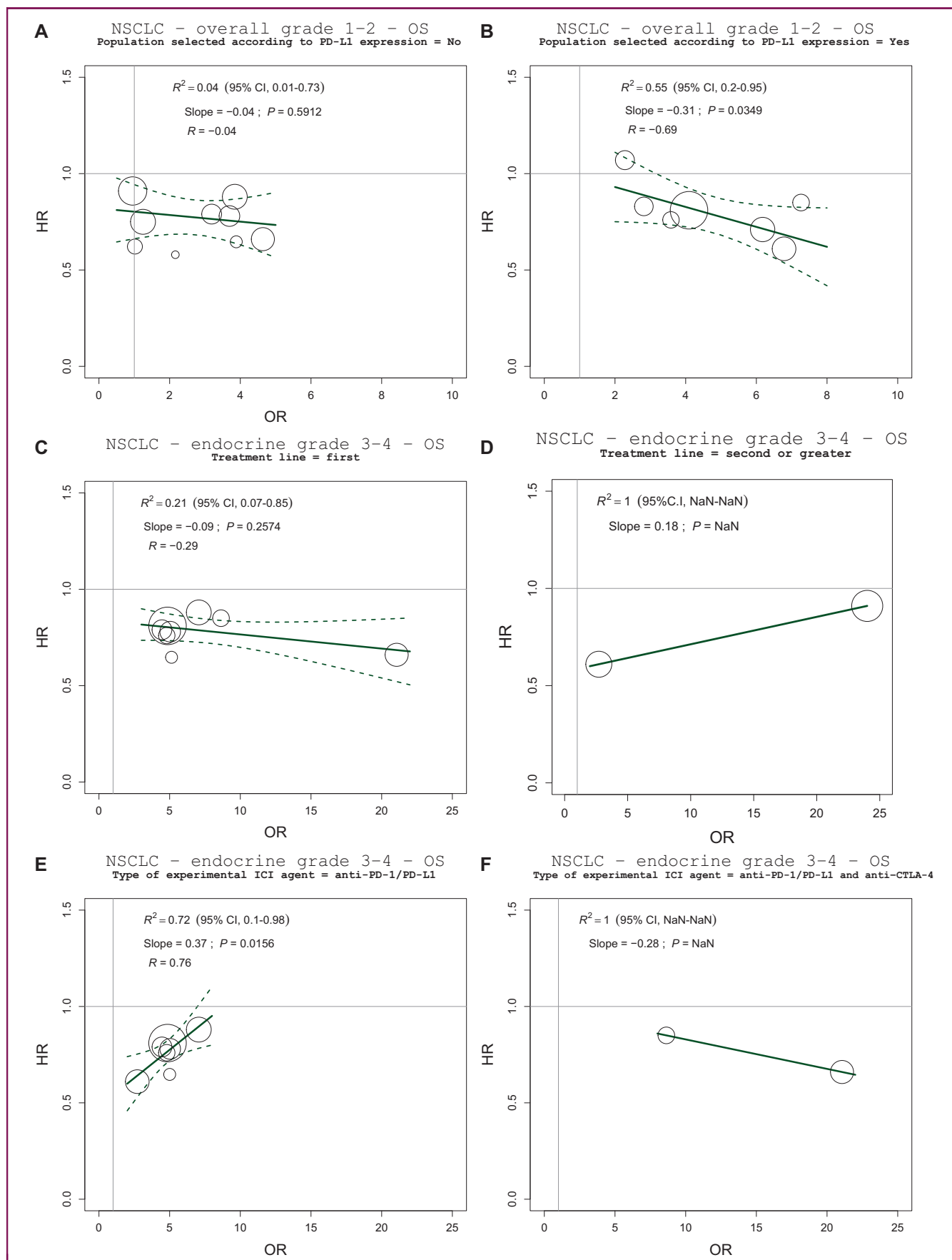
Figure 2. Explorative interaction analyses according to prespecified factors in NSCLC trial subset. Treatment effects are expressed as ORs for irAEs and HRs for OS. Every circle represents a comparison of the experimental group versus the control group, with the size of the circles representing the weight of the comparison,

for either the overall (any grade, grade 1-2, grade 3-4) irAE rates or the specific organ (skin, GI, endocrine) irAE rates (Figure 1M-O; Supplementary Figure S8, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel A to I). We observed a weak association, not statistically significant, between treatment effects on the overall grade 1-2 irAE rate and treatment effects on OS. The R^2 value of the weighted regression line was 0.44 (95% CI 0.10-0.99; slope $= -0.16$; P value $= 0.10$). The correlation between treatment effects was of low strength ($R = -0.69$) (Figure 1N). Due to the low number of available studies, the validation analysis was not carried out (Supplementary Figure S3, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel E).

The interaction analysis showed that surrogacy was not different according to the treatment line factor. The other stratification factors were not considered in the interaction analysis.

Since we could not examine the impact of missing data imputation on the primary analysis, these findings must be regarded as preliminary and warrant further validation.

Urothelial cancer

The analysis, regressing the log(HR) for OS on the log(OR) for irAEs using the complete set of data, demonstrated no association between the two effects for both the overall (any grade, grade 1-2, grade 3-4) irAE rates and the specific organ (skin, GI, endocrine) irAE rates (Figure 1P-R; Supplementary Figure S9, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel A to I). Sensitivity and validation analyses gave similar results for the irAEs of any grade and grade 3-4 (Supplementary Figure S3, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel F). Due to the low number of trials, the interaction analysis was not carried out.

Other tumors

We grouped ICI therapy randomized studies of cancer types with fewer than four trials (head and neck cancer, breast cancer, prostate cancer, hepatocellular carcinoma, glioblastoma, and mesothelioma). In this subset of 12 contrasts, the experimental arm was an anti-PD-1/PD-L1 antibody in six comparisons, an anti-CTLA-4 antibody in three studies, and a combination of anti-PD-1/PD-L1 and anti-CTLA-4 agents in 3 contrasts (Supplementary Table S1, available at https://doi.org/10.1016/j.esmoop.2023.100787).

The analysis, regressing the log(HR) for OS on the log(OR) for irAE using the complete set of data, did not demonstrate any association between the two effects for either overall (any grade, grade 1-2, grade 3-4) irAE rates or specific organ (skin, GI, endocrine) irAE rates (Figure 1S-U; Supplementary Figure S10, available at https://doi.org/10.1016/j.esmoop.

2023.100787, panel A to I). Sensitivity and validation analyses gave similar results. The regression coefficients obtained from the main, sensitivity, and validation analyses were similar according to the immune-related toxicity grade. Also, a large 95% CI of the regression coefficient was observed in sensitivity analysis for irAEs of any grade (Supplementary Figure S3, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel G).

The surrogacy was significantly different according to the type of experimental agent and control arm (ICI-based or not) factors (P value for interaction: 0.0347 and 0.0499, respectively).

The regression lines, estimated separately for the different experimental regimens, indicated a positive association, not statistically significant, between treatment effects on the overall any grade irAE rate and treatment effects on OS in the trial subset assessing an anti-PD-1/PD-L1 antibody as experimental arm ($R^2 = 0.52$, 95% CI 0.08-0.99; slope $= 0.19$; P value $= 0.10$; $R = 0.73$) (Supplementary Figure S11, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel B). In the trial subset with the combination of anti-PD-1/PD-L1 and anti-CTLA-4 antibodies as experimental arm, based on three randomized studies, we observed a weak association, not statistically significant, between treatment effects on the overall any grade irAE rate and treatment effects on OS ($R^2 = 0.60$, 95% CI 0.09-0.90; slope $= -0.05$; P value $= 0.44$; $R = -0.8$) (Supplementary Figure S11, available at https://doi.org/10.1016/j.esmoop.2023.100787, panel C).

In the subset of three trials comparing two different ICI-based regimens, the regression line showed a weak association, not statistically significant, between treatment effects on the overall grade 1-2 irAEs and treatment effects on OS ($R^2 = 0.79$, 95% CI 0.31-0.99; slope $= -0.08$; P value $= 0.31$; $R = -0.91$) (Supplementary Figure S11, available at https://doi.org/10.1016/j.esmoop.2023.100787; panel E).

DISCUSSION

Several clinical studies and an individual patient data (IPD) meta-analysis have shown an association between the occurrence of the irAEs and the survival outcomes of patients undergoing ICI therapy. To assess whether a new intervention significantly affects survival based on the effect on a surrogate measure, however, a reliable and plausible prediction equation should be estimated by a meta-analysis of several randomized trials.[87]

The present meta-analysis collected randomized studies that tested ICI therapy across several advanced-stage cancers, such as NSCLC, SCLC, melanoma, renal cell carcinoma, urothelial, and upper GI cancers. The median follow-up of the included studies (19.0 months; IQR, 11.8-27.5 months) was numerically longer than previous meta-analyses

---

proportional to the number of events (deaths). The straight lines represent the weighted regressions, which show the effects on OS predicted by the observed effects on irAEs.

CI, confidence interval; CTLA-4, cytotoxic T-lymphocyte antigen 4; HR, hazard ratio; ICI, immune checkpoint inhibitor; irAEs, immune-related adverse events; NaN, not a number; NSCLC, non-small-cell lung cancer; OR, odds ratio; OS, overall survival; PD-1, programmed cell death protein 1; PD-L1, programmed death-ligand 1.
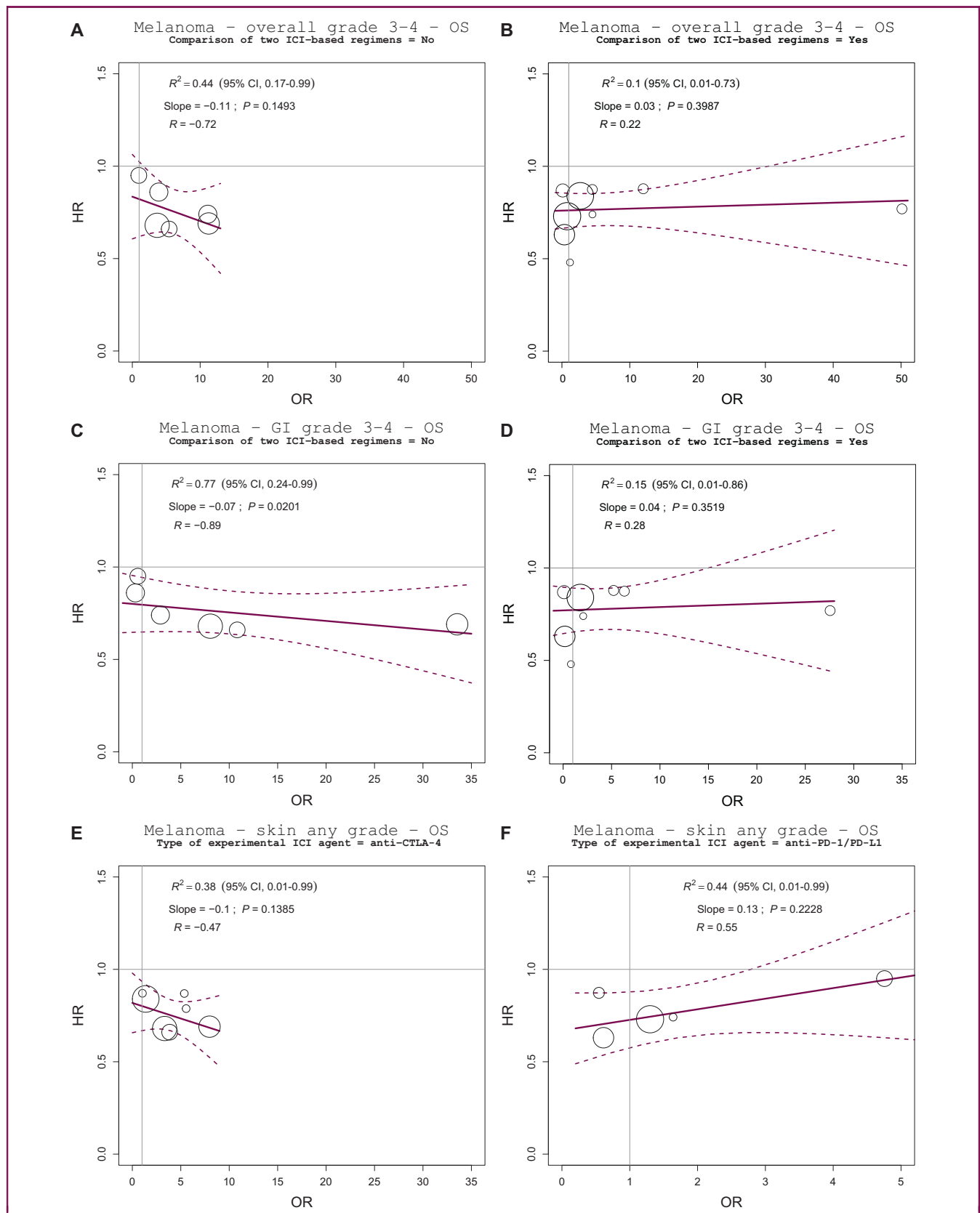
Figure 3. Explorative interaction analyses according to prespecified factors in melanoma trial subset. Treatment effects are expressed as ORs for irAEs and HRs for OS. Every circle represents a comparison of the experimental group versus the control group, with the size of the circles representing the weight of the comparison, proportional to the number of events (deaths). The straight lines represent the weighted regressions, which show the effects on OS predicted by the observed effects on irAEs.

CI, confidence interval; CTLA-4, cytotoxic T-lymphocyte antigen 4; HR, hazard ratio; ICI, immune checkpoint inhibitor; irAEs, immune-related adverse events; OR, odds ratio; OS, overall survival; PD-1, programmed cell death protein 1; PD-L1, programmed death-ligand 1.
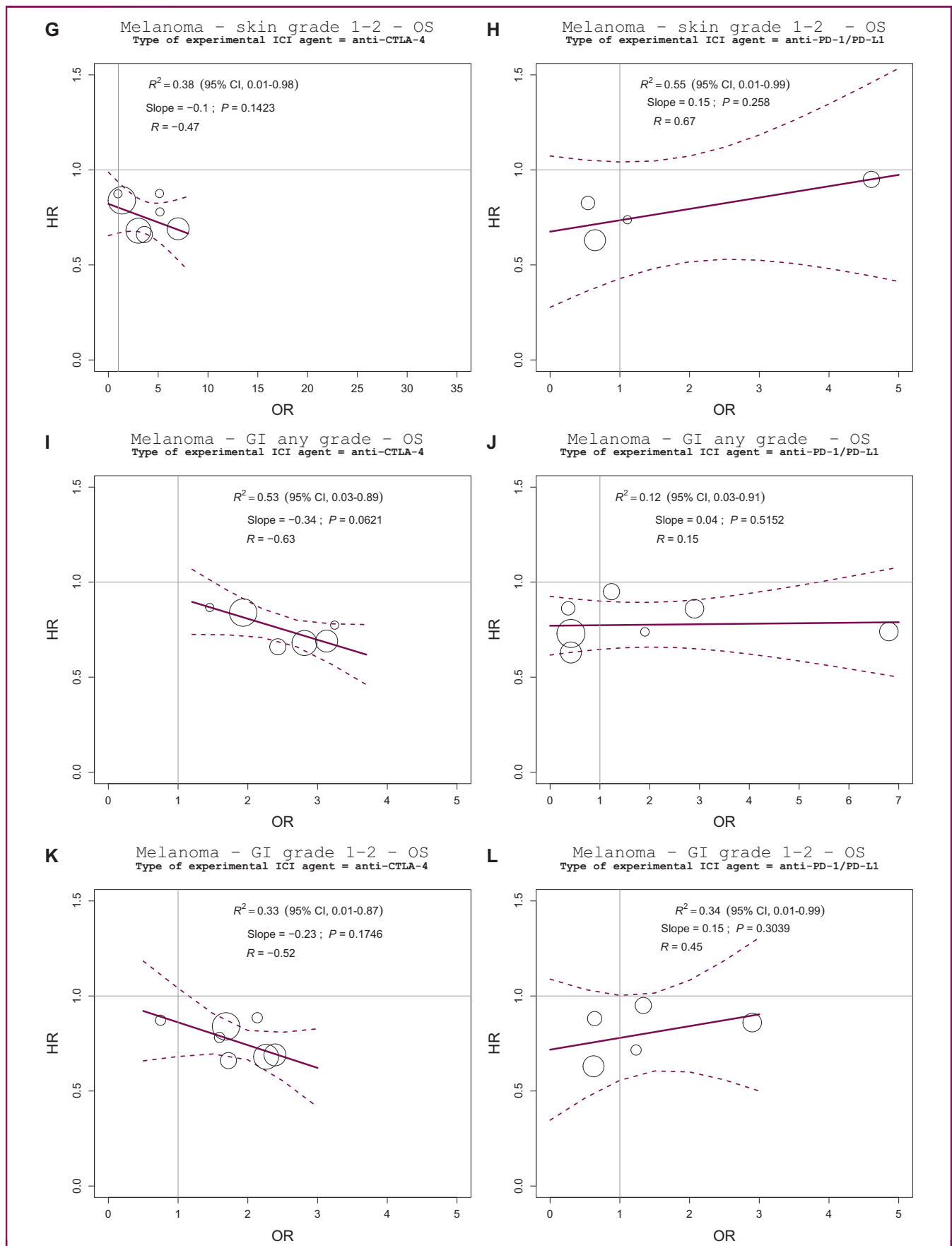
G Melanoma – skin grade 1–2 – OS
Type of experimental ICI agent = anti-CTLA-4

$R^2 = 0.38$ (95% CI, 0.01-0.98)
Slope = −0.1 ; $P = 0.1423$
$R = -0.47$

H Melanoma – skin grade 1–2 – OS
Type of experimental ICI agent = anti-PD-1/PD-L1

$R^2 = 0.55$ (95% CI, 0.01-0.99)
Slope = 0.15 ; $P = 0.258$
$R = 0.67$

I Melanoma – GI any grade – OS
Type of experimental ICI agent = anti-CTLA-4

$R^2 = 0.53$ (95% CI, 0.03-0.89)
Slope = −0.34 ; $P = 0.0621$
$R = -0.63$

J Melanoma – GI any grade – OS
Type of experimental ICI agent = anti-PD-1/PD-L1

$R^2 = 0.12$ (95% CI, 0.03-0.91)
Slope = 0.04 ; $P = 0.5152$
$R = 0.15$

K Melanoma – GI grade 1–2 – OS
Type of experimental ICI agent = anti-CTLA-4

$R^2 = 0.33$ (95% CI, 0.01-0.87)
Slope = −0.23 ; $P = 0.1746$
$R = -0.52$

L Melanoma – GI grade 1–2 – OS
Type of experimental ICI agent = anti-PD-1/PD-L1

$R^2 = 0.34$ (95% CI, 0.01-0.99)
Slope = 0.15 ; $P = 0.3039$
$R = 0.45$

Figure 3. Continued.

assessing the surrogacy of other intermediate endpoints in ICI-based trials.[17,19,88]

The present analysis showed correlations between the ICI therapy effects on the overall irAE rate and treatment effects on OS of varying magnitude (R ranged from −0.63 to 0.25) in different cancer types. All estimated correlations, however, were of low strength, according to the predefined R values necessary for validating a candidate surrogate endpoint. The coefficients of determination of the calculated linear regression models (R^2) were between 0.01 and 0.29, meaning that, in general, less than one-third of the variability of ICI therapy survival benefit could be explained by the variability in the treatment effects on the irAE rate.

We aimed to explore the surrogacy of the irAEs starting from the biological plausibility of the relationship between the irAEs and the OS outcome.[89] The biological hypothesis underlying the potential surrogacy of irAEs in patients treated with immunotherapy is that the stronger the autoimmune response induced by ICI therapy is, the more robust is antitumor immunity. In that case, the irAE development could lie in the main causal pathway of the disease process and thus would entirely capture the ICI effect on the OS. In addition, it is possible that certain irAEs are more directly related to antitumor efficacy than others, so that, along this line, not all irAEs are required to occur to obtain a benefit from ICI therapy.[90] Therefore, in the primary analysis, we explored the surrogacy of irAEs according to the severity (grade 1-2 or grade 3-4) and specific organ involvement (skin, GI, and endocrine system). We only observed a low-strength negative correlation (R = −0.69) among the treatment effects on the overall grade 1-2 irAE rate and treatment effects on OS in the upper GI trial subset, which was not statistically significant (P value = 0.10).

Considering that a candidate surrogate should be specific for the disease setting and the type of treatment administered, we also planned an interaction analysis by patient selection for PD-L1 expression or not, class of drugs used (anti-PD-1/PD-L1 versus anti-CTLA-4 versus a combination of anti-PD-1/PD-L1 and anti-CTLA-4 antibodies), type of control arm (ICI-based or not) and treatment line. The results of the different analyses suggest some hypothetical correlations. In the NSCLC trial subset, we did show a significant negative correlation between the overall irAEs of grade 1-2 and OS in trials with patients selected for PD-L1 expression (R = −0.69). In the melanoma trial subgroup with a non-ICI-based control arm, we reported a negative medium-strength correlation between treatment effects on the overall grade 3-4 irAEs and treatment effects on OS (R = −0.72) that was not statistically significant and otherwise a significant high-strength negative correlation between treatment effects on GI grade 3-4 irAEs and treatment effects on OS (R = −0.89).

Immunotherapy with ICIs has significantly improved the outcomes of patients with different cancers, although biomarkers available to predict ICI efficacy remain unsatisfactory.[91] As a whole, our findings did not support the hypothesis of surrogacy of irAEs in cancer patients undergoing immunotherapy. Therefore, irAEs' development possibly captures only part of the mechanism of action by which ICI therapy improves survival. Other important determinants of ICI efficacy could be related, for example, to tumor characteristics, such as PD-L1 expression, tumor mutational burden, tumor-infiltrating lymphocytes, or specific gene signatures.[91] The weak surrogacy of grades 1 or 2 irAEs confined to NSCLC patients with PD-L1 positivity, observed in our study, goes in this direction.

Indeed, our results suggest also that ICI treatment effects on grade 3 or 4 irAE rates might be positively associated with treatment effects on OS in at least some patient subsets. This observation might be explained by the fact that cancer patients can experience severe morbidity and increased mortality related to immune toxicities, which may confound the survival difference between patients with and without irAEs. In addition, treatment interruptions or discontinuations for severe immune toxicities may reduce the effectiveness of ICI therapy.[92] Clinicians should promptly manage the lower grade irAEs with supportive care as recommended by current guidelines,[6] to prevent the development of more severe toxicities.

The present study has several limitations. Firstly, we relied on data summarized in the published literature. One limit of using AD in meta-analyses is that errors in estimating treatment effects on the surrogate and the long-term outcome cannot be fully accounted for, resulting in a systematic bias toward lower association values.[93] Moreover, the irAE rate is a binary endpoint whose correlation with OS can be quantified with HRs, log-rank test, or survival ORs from a joint copula only in IPD and not in AD meta-analyses.[20]

Secondly, the premise of our trial-based surrogate validation study was that the 'outcome surrogacy' of the irAEs (that is, the demonstration of a strong correlation between the irAE occurrence and OS) had been adequately demonstrated in different cancer types.[9,11,94,95] In melanoma patients receiving ipilimumab, however, OS was not affected by the occurrence of irAEs in a retrospective and mono-institutional study.[96] Therefore, whether the irAE development correlates with improved outcomes is partly controversial.

Another limitation concerns the assessment of irAEs, which cannot be subject to verification. The irAEs are usually reported by patients, noted, and graded by investigators according to the CTCAE system. We collected only clinician-reported AEs that were classified as irAEs in the published studies. Patients, however, generally report adverse symptoms earlier and more frequently than clinicians. Consequently, the most accurate information on irAEs would derive from the patient, using validated patient-reported outcome (PRO) CTCAEs.[97] Nevertheless, there is a need to strengthen the validity of PRO measures used to monitor patients receiving ICI therapy.[98] In addition, it may be difficult to distinguish the immune pathogenesis of an adverse event from other causes, especially if the ICI is combined with other therapies. This bias might have influenced the results of our analyses.

In the present meta-analysis, we used the HR of death between the experimental and control arm as the treatment effect estimate on the clinical outcome OS. Reporting HRs may not adequately represent the ICI treatment benefit due to delayed separation or crossing of survival curves observed in many randomized trials of ICI therapy.[99] We cannot rule out that further analyses with more studies might better account for the potential loss of power due to the non-proportionality of survival curves in ICI trials. Moreover, we planned to explore the heterogeneity of the trial-level associations according to specific trial characteristics, but many of these subgroup analyses are difficult to interpret due to the limited number of available studies.

Finally, we have set the irAE rate as a potential efficacy endpoint of ICI therapy in the present meta-analysis, although it is typically a safety endpoint in the included randomized studies. As intermediate efficacy endpoints are generally validated to replace OS as the primary endpoint in trials assessing novel therapies, this cannot be the case for the irAE rate. Therefore, our analyses have only the speculative intent to provide further valuable insight into the ICI therapy's mechanism of action and not suggest alternative primary endpoints for future clinical studies.[13]

In conclusion, the present study showed that the irAE rate should not be considered a valid surrogate for OS when evaluating ICI therapy efficacy in NSCLC, SCLC, melanoma, upper GI cancers, renal cell carcinoma, urothelial cancers, and other solid tumors. To our knowledge, this is the first trial-based meta-analysis addressing this intriguing topic. Further research using IPD from a significant number of randomized studies might contribute to examining in depth the association of the irAEs with long-term outcomes at both patient and trial level.

FUNDING

DISCLOSURE

REFERENCES

1. Koon H, Atkins M. Autoimmunity and immunotherapy for cancer. *N Engl J Med*. 2006;354(7):758-760.
2. Zimmer L, Goldinger SM, Hofmann L, et al. Neurological, respiratory, musculoskeletal, cardiac and ocular side-effects of anti-PD-1 therapy. *Eur J Cancer*. 2016;60:210-225.
3. Hofmann L, Forschner A, Loquai C, et al. Cutaneous, gastrointestinal, hepatic, endocrine, and renal side-effects of anti-PD-1 therapy. *Eur J Cancer*. 2016;60:190-209.
4. Boutros C, Tarhini A, Routier E, et al. Safety profiles of anti-CTLA-4 and anti-PD-1 antibodies alone and in combination. *Nat Rev Clin Oncol*. 2016;13(8):473-486.
5. Haanen J, Carbonnel F, Robert C, et al. Management of toxicities from immunotherapy: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2018;29(suppl 4):iv264-iv266.
6. Schneider BJ, Naidoo J, Santomasso BD, et al. Management of immune-related adverse events in patients treated with immune checkpoint inhibitor therapy: ASCO Guideline Update. *J Clin Oncol*. 2021;39(36):4073-4126.
7. Gogas H, Ioannovich J, Dafni U, et al. Prognostic significance of auto-immunity during treatment of melanoma with interferon. *N Engl J Med*. 2006;354(7):709-718.
8. Sun X, Roudi R, Dai T, et al. Immune-related adverse events associated with programmed cell death protein-1 and programmed cell death ligand 1 inhibitors for non-small cell lung cancer: a PRISMA systematic review and meta-analysis. *BMC Cancer*. 2019;19(1):558.
9. Berner F, Bomze D, Diem S, et al. Association of checkpoint inhibitor-induced toxic effects with shared cancer and tissue antigens in non-small cell lung cancer. *JAMA Oncol*. 2019;5(7):1043-1047.
10. Kawai T, Sato Y, Makino K, et al. Immune-related adverse events predict the therapeutic efficacy of pembrolizumab in urothelial cancer patients. *Eur J Cancer*. 2019;116:114-115.
11. Socinski MA, Jotte RM, Cappuzzo F, et al. Pooled analyses of immune-related adverse events (irAEs) and efficacy from the phase 3 trials IMpower130, IMpower132, and IMpower150. *J Clin Oncol*. 2021;39(suppl 15):9002.
12. Buyse M, Molenberghs G. Criteria for the validation of surrogate endpoints in randomized experiments. *Biometrics*. 1998;54(3):1014-1029.
13. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med*. 1996;125(7):605-613.
14. Baker SG. Five criteria for using a surrogate endpoint to predict treatment effect based on data from multiple previous trials. *Stat Med*. 2018;37(8):1406.
15. Katz R. Biomarkers and surrogate markers: an FDA perspective. *NeuroRx*. 2004;1(2):189-195.
16. Zhang J, Liang W, Liang H, Wang X, He J. Endpoint surrogacy in oncological randomized controlled trials with immunotherapies: a systematic review of trial-level and arm-level meta-analyses. *Ann Transl Med*. 2019;7(11):244.
17. Kok PS, Cho D, Yoon WH, et al. Validation of progression-free survival rate at 6 months and objective response for estimating overall survival in immune checkpoint inhibitor trials: a systematic review and meta-analysis. *JAMA Netw Open*. 2020;3(9):e2011809.
18. Kaufman HL, Schwartz LH, William WN Jr, et al. Evaluation of classical clinical endpoints as surrogates for overall survival in patients treated with immune checkpoint blockers: a systematic review and meta-analysis. *J Cancer Res Clin Oncol*. 2018;144(11):2245-2261.
19. 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;24(10):2268-2275.
20. Xie W, Halabi S, Tierney JF, et al. A systematic review and recommendation for reporting of surrogate endpoint evaluation using meta-analyses. *JNCI Cancer Spectr*. 2019;3(1):pkz002.
21. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.
22. Ciani O, Davis S, Tappenden P, et al. Validation of surrogate endpoints in advanced solid tumors: systematic review of statistical methods, results, and implications for policy makers. *Int J Technol Assess Health Care*. 2014;30(3):312-324.
23. Validity of surrogate endpoints in oncology: Executive summary of rapid report A10-05, Version 1.1. *Institute for Quality and Efficiency in Health Care: Executive Summaries*. Cologne, Germany: Institute for Quality and Efficiency in Health Care (IQWiG) © IQWiG (Institute for Quality and Efficiency in Health Care). 2005.
24. 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(8):1389-1398.
25. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2021. Available at https://www.r-project.org/. Accessed November 1, 2021.

26. Kojima T, Shah MA, Muro K, et al. Randomized phase III KEYNOTE-181 study of pembrolizumab versus chemotherapy in advanced esophageal cancer. *J Clin Oncol*. 2020;38(35):4138-4148.

27. Chau I, Doki Y, Ajani JA, et al. Nivolumab (NIVO) plus ipilimumab (IPI) or NIVO plus chemotherapy (chemo) versus chemo as first-line (1L) treatment for advanced esophageal squamous cell carcinoma (ESCC): First results of the CheckMate 648 study. *J Clin Oncol*. 2021;39(suppl 18):LBA4001.

28. Chen LT, Satoh T, Ryu MH, et al. A phase 3 study of nivolumab in previously treated advanced gastric or gastroesophageal junction cancer (ATTRACTION-2): 2-year update data. *Gastric Cancer*. 2020;23(3):510-519.

29. Shitara K, Ozguroglu M, Bang YJ, et al. Pembrolizumab versus paclitaxel for previously treated, advanced gastric or gastro-oesophageal junction cancer (KEYNOTE-061): a randomised, open-label, controlled, phase 3 trial. *Lancet*. 2018;392(10142):123-133.

30. Shitara K, Van Cutsem E, Bang YJ, et al. Efficacy and safety of pembrolizumab or pembrolizumab plus chemotherapy vs chemotherapy alone for patients with first-line, advanced gastric cancer: the KEYNOTE-062 phase 3 randomized clinical trial. *JAMA Oncol*. 2020;6(10):1571-1580.

31. Janjigian YY, Shitara K, Moehler M, et al. First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal adenocarcinoma (CheckMate 649): a randomised, open-label, phase 3 trial. *Lancet*. 2021;398(10294):27-40.

32. Motzer RJ, Rini BI, McDermott DF, et al. Nivolumab for metastatic renal cell carcinoma: results of a randomized phase II trial. *J Clin Oncol*. 2015;33(13):1430-1437.

33. Motzer RJ, Rini BI, McDermott DF, et al. Nivolumab plus ipilimumab versus sunitinib in first-line treatment for advanced renal cell carcinoma: extended follow-up of efficacy and safety results from a randomised, controlled, phase 3 trial. *Lancet Oncol*. 2019;20(10):1370-1385.

34. Rini BI, Powles T, Atkins MB, et al. Atezolizumab plus bevacizumab versus sunitinib in patients with previously untreated metastatic renal cell carcinoma (IMmotion151): a multicentre, open-label, phase 3, randomised controlled trial. *Lancet*. 2019;393(10189):2404-2415.

35. Powles T, Plimack ER, Soulieres D, et al. Pembrolizumab plus axitinib versus sunitinib monotherapy as first-line treatment of advanced renal cell carcinoma (KEYNOTE-426): extended follow-up from a randomised, open-label, phase 3 trial. *Lancet Oncol*. 2020;21(12):1563-1573.

36. Motzer R, Alekseev B, Rha SY, et al. Lenvatinib plus pembrolizumab or everolimus for advanced renal cell carcinoma. *N Engl J Med*. 2021;384(14):1289-1300.

37. Choueiri TK, Powles T, Burotto M, et al. Nivolumab plus cabozantinib versus sunitinib for advanced renal-cell carcinoma. *N Engl J Med*. 2021;384(9):829-841.

38. Hodi FS, O'Day SJ, McDermott DF, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med*. 2010;363(8):711-723.

39. Maio M, Grob JJ, Aamdal S, et al. Five-year survival rates for treatment-naive patients with advanced melanoma who received ipilimumab plus dacarbazine in a phase III trial. *J Clin Oncol*. 2015;33(10):1191-1196.

40. Wolchok JD, Neyns B, Linette G, et al. Ipilimumab monotherapy in patients with pretreated advanced melanoma: a randomised, double-blind, multicentre, phase 2, dose-ranging study. *Lancet Oncol*. 2010;11(2):155-164.

41. Long GV, Schachter J, Ribas A, et al. 4-year survival and outcomes after cessation of pembrolizumab (pembro) after 2-years in patients (pts) with ipilimumab (ipi)-naive advanced melanoma in KEYNOTE-006. *J Clin Oncol*. 2018;36(suppl 15):9503.

42. Hamid O, Puzanov I, Dummer R, et al. Final analysis of a randomised trial comparing pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory advanced melanoma. *Eur J Cancer*. 2017;86:37-45.

43. Hodi FS, Chesney J, Pavlick AC, et al. Combined nivolumab and ipilimumab versus ipilimumab alone in patients with advanced melanoma: 2-year overall survival outcomes in a multicentre, randomised, controlled, phase 2 trial. *Lancet Oncol*. 2016;17(11):1558-1568.

44. Wolchok JD, Chiarion-Sileni V, Gonzalez R, et al. Overall survival with combined nivolumab and ipilimumab in advanced melanoma. *N Engl J Med*. 2017;377(14):1345-1356.

45. Ribas A, Kefford R, Marshall MA, et al. Phase III randomized clinical trial comparing tremelimumab with standard-of-care chemotherapy in patients with advanced melanoma. *J Clin Oncol*. 2013;31(5):616-622.

46. Weber JS, Gibney G, Sullivan RJ, et al. Sequential administration of nivolumab and ipilimumab with a planned switch in patients with advanced melanoma (CheckMate 064): an open-label, randomised, phase 2 trial. *Lancet Oncol*. 2016;17(7):943-955.

47. Ascierto PA, Del Vecchio M, Robert C, et al. Ipilimumab 10 mg/kg versus ipilimumab 3 mg/kg in patients with unresectable or metastatic melanoma: a randomised, double-blind, multicentre, phase 3 trial. *Lancet Oncol*. 2017;18(5):611-622.

48. Larkin J, Minor D, D'Angelo S, et al. Overall survival in patients with advanced melanoma who received nivolumab versus investigator's choice chemotherapy in CheckMate 037: a randomized, controlled, open-label phase III trial. *J Clin Oncol*. 2018;36(4):383-390.

49. Ascierto PA, Del Vecchio M, Mandala M, et al. Adjuvant nivolumab versus ipilimumab in resected stage IIIB-C and stage IV melanoma (CheckMate 238): 4-year results from a multicentre, double-blind, randomised, controlled, phase 3 trial. *Lancet Oncol*. 2020;21(11):1465-1477.

50. Horn L, Spigel DR, Vokes EE, et al. Nivolumab versus docetaxel in previously treated patients with advanced non-small-cell lung cancer: two-year outcomes from two randomized, open-label, phase III trials (CheckMate 017 and CheckMate 057). *J Clin Oncol*. 2017;35(35):3924-3933.

51. Carbone DP, Reck M, Paz-Ares L, et al. First-line nivolumab in stage IV or recurrent non-small-cell lung cancer. *N Engl J Med*. 2017;376(25):2415-2426.

52. Borghaei H, Langer CJ, Gadgeel S, et al. 24-Month overall survival from KEYNOTE-021 cohort G: pemetrexed and carboplatin with or without pembrolizumab as first-line therapy for advanced nonsquamous non-small cell lung cancer. *J Thorac Oncol*. 2019;14(1):124-129.

53. Brahmer JR, Rodriguez-Abreu D, Robinson AG, et al. Progression after the next line of therapy (PFS2) and updated OS among patients (pts) with advanced NSCLC and PD-L1 tumor proportion score (TPS) ≥50% enrolled in KEYNOTE-024. *J Clin Oncol*. 2017;35(suppl 15):9000.

54. Herbst RS, Baas P, Kim DW, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet*. 2016;387(10027):1540-1550.

55. Gandhi L, Rodriguez-Abreu D, Gadgeel S, et al. Pembrolizumab plus chemotherapy in metastatic non-small-cell lung cancer. *N Engl J Med*. 2018;378(22):2078-2092.

56. Socinski MA, Jotte RM, Cappuzzo F, et al. Atezolizumab for first-line treatment of metastatic nonsquamous NSCLC. *N Engl J Med*. 2018;378(24):2288-2301.

57. Jotte R, Cappuzzo F, Vynnychenko I, et al. Atezolizumab in combination with carboplatin and nab-paclitaxel in advanced squamous NSCLC (IMpower131): results from a randomized phase III trial. *J Thorac Oncol*. 2020;15(8):1351-1360.

58. Paz-Ares L, Luft A, Vicente D, et al. Pembrolizumab plus chemotherapy for squamous non-small-cell lung cancer. *N Engl J Med*. 2018;379(21):2040-2051.

59. Herbst RS, Giaccone G, de Marinis F, et al. Atezolizumab for first-line treatment of PD-L1-selected patients with NSCLC. *N Engl J Med*. 2020;383(14):1328-1339.

60. Rizvi NA, Cho BC, Reinmuth N, et al. Durvalumab with or without tremelimumab vs standard chemotherapy in first-line treatment of metastatic non-small cell lung cancer: the MYSTIC phase 3 randomized clinical trial. *JAMA Oncol*. 2020;6(5):661-674.

61. West H, McCleod M, Hussein M, et al. Atezolizumab in combination with carboplatin plus nab-paclitaxel chemotherapy compared with chemotherapy alone as first-line treatment for metastatic non-squamous non-small-cell lung cancer (IMpower130): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol*. 2019;20(7):924-937.

62. Paz-Ares L, Ciuleanu TE, Cobo M, et al. First-line nivolumab plus ipilimumab combined with two cycles of chemotherapy in patients with non-small-cell lung cancer (CheckMate 9LA): an international, randomised, open-label, phase 3 trial. *Lancet Oncol*. 2021;22(2):198-211.

63. Mok TSK, Wu YL, Kudaba I, et al. Pembrolizumab versus chemotherapy for previously untreated, PD-L1-expressing, locally advanced or metastatic non-small-cell lung cancer (KEYNOTE-042): a randomised, open-label, controlled, phase 3 trial. *Lancet*. 2019;393(10183):1819-1830.

64. Govindan R, Szczesna A, Ahn MJ, et al. Phase III trial of ipilimumab combined with paclitaxel and carboplatin in advanced squamous non-small-cell lung cancer. *J Clin Oncol*. 2017;35(30):3449-3457.

65. Liu SV, Reck M, Mansfield AS, et al. Updated overall survival and PD-L1 subgroup analysis of patients with extensive-stage small-cell lung cancer treated with atezolizumab, carboplatin, and etoposide (IMpower133). *J Clin Oncol*. 2021;39(6):619-630.

66. Reck M, Bondarenko I, Luft A, et al. Ipilimumab in combination with paclitaxel and carboplatin as first-line therapy in extensive-disease-small-cell lung cancer: results from a randomized, double-blind, multicenter phase 2 trial. *Ann Oncol*. 2013;24(1):75-83.

67. Rudin CM, Awad MM, Navarro A, et al. Pembrolizumab or placebo plus etoposide and platinum as first-line therapy for extensive-stage small-cell lung cancer: randomized, double-blind, phase III KEYNOTE-604 study. *J Clin Oncol*. 2020;38(21):2369-2379.

68. Goldman JW, Dvorkin M, Chen Y, et al. Durvalumab, with or without tremelimumab, plus platinum-etoposide versus platinum-etoposide alone in first-line treatment of extensive-stage small-cell lung cancer (CASPIAN): updated results from a randomised, controlled, open-label, phase 3 trial. *Lancet Oncol*. 2021;22(1):51-65.

69. Reck M, Luft A, Szczesna A, et al. Phase III randomized trial of ipilimumab plus etoposide and platinum versus placebo plus etoposide and platinum in extensive-stage small-cell lung cancer. *J Clin Oncol*. 2016;34(31):3740-3748.

70. Fradet Y, Bellmunt J, Vaughn DJ, et al. Randomized phase III KEYNOTE-045 trial of pembrolizumab versus paclitaxel, docetaxel, or vinflunine in recurrent advanced urothelial cancer: results of >2 years of follow-up. *Ann Oncol*. 2019;30(6):970-976.

71. Powles T, Duran I, van der Heijden MS, et al. Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial. *Lancet*. 2018;391(10122):748-757.

72. Powles T, van der Heijden MS, Castellano D, et al. Durvalumab alone and durvalumab plus tremelimumab versus chemotherapy in previously untreated patients with unresectable, locally advanced or metastatic urothelial carcinoma (DANUBE): a randomised, open-label, multicentre, phase 3 trial. *Lancet Oncol*. 2020;21(12):1574-1588.

73. Galsky MD, Arija JAA, Bamias A, et al. Atezolizumab with or without chemotherapy in metastatic urothelial cancer (IMvigor130): a multicentre, randomised, placebo-controlled phase 3 trial. *Lancet*. 2020;395(10236):1547-1557.

74. Powles T, Park SH, Voog E, et al. Avelumab maintenance therapy for advanced or metastatic urothelial carcinoma. *N Engl J Med*. 2020;383(13):1218-1230.

75. Powles T, Csoszi T, Ozguroglu M, et al. Pembrolizumab alone or combined with chemotherapy versus chemotherapy as first-line therapy for advanced urothelial carcinoma (KEYNOTE-361): a randomised, open-label, phase 3 trial. *Lancet Oncol*. 2021;22(7):931-945.

76. Schmid P, Rugo HS, Adams S, et al. Atezolizumab plus nab-paclitaxel as first-line treatment for unresectable, locally advanced or metastatic triple-negative breast cancer (IMpassion130): updated efficacy results from a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol*. 2020;21(1):44-59.

77. Winer EP, Lipatov O, Im SA, et al. Pembrolizumab versus investigator-choice chemotherapy for metastatic triple-negative breast cancer (KEYNOTE-119): a randomised, open-label, phase 3 trial. *Lancet Oncol*. 2021;22(4):499-511.

78. Reardon DA, Brandes AA, Omuro A, et al. Effect of nivolumab vs bevacizumab in patients with recurrent glioblastoma: the checkMate 143 phase 3 randomized clinical trial. *JAMA Oncol*. 2020;6(7):1003-1010.

79. Scherpereel A, Mazieres J, Greillier L, et al. Nivolumab or nivolumab plus ipilimumab in patients with relapsed malignant pleural mesothelioma (IFCT-1501 MAPS2): a multicentre, open-label, randomised, non-comparative, phase 2 trial. *Lancet Oncol*. 2019;20(2):239-253.

80. Maio M, Scherpereel A, Calabro L, et al. Tremelimumab as second-line or third-line treatment in relapsed malignant mesothelioma (DETERMINE): a multicentre, international, randomised, double-blind, placebo-controlled phase 2b trial. *Lancet Oncol*. 2017;18(9):1261-1273.

81. Kwon ED, Drake CG, Scher HI, et al. Ipilimumab versus placebo after radiotherapy in patients with metastatic castration-resistant prostate cancer that had progressed after docetaxel chemotherapy (CA184-043): a multicentre, randomised, double-blind, phase 3 trial. *Lancet Oncol*. 2014;15(7):700-712.

82. Beer TM, Kwon ED, Drake CG, et al. Randomized, double-blind, phase III trial of ipilimumab versus placebo in asymptomatic or minimally symptomatic patients with metastatic chemotherapy-naive castration-resistant prostate cancer. *J Clin Oncol*. 2017;35(1):40-47.

83. Finn RS, Ryoo BY, Merle P, et al. Pembrolizumab as second-line therapy in patients with advanced hepatocellular carcinoma in KEYNOTE-240: a randomized, double-blind, phase III trial. *J Clin Oncol*. 2020;38(3):193-202.

84. Ferris RL, Blumenschein G Jr, Fayette J, et al. Nivolumab vs investigator's choice in recurrent or metastatic squamous cell carcinoma of the head and neck: 2-year long-term survival update of CheckMate 141 with analyses by tumor PD-L1 expression. *Oral Oncol*. 2018;81:45-51.

85. Cohen EEW, Soulieres D, Le Tourneau C, et al. Pembrolizumab versus methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck squamous cell carcinoma (KEYNOTE-040): a randomised, open-label, phase 3 study. *Lancet*. 2019;393(10167):156-167.

86. Siu LL, Even C, Mesia R, et al. Safety and efficacy of durvalumab with or without tremelimumab in patients with PD-L1-Low/Negative recurrent or metastatic HNSCC: the phase 2 CONDOR randomized clinical trial. *JAMA Oncol*. 2019;5(2):195-203.

87. 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 Clin Epidemiol*. 2015;68(7):833-842.

88. Nie RC, Chen FP, Yuan SQ, et al. Evaluation of objective response, disease control and progression-free survival as surrogate end-points for overall survival in anti-programmed death-1 and anti-programmed death ligand 1 trials. *Eur J Cancer*. 2019;106:1-11.

89. Ciani O, Buyse M, Drummond M, Rasi G, Saad ED, Taylor RS. Use of surrogate end points in healthcare policy: a proposal for adoption of a validation framework. *Nat Rev Drug Discov*. 2016;15(7):516.

90. Postow MA, Sidlow R, Hellmann MD. Immune-related adverse events associated with immune Checkpoint blockade. *N Engl J Med*. 2018;378(2):158-168.

91. Sharma P, Siddiqui BA, Anandhan S, et al. The next decade of immune Checkpoint therapy. *Cancer Discov*. 2021;11(4):838-857.

92. Ksienski D, Wai ES, Croteau N, et al. Efficacy of nivolumab and pembrolizumab in patients with advanced non-small-cell lung cancer needing treatment interruption because of adverse events: a retrospective multicenter analysis. *Clin Lung Cancer*. 2019;20(1):e97-e106.

93. Buyse M, Molenberghs G, Paoletti X, et al. Statistical evaluation of surrogate endpoints with examples from cancer clinical trials. *Biom J*. 2016;58(1):104-132.

94. Eggermont AMM, Kicinski M, Blank CU, et al. Association between immune-related adverse events and recurrence-free survival among patients with stage III melanoma randomized to receive pembrolizumab or placebo: a secondary analysis of a randomized clinical trial. *JAMA Oncol*. 2020;6(4):519-527.

95. Tang K, Seo J, Tiu BC, et al. Association of cutaneous immune-related adverse events with increased survival in patients treated with anti-

programmed cell death 1 and anti-programmed cell death ligand 1 therapy. *JAMA Dermatol*. 2022;158(2):189-193.

96. Horvat TZ, Adel NG, Dang TO, et al. Immune-related adverse events, need for systemic immunosuppression, and effects on survival and time to treatment failure in patients with melanoma treated with ipilimumab at Memorial Sloan Kettering Cancer Center. *J Clin Oncol*. 2015;33(28):3193-3198.

97. Dueck AC, Mendoza TR, Mitchell SA, et al. Validity and reliability of the US national cancer institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). *JAMA Oncol*. 2015;1(8):1051-1059.

98. Da Silva Lopes AM, Colomer-Lahiguera S, Mederos Alfonso N, et al. Patient-reported outcomes for monitoring symptomatic toxicities in cancer patients treated with immune-checkpoint inhibitors: a Delphi study. *Eur J Cancer*. 2021;157:225-237.

99. Anagnostou V, Yarchoan M, Hansen AR, et al. Immuno-oncology trial endpoints: capturing clinically meaningful activity. *Clin Cancer Res*. 2017;23(17):4959-4969.