



Randomized Versus Real-World Evidence on the Efficacy and Toxicity of Checkpoint Inhibitors in Cancer in Patients with Advanced Non-small Cell Lung Cancer or Melanoma: A Meta-analysis

Evangelos Digkas¹ · Anthony Jagri Tabiim² · Daniel Smith³ · Antonis Valachis⁴ 

Accepted: 11 July 2022 / Published online: 1 August 2022
© The Author(s) 2022

Abstract

Background Both randomized controlled trials (RCTs) and real-world evidence (RWE) studies provide results regarding the efficacy and toxicity of checkpoint inhibitors in cancer patients. The results from these two sources are considered complementary but whether they are comparable remains unknown.

Objective The aim of this study was to compare the efficacy and toxicity of checkpoint inhibitors between RCTs and RWE studies in patients with advanced non-small cell lung cancer (NSCLC) or melanoma.

Patients and Methods Two electronic databases were searched to identify eligible studies, either RCTs or RWE studies, investigating the efficacy or toxicity of checkpoint inhibitors given for indications that were approved by the European Medicines Agency (EMA) at the date of the last search. A meta-analysis was performed and the pooled estimates of objective response rates (ORR), progression-free survival (PFS), overall survival (OS), and toxicity and treatment discontinuation between RCTs and RWE studies were compared.

Results In total, 43 RWE studies and 15 RCTs were eligible, with adequate data for pooled estimates for immunotherapy indications regarding NSCLC and melanoma. No statistically significant or clinically meaningful differences in terms of pooled PFS, OS, or rates of treatment discontinuation due to toxicity between RCTs and RWE studies were observed. In some indications, a higher rate of response rates and lower rate of toxicity in favor of RWE was observed.

Conclusion In patients with melanoma or NSCLC, the clinical value of checkpoint inhibitors is evident in both RCTs and real-world settings. Some differences in response or toxicity rates in favor of RWE mainly reflects the inherent difficulties in evaluating these outcomes in RWE studies.

Evangelos Digkas and Anthony Jagri Tabiim contributed equally to this work.

✉ Antonis Valachis
Antonios.valachis@oru.se

- ¹ Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden
- ² Department of Surgery, Korle-Bu Teaching Hospital, Accra, Ghana
- ³ Clinical Epidemiology and Biostatistics, School of Medical Sciences, Örebro University, Örebro, Sweden
- ⁴ Department of Oncology, Faculty of Medicine and Health, Örebro University, Örebro, Sweden

Key Points

By comparing the efficacy and toxicity of checkpoint inhibitors between randomized controlled trials (RCTs) and real-world evidence (RWE) studies, we found no statistically significant or clinically relevant differences in terms of progression-free survival or overall survival within the same indication.

In some indications, a higher rate of response rates and lower rate of toxicity in favor of RWE was observed that mainly reflects the inherent difficulties in evaluating these outcomes in RWE studies.

1 Introduction

Since 2010 and the first randomized evidence on improved overall survival (OS) in melanoma patients treated with ipilimumab [1], immunotherapy with checkpoint inhibitors has revolutionized cancer treatment. In fact, the number of indications for checkpoint inhibitors has increased substantially over the years [2] along with the estimated percentage of cancer patients eligible for checkpoint inhibitors in various cancer types and treatment settings [3].

Interestingly, the efficacy of checkpoint inhibitors varied among different cancer types, from substantial survival benefit in patients with advanced melanoma treated with combined nivolumab/ipilimumab [4], to modest and clinically questionable survival benefit as second-line therapy in recurrent squamous cell carcinoma of the head and neck [5] or in advanced urothelial carcinoma [6]. Varying efficacy along with toxicity risk and economic burden raises concerns on the implementation of immunotherapy in clinical practice.

The implementation and economic assessment of immunotherapy in clinical practice is based on evidence from randomized controlled trials (RCTs). In these trials, strict inclusion and exclusion criteria are adopted to define the patient population. Although RCTs are the gold-standard practice to evaluate the efficacy and toxicity of new treatment strategies, the generalizability of their results in a real-world setting can be questioned since nearly two-thirds of patients in clinical settings would not be eligible for randomized trials [7–9].

One could argue that this potential discrepancy between randomized and real-world evidence (RWE) is more likely to be a source of concern in indications where the benefit of a new treatment strategy according to randomized evidence is modest. Furthermore, health economic analyses using assumptions only from randomized evidence might also be prone to misleading results regarding the cost effectiveness of a new therapeutic agent in real-world settings if the discrepancy between randomized evidence used in assumptions and RWE is considerable [10].

The aim of the present systematic review and meta-analysis was to assess and compare the efficacy and toxicity of checkpoint inhibitors between randomized controlled trials (RCTs) and RWE studies in all current treatment indications for checkpoint inhibitors according to the European Medicines Agency (EMA).

2 Methods

2.1 Search Strategy

The present systematic review and meta-analysis was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The study protocol was prospectively published to the PROSPERO database (CRD42020180883).

Two separate search strategies were used to identify RCTs and RWE studies. For RCTs, two investigators (ED, AJT) searched the PubMed and ISI Web of Science databases independently using different algorithms including the following keywords: immunotherapy, checkpoint inhibitor, PD-1 inhibitor, PD-L1 inhibitor, RCT, randomized, randomized controlled trial, random*, cancer, and malignancy. For RWE studies, two investigators (ED, AJT) searched the PubMed and ISI Web of Science databases using different algorithms including the following keywords: immunotherapy, checkpoint inhibitor, PD-1 inhibitor, PD-L1 inhibitor, real-world evidence, real-world, real-world setting, real-world data, real-life, RWD, and RWE.

No year restriction was set but only studies published in English were considered eligible. The last search was performed in May 2020.

To complement the search strategy, reference lists from selected systematic reviews on this topic, as well as reference lists from eligible studies, were scrutinized for potentially eligible studies.

2.2 Study Selection Process

Inclusion criteria were used to screen the studies based on their title and abstract. Full-text articles were evaluated for eligibility using the following inclusion and exclusion criteria. For RCTs, studies including patients with metastatic disease of a solid malignancy; treated with a checkpoint inhibitor as a sole therapy or in combination with another checkpoint inhibitor in at least one arm; the checkpoint inhibitor is approved by the EMA (at the time of searching) for clinical use; and available data for at least one of the efficacy outcomes. For RWE studies, studies including patients with metastatic disease of solid malignancy; in cases of multiple tumors, a separate analysis for each tumor type should be available; treated with checkpoint inhibitors as a sole therapy or in combination with another checkpoint inhibitor; the checkpoint inhibitor is approved by the EMA (at the time of searching) for clinical use; and available data for at least one of the efficacy outcomes of interest.

We excluded studies that investigated the efficacy of checkpoint inhibitors outside of current EMA indications,

studies that investigated hematologic malignancies, and studies with only toxicity data and no efficacy data.

In case of discrepancies between the two investigators, a third investigator (AV) was consulted and consensus was reached among the three investigators regarding study eligibility.

2.3 Data Collection Strategy

Data were extracted independently by two investigators (ED, AJT), while a third investigator (AV) resolved any discrepancies. From each eligible trial, the following data were extracted: first author, year of publication, journal; inclusion period, number of patients, type of cancer, type of checkpoint inhibitor used, line of treatment; median follow-up, efficacy outcomes for patients treated with checkpoint inhibitors (objective response rate [ORR], median progression-free survival [PFS], median OS, rate of grade 3/4 toxicities due to checkpoint inhibitors [grading as per study definition], and rate of discontinuation of checkpoint inhibitors). The median PFS or OS was extracted from Kaplan–Meier curves if it was not stated in the manuscript.

2.4 Outcomes and Definitions

Three efficacy outcomes (ORR, PFS, OS) and two safety outcomes (rate of grade 3 or 4 toxicity, rate of discontinuation) were analyzed in the present meta-analysis.

ORR was defined as the rate of complete response or partial response according to the Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 criteria for RCTs. In RWE studies, all definitions that were used for ORR were accepted. PFS was defined as the time from initiation of checkpoint inhibitors until disease progression or death due to any cause, while OS was defined as the time from initiation of checkpoint inhibitors until death due to any cause.

Rate of grade 3 or 4 toxicity (RG 3–4 Tox) was defined as any treatment-related toxicity observed during treatment with checkpoint inhibitors and reported as grade 3 or 4 according to the Common Terminology Criteria for Adverse Events (CTCAE) criteria (irrespective of the version used in each study). Rate of discontinuation to checkpoint inhibitors (RODI) was defined as the number of patients who did not complete their planned treatment with checkpoint inhibitors due to toxicity.

2.5 Data Synthesis

Statistical analysis was performed in R (version 4.1.0) using R studio (version 1.4.1717) [R code to fully reproduce the analysis is included as electronic supplementary material (ESM)]. We prespecified select combinations of outcome type, study type, cancer type, treatment type and treatment

Table 1 Pooled rates of grade 3–4 immune-related toxicity and discontinuation in RCTs and RWE studies in different clinical indications of immunotherapy in oncology

Indications	Pooled rates	95% CI
Grade 3–4 immune-related toxicity		
First-line NSCLC		
RCT	19.2	14.3–25.2
RWE	NC	NC
Second-line NSCLC		
RCT	12.2	9.4–15.7
RWE	8.1	6.9–9.5
Second-line melanoma		
RCT	19.6	15.8–24.0
RWE	10.2	7.1–14.3
Ipilimumab in melanoma		
RCT	20.5	16.8–25.0
RWE	NC	NC
Discontinuation of immunotherapy due to toxicity		
2nd-line NSCLC		
RCT	6.7	4.9–9.2
RWE	9.5	7.8–11.4
Second-line melanoma		
RCT	7.1	5.3–9.5
RWE	NC	NC
Ipilimumab in melanoma		
RCT	12.3	9.4–15.7
RWE	NC	NC

NSCLC non-small cell lung cancer, *NC* not calculated, *RCTs* randomized controlled trials, *RWE* real-world evidence, *CI* confidence interval

line, hereafter referred to as *subsets* (Table 1). We used multilevel meta-analytic models to compute pooled metrics and 95% confidence intervals (CIs) for each subset containing at least three estimates from primary studies.

2.5.1 Median Survival Outcomes

OS and PFS were analyzed using a linear mixed-effects model assuming a Gaussian error distribution, with weighting using the inverse variance method. The response variable was the logarithm of the median survival time. We estimated standard errors from the CIs provided in primary studies. Where both upper- and lower-bound estimates were available, we used the full width of the CI to estimate the standard error of the log median survival (Eq. 1):

$$\hat{s}e_{\log} = \frac{\log\{\hat{\theta}_U\} - \log\{\hat{\theta}_L\}}{2 \times 1.96} \quad (1)$$

where $\hat{\theta}_U$ and $\hat{\theta}_L$ denote the upper- and lower-bound estimates of the median survival. If both the upper- and

lower-bound estimates were unavailable, we used the half width of the CI between the lower-bound and point estimate (Eq. 2):

$$\hat{s}e_{\log} = \frac{\log\{\hat{\theta}\} - \log\{\hat{\theta}_L\}}{1.96} \quad (2)$$

where $\hat{\theta}$ denotes the point estimate of the median survival.

Parameters were estimated using restricted maximum likelihood (REML). The subset was modelled as a fixed effect using a series of dummy-coded binary variables. We included study- and observation-level random effects as well as a random slope for the subset. The selected model used a block diagonal variance-covariance matrix that was shown to improve model fit compared with an identity variance-covariance matrix (comparison of the two nested models fitted using maximum likelihood estimation: likelihood ratio test: Chi-square = 33.42, $p = 0.003$). Pooled metrics (i.e. the mean of the median survival) $\pm 95\%$ CIs were back-transformed for presentation. Prespecified contrasts were conducted on the logarithmic scale, with CIs adjusted for simultaneous inference.

2.5.2 Proportion Outcomes

ORR, RG 3–4 Tox, and RODI were analyzed using a generalized linear mixed-effects model assuming a binomial error distribution with a logit link function. Parameters were estimated using maximum likelihood. The response variable was the number of cases, while the binomial denominator was the sample size for each observation. As above, subset was modelled as a fixed effect using a series of dummy-coded binary variables. We included study- and observation-level random effects (the latter was shown to improve performance for a number of diagnostic checks on residuals simulated from the model). Pooled metrics (i.e. the mean logit proportions) $\pm 95\%$ CIs were back-transformed for presentation.

3 Results

3.1 Literature Search

A total of 9551 studies were initially identified through searching algorithms. Through exclusion by reading the title and/or abstract, 194 RWE studies and 35 RCTs were considered as potentially eligible and retrieved as full text. After reading the full text, 64 RWE studies and 20 RCTs fulfilled the inclusion criteria. From all the studies that fulfilled the inclusion criteria, 43 RWE studies and 15 RCTs were included in the meta-analysis. A flowchart of the study selection process is presented in Fig. 1. The

study characteristics for eligible studies included in the meta-analysis are shown in ESM Table 1.

The indications of checkpoint inhibitors with adequate data from pooled analyses in both RWE studies and RCTs were in advanced melanoma and non-small cell lung cancer (NSCLC) either as first or second or later line.

3.2 Objective Response Rates in Randomized Controlled Trials (RCTs) and Real-World Evidence (RWE) Studies

A diagram of all eligible studies that presented data on ORR based on indication and source of evidence is presented in ESM Fig. 1. Pooled analyses of ORR in both RCTs and RWE studies was possible for three indications, namely programmed death-1/programmed death-ligand 1 (PD-1/PD-L1) as first line in NSCLC, PD-1/PD-L1 as second or later line in NSCLC, and PD-1/PD-L1 as second or later line in melanoma.

As first line in NSCLC, the pooled ORR was 40.5% (95% CI 32.2–49.3) in RCTs and 48.7% (95% CI 39.5–58.0) in RWE studies. In patients treated with PD-1/PD-L1 as second or later line in NSCLC, the pooled ORR in RCTs and RWE studies was 31.8% (95% CI 26.6–37.5) and 34.9% (95% CI 26.6–44.3), respectively. In melanoma patients treated with PD-1/PD-L1 as second or later line, the pooled ORR was 16.3% (95% CI 12.8–20.5) in RCTs and 23.1% (95% CI 20.6–25.9) in melanoma patients. The latter difference in pooled ORR between RCTs and RWE studies was statistically significant in favor of RWE studies.

3.3 Progression-Free Survival in RCTs and RWE Studies

A diagram of all eligible studies that presented data on PFS based on indication and source of evidence is presented in ESM Fig. 2.

Figure 2 illustrates the pooled modified PFS (mPFS) between RCTs and RWE studies in different treatment strategies and indications. Adequate data for comparisons were available for the same three indications as in ORR analysis.

Pooled mPFS was similar in RCTs and RWE studies including PD-1/PD-L1 as first line in NSCLC (mPFS_{RCT} 7.96 [95% CI 6.67–9.52] vs. mPFS_{RWE} 7.79 [95% CI 5.59–10.87]), PD-1/PD-L1 as second or later line in NSCLC (mPFS_{RCT} 3.00 [95% CI 2.27–3.98] vs. mPFS_{RWE} 3.93 [95% CI 3.14–4.92]), and PD-1/PD-L1 as second or later line in melanoma (mPFS_{RCT} 4.77 [95% CI 2.93–7.76] vs. mPFS_{RWE} 3.93 [95% CI 2.05–7.53]). These contrasts were not statistically significant ($p > 0.05$) [ESM Fig. 3].

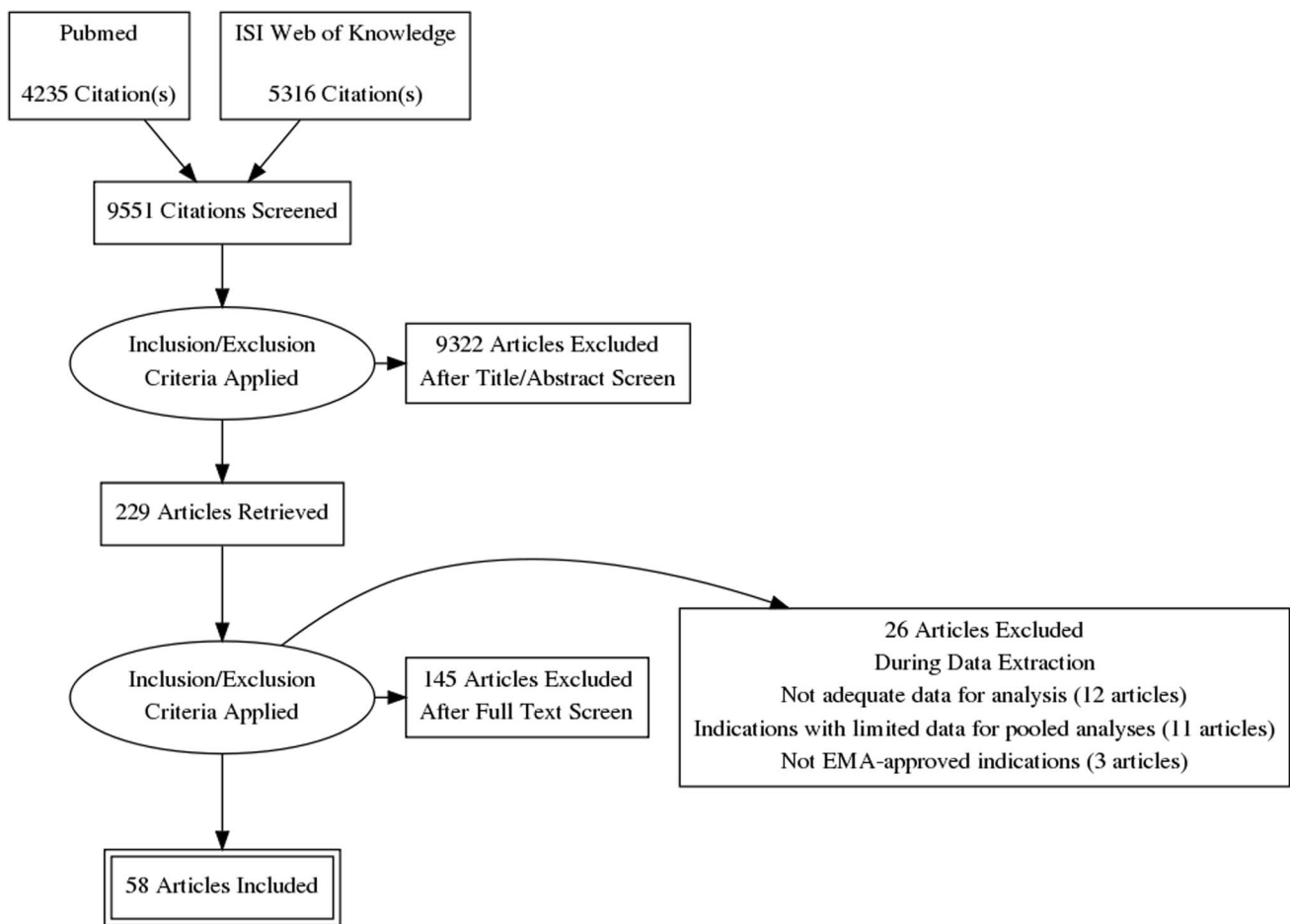


Fig. 1 Study selection process. EMA European Medicines Agency

3.4 Overall Survival in RWE and RCTs

A diagram of eligible studies that presented data on OS is shown in ESM Fig. 4.

Adequate data for comparisons in terms of median OS (mOS) were retrieved for four indications that were the same as the pooled analyses for mPFS with the addition of ipilimumab in melanoma patients (Fig. 2).

Pooled mOS was similar in RCTs and RWE studies in the four indications with adequate data. PD-1/PD-L1 as first line in NSCLC (mOS_{RCT} 21.22 [95% CI 17.76–25.36] vs. mOS_{RWE} 18.52 [95% CI 12.98–21.02]), PD-1/PD-L1 as second or later line in NSCLC (mOS_{RCT} 11.55 [95% CI 10.02–13.31] vs. mOS_{RWE} 10.54 [95% CI 9.62–11.56]), PD-1/PD-L1 as second or later line in melanoma (mOS_{RCT} 22.35 [95% CI 13.63–36.85] vs. mOS_{RWE} 17.33 [95% CI 6.14–28.91]), and ipilimumab in melanoma patients (mOS_{RCT} 13.93 [95% CI 10.36–18.74] vs. mOS_{RWE} 12.18 [95% CI 6.52–22.75]). These contrasts were not statistically significant ($p > 0.05$) [ESM Fig. 3].

3.5 High-Grade Immune-Related Toxicity in RCTs and RWE Studies

A summary of the pooled rates for high-grade immune-related toxicities in RCTs and RWE studies in different clinical indications is presented in Table 1.

ESM Fig. 5 presents a list of eligible studies with information about RG 3–4 Tox. A comparison between pooled rates in RCTs and RWE studies was only feasible in two indications, namely second line of PD-1/PD-L1 in NSCLC and second line of PD-1/PD-L1 in melanoma. In patients with NSCLC and melanoma, the contrasts of pooled toxicity rates showed a statistically significant difference between RCTs and RWE studies, with lower toxicity rates in RWE studies (ESM Fig. 6).

3.6 Rate of Treatment Discontinuation Due to Toxicity in RWE Studies and RCTs

ESM Fig. 7 presents a list of eligible studies with information about RODI. Among these studies, adequate data for

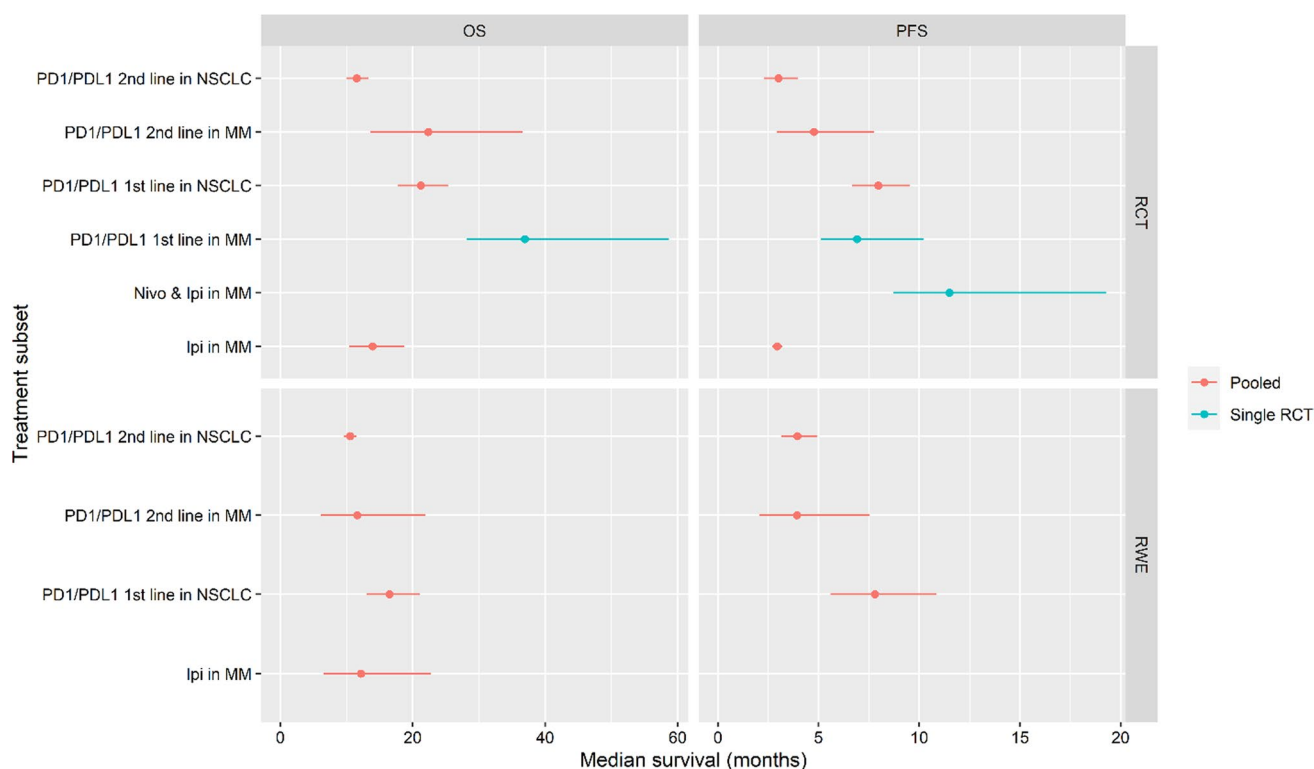


Fig. 2 Pooled median survivals (in months) and corresponding 95% confidence intervals for PFS and OS between RCTs and RWE studies for different immunotherapy indications. *NSCLC* non-small cell lung cancer, *MM* malignant melanoma, *Ipi* ipilimumab, *Nivo* nivolumab,

PFS progression-free survival, *OS* overall survival, *RCT* randomized controlled trial, *RWE* real-world evidence, *PDI* programmed death-1, *PDL1* programmed death-ligand 1

comparison between RCTs and RWE studies in terms of treatment discontinuation due to toxicity were found in only one indication (second-line PD-1/PD-L1 in NSCLC) where no statistically significant difference was observed (pooled $RODI_{RCT}$ 6.7% [95% CI 4.9–9.2] vs. pooled $RODI_{RWE}$ 9.5% [95% CI 7.8–11.4]) [Table 1 and ESM Fig. 6].

4 Discussion

In our meta-analysis, aimed at comparing the efficacy and toxicity of checkpoint inhibitors in different indications, as observed in RCTs compared with RWE studies, we did not find any statistically significant or clinically relevant differences in terms of PFS or OS within the same indication. In some indications, a higher rate of ORR and lower rate of toxicity in favor of RWE was observed that mainly reflects the inherent difficulties in evaluating these outcomes in RWE studies.

RWE studies have an established role as part of postmarket effectiveness and safety monitoring of new therapies in a real-world setting. Recently, RWE seems to have gained a more central role as part of evidence to support regulatory decision making. In fact, the number of applications

including RWE as supporting evidence to both the US FDA and the EMA has increased [11, 12]. At the same time, the number of approvals where RWE influenced regulatory decision making and was included in product labels has also increased [12, 13], highlighting the emerging role of RWE in the regulatory decision-making process. The potential role of RWE on cost-effectiveness analyses has also been highlighted as a valuable aspect that could impact health technology assessment decisions [9, 14].

Considering the emerging role of RWE in regulatory decision making, we sought to investigate whether there is a clinically relevant difference in the expected, according to RCTs, efficacy and toxicity of immunotherapy in cancer patients compared with evidence derived from studies in a real-world setting. Our hypothesis was that we would not observe any clinically meaningful difference in outcomes between the two sources of evidence in indications where immunotherapy has shown a substantial benefit, as in melanoma, NSCLC, and renal cell carcinoma, whereas there might be a difference in indications with modest benefit, as in squamous cell carcinoma of the head and neck or in advanced urothelial carcinoma [15]. Given the lack of RWE in several indications, we were only able to test the first part of our hypothesis, namely whether there is a difference

in outcomes between RCTs and RWE studies in melanoma and NSCLC indications where immunotherapy has revolutionized the treatment strategy. The results of the pooled analyses confirm our hypothesis and further strengthen the robustness of randomized evidence on the benefit of immunotherapy in these indications that can be directly implemented into clinical practice, thus minimizing the efficacy-effectiveness gap [16].

The observed higher rate of ORR in RWE studies compared with RCTs in some indications can be explained by the well-documented problematic application of RECIST criteria in a real-world setting compared with the objectivity that can be achieved on response outcomes by using RECIST in RCTs [17, 18]. Recently, a framework for evaluating tumor responses in a real-world setting has been developed and tested in patients with NSCLC, showing a high correlation with response rates from RCTs [18]. However, the RWE studies included in the current meta-analysis did not use similar frameworks to evaluate response rates, thus increasing the risk for overestimation of response rates based on unstandardized medical records.

Among all immunotherapy indications where a comparison between RCTs and RWE studies was feasible, a somewhat lower toxicity rate in RWE studies was observed. Using RWE to describe toxicity to specific treatment strategies has several advantages, as the possibility of capturing long-term or rare toxicities and to investigate tolerability in an unselected population that better reflects clinical practice than patient cohorts in RCTs does [19]. On the other hand, capturing toxicity through medical records is challenging, with risk for lower accuracy compared with the close monitoring and prospective collection within RCTs [19]. The latter challenge can be overcome with prospective collection of toxicity data in a real-world setting, but this was not the case in any of the eligible studies in the current meta-analysis. Although the lower toxicity rates in RWE studies in our pooled analyses can be explained by the challenges in capturing toxicity events in medical records, our findings are reassuring in relation to the safety profile of immunotherapy in clinical practice.

Could these data oppose that RWE might not be valuable in indications where a treatment strategy offers substantial benefit according to randomized evidence? We argue that RWE studies still offer valuable information in several aspects of immunotherapy in clinical practice. In fact, the use of immunotherapy in patients with pre-existing rheumatic disease [20, 21], the possibility of immunotherapy rechallenge after discontinuation due to high-grade toxicity [22, 23], and the risk for rare immune-related toxicities such as myocarditis [24] are some of the aspects where the current evidence relies only on RWE.

Our study has several limitations that should be discussed and considered when interpreting the results. First,

we restricted the immunotherapy indications to those with EMA approval at the time of the last search. Considering the expanding range of cancer types and settings where immunotherapy gains an important role as a treatment strategy, several new indications have not been captured in the present meta-analysis. Furthermore, few RWE studies were found in some indications, thus limiting the possibility of comparing the evidence from RCTs and RWE studies in all EMA-approved indications. As a result, our findings are restricted only to patients with melanoma and NSCLC, namely indications where immunotherapy has shown a substantial benefit. Another limitation of the present study is that we restricted our analyses to indications where immunotherapy is given as monotherapy or as combination immunotherapy (although we did not find enough data to perform pooled analyses) and excluded indications where immunotherapy is combined with chemotherapy or targeted therapies. Finally, the clinical heterogeneity of eligible studies in terms of patient-related and tumor characteristics could negatively influence the reliability of pooled analyses.

5 Conclusion

In summary, our comparison of randomized and real-world evidence in terms of efficacy-effectiveness and toxicity of immunotherapy in melanoma and NSCLC revealed no clinically significant difference between the two sources of evidence. The results are reassuring that the clinical value of immunotherapy, according to randomized evidence, in patients with melanoma or NSCLC is evident in real-world setting as well. Our results give insights on the expected results from RWE in treatment strategies where a substantial benefit has been shown on randomized evidence. A comparison between RCTs and RWE might be of importance in situations where the clinical benefit of a treatment strategy is modest and future studies should be focused on such indications. In any case, RWE is crucial for specific clinical scenarios when a new treatment strategy is implemented to clinical practice.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11523-022-00901-1>.

Declarations

Funding Open access funding provided by Örebro University.

Conflict of interest Evangelos Digkas, Anthony Jagri Tabiim, Daniel Smith, and Antonis Valachis declare they have no conflicts of interest that might be relevant to the contents of this manuscript.

Ethics approval Not applicable for a systematic review and meta-analysis.

Consent to participate Not applicable for a systematic review and meta-analysis

Consent for publication Not applicable for a systematic review and meta-analysis

Availability of data, material, or code The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author contributions All authors contributed to the study conception and design. Evangelos Digkas and Anthony Jagri Tabiim performed the literature search and data extraction, and Antonis Valachis resolved any discrepancies during the searching and data extraction phases. Daniel Smith performed the data analysis. The first draft of the manuscript was written by Evangelos Digkas and Anthony Jagri Tabiim and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

References

- Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med*. 2010;363:711–23. <https://doi.org/10.1056/NEJMoa1003466>.
- Vaddepally RK, Kharel P, Pandey R, Garje R, Chandra AB. Review of indications of FDA-approved immune checkpoint inhibitors per NCCN guidelines with the level of evidence. *Cancers (Basel)*. 2020;12:738. <https://doi.org/10.3390/cancers12030738>.
- Haslam A, Prasad V. Estimation of the percentage of US patients with cancer who are eligible for and respond to checkpoint inhibitor immunotherapy drugs. *JAMA Netw Open*. 2019;2: e192535. <https://doi.org/10.1001/jamanetworkopen.2019.2535>.
- Larkin J, Chiarion-Sileni V, Gonzalez R, Grob JJ, Rutkowski P, Lao CD, et al. Five-year survival with combined nivolumab and ipilimumab in advanced melanoma. *N Engl J Med*. 2019;381:1535–46. <https://doi.org/10.1056/NEJMoa1910836>.
- Ferris RL, Blumenschein G Jr, Fayette J, Guigay J, Colevas AD, Licitra L, et al. Nivolumab for recurrent squamous-cell carcinoma of the head and neck. *N Engl J Med*. 2016;375:1856–67. <https://doi.org/10.1056/NEJMoa1602252>.
- Bellmunt J, de Wit R, Vaughn DJ, Fradet Y, Lee JL, Fong L, KEY-NOTE-045 Investigators, et al. Pembrolizumab as second-line therapy for advanced urothelial carcinoma. *N Engl J Med*. 2017;376:1015–26. <https://doi.org/10.1056/NEJMoa1613683>.
- Yoo SH, Keam B, Kim M, Kim TM, Kim DW, Heo DS. Generalization and representativeness of phase III immune checkpoint blockade trials in non-small cell lung cancer. *Thorac Cancer*. 2018;9:736–44. <https://doi.org/10.1111/1759-7714.12641>.
- Sorbye H, Pfeiffer P, Cavalli-Björkman N, Qvortrup C, Holsen MH, Wentzel-Larsen T, et al. Clinical trial enrollment, patient characteristics, and survival differences in prospectively registered metastatic colorectal cancer patients. *Cancer*. 2009;115:4679–87. <https://doi.org/10.1002/cncr.24527>.
- Donia M, Kimper-Karl ML, Høyer KL, Bastholt L, Schmidt H, Svane IM. The majority of patients with metastatic melanoma are not represented in pivotal phase III immunotherapy trials. *Eur J Cancer*. 2017;74:89–95. <https://doi.org/10.1016/j.ejca.2016.12.017>.
- Bowrin K, Briere JB, Levy P, Toumi M, Millier A. Use of real-world evidence in meta-analyses and cost-effectiveness models. *J Med Econ*. 2020;23:1053–60. <https://doi.org/10.1080/13696998.2020.1792917>.
- Purpura CA, Garry EM, Honig N, Case A, Rassen JA. The role of real-world evidence in FDA-approved new drug and biologics license applications. *Clin Pharmacol Ther*. 2022;111:135–44. <https://doi.org/10.1002/cpt.2474>.
- Flynn R, Plueschke K, Quinten C, Strassmann V, Duijnhoven RG, Gordillo-Marañón M, et al. Marketing authorization applications made to the European Medicines Agency in 2018–2019: what was the contribution of real-world evidence? *Clin Pharmacol Ther*. 2022;111:90–7. <https://doi.org/10.1002/cpt.2461>.
- Arondekar B, Duh MS, Bhak RH, DerSarkissian M, Huynh L, Wang K, et al. Real-world evidence in support of oncology product registration: a systematic review of new drug application and biologics license application approvals from 2015–2020. *Clin Cancer Res*. 2022;28:27–35. <https://doi.org/10.1158/1078-0432.CCR-21-2639>.
- Bowrin K, Briere JB, Levy P, Millier A, Clay E, Toumi M. Cost-effectiveness analyses using real-world data: an overview of the literature. *J Med Econ*. 2019;22:545–53. <https://doi.org/10.1080/13696998.2019.1588737>.
- Sun L, Zhang L, Yu J, Zhang Y, Pang X, Ma C, et al. Clinical efficacy and safety of anti-PD-1/PD-L1 inhibitors for the treatment of advanced or metastatic cancer: a systematic review and meta-analysis. *Sci Rep*. 2020;10:2083. <https://doi.org/10.1038/s41598-020-58674-4>.
- Eichler HG, Abadie E, Breckenridge A, Flamion B, Gustafsson LL, Leufkens H, et al. Bridging the efficacy-effectiveness gap: a regulator's perspective on addressing variability of drug response. *Nat Rev Drug Discov*. 2011;10:495–506. <https://doi.org/10.1038/nrd3501>.
- Griffith SD, Tucker M, Bowser B, Calkins G, Chang CJ, Guardino E, et al. Generating real-world tumor burden endpoints from electronic health record data: comparison of RECIST, radiology-anchored, and clinician-anchored approaches for abstracting real-world progression in non-small cell lung cancer. *Adv Ther*. 2019;36:2122–36. <https://doi.org/10.1007/s12325-019-00970-1>.
- Ma X, Bellomo L, Magee K, Bennette CS, Tymejczyk O, Samant M, et al. Characterization of a real-world response variable and comparison with RECIST-based response rates from clinical trials in advanced NSCLC. *Adv Ther*. 2021;38:1843–59. <https://doi.org/10.1007/s12325-021-01659-0>.
- Di Maio M, Perrone F, Conte P. Real-world evidence in oncology: opportunities and limitations. *Oncologist*. 2020;25:e746–52. <https://doi.org/10.1634/theoncologist.2019-0647>.
- Tison A, Quéré G, Misery L, Funck-Brentano E, Danlos FX, Routier E, et al. Groupe de Cancérologie Cutanée, Groupe Français de Pneumo-Cancérologie, and Club Rhumatismes et Inflammations.

- Safety and efficacy of immune checkpoint inhibitors in patients with cancer and preexisting autoimmune disease: a nationwide, multicenter cohort study. *Arthritis Rheumatol.* 2019;71:2100–11. <https://doi.org/10.1002/art.41068>.
21. Yamaguchi A, Saito Y, Okamoto K, Narumi K, Furugen A, Takekuma Y, et al. Preexisting autoimmune disease is a risk factor for immune-related adverse events: a meta-analysis. *Support Care Cancer.* 2021;29:7747–53. <https://doi.org/10.1007/s00520-021-06359-7>.
 22. Dolladille C, Ederhy S, Sassier M, Cautela J, Thuny F, Cohen AA, et al. Immune checkpoint inhibitor rechallenge after immune-related adverse events in patients with cancer. *JAMA Oncol.* 2020;6:865–71. <https://doi.org/10.1001/jamaoncol.2020.0726>.
 23. Allouchery M, Lombard T, Martin M, Rouby F, Sassier M, Bertin C, French Network of Regional Pharmacovigilance Centers, et al. Safety of immune checkpoint inhibitor rechallenge after discontinuation for grade ≥ 2 immune-related adverse events in patients with cancer. *J Immunother Cancer.* 2020;8: e001622. <https://doi.org/10.1136/jitc-2020-001622>.
 24. Makunts T, Saunders IM, Cohen IV, Li M, Moumedjian T, Issa MA, et al. Myocarditis occurrence with cancer immunotherapy across indications in clinical trial and post-marketing data. *Sci Rep.* 2021;11:17324. <https://doi.org/10.1038/s41598-021-96467-5>.