



HAL
open science

Adverse events reporting in phase 3 oncology clinical trials of checkpoint inhibitors: A systematic review

Aline Barhli, Marie-Liesse Joulia, Christophe Tournigand, Emmanuelle Kempf

► To cite this version:

Aline Barhli, Marie-Liesse Joulia, Christophe Tournigand, Emmanuelle Kempf. Adverse events reporting in phase 3 oncology clinical trials of checkpoint inhibitors: A systematic review. *Critical Reviews in Oncology/Hematology*, 2021, 157, pp.103162 -. 10.1016/j.critrevonc.2020.103162 . hal-03560258

HAL Id: hal-03560258

<https://hal.science/hal-03560258>

Submitted on 15 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1. AP-HP, Hôpital Henri Mondor, Service d'oncologie médicale, , 52 avenue du Maréchal de Lattre de Tassigny, 94000 Créteil, France.

2. Univ Paris Est Créteil, INSERM, IMRB, F-94010 Créteil, France

E-mails of the authors : aline.barhli@hotmail.com; joulia.marie@gmail.com; christophe.tournigand@aphp.fr; emmanuelle.kempf@aphp.fr

Corresponding author: Prof Christophe Tournigand, MD, PhD, Department of Medical Oncology, Henri Mondor Teaching Hospital, Assistance Publique – Hôpitaux de Paris, 52 avenue du Maréchal de Lattre de Tassigny, 94000 Créteil, France.

Tel: +33 1 49 81 25 75, E-mail: christophe.tournigand@aphp.fr, Fax: +33 1 49 81 25 76.

Department of Medical Oncology

Henri Mondor Teaching Hospital, Assistance Publique – Hôpitaux de Paris

52 avenue du Maréchal de Lattre de Tassigny

94000 Créteil

France.

Total word count: 2792

1 figure / 4 tables

Aline Barhli is a resident in medical oncology of Assistance Publique-Hôpitaux de Paris, France. She is a member of the American Society of Clinical Oncology and the European Society of Medical Oncology.

Marie-Liesse Joulia is a resident in medical oncology of Assistance Publique-Hôpitaux de Paris, France. She is a member of the American Society of Clinical Oncology and the European Society of Medical Oncology.

Christophe Tournigand, MD, PhD, is head of the medical oncology department in Henri Mondor Hospital, Assistance Publique-Hôpitaux de Paris, France. He is Professor at the Université Paris-Est Créteil. He also worked as Associate Professor at Sherbrooke University and Charles LeMoyné Hospital, Quebec, Canada. He is a member of the steering committee of the GERCOR research Group (Groupe d'Etude et de Recherche Clinique en Oncologie et Radiothérapie).

Prof. Tournigand is an active member of the American Society of Clinical Oncology and the European Society of Medical Oncology. He is the co-author of more than 180 peer-reviewed publications. He is also medical reviewer for a number of journals, including *Annals of Oncology*, *European Journal of Oncology* and *Journal of Clinical Oncology*. He was a member of the editorial committee of *Journal of Clinical Oncology*. He is currently the PI of several clinical trials, from phase I to phase III. His clinical research interests are mainly in gastro-intestinal malignancies, care organisation and end-of-life.

Emmanuelle Kempf, MD, MSc, is a medical oncologist of the medical oncology department in Henri Mondor Hospital, Assistance Publique-Hôpitaux de Paris, France.

Doctor Kempf is a member of the American Society of Clinical Oncology, the European Society of Medical Oncology and the Society for Immunotherapy of Cancer. She performed her MSc' degree in Oxford, UK and a special fellowship at the European Organization for Research and Treatment of Cancer (EORTC), Brussels, Belgium. Her clinical research interests are mainly the methodology of drug development and artificial intelligence.. She is a permanent member of the Cancer Committee at the Agence Nationale de Sécurité du Médicament (ANSM), in charge of the early access programs for anticancer innovations, at the national level.

Total word count: 2792

1 figure / 4 tables

TABLE OF CONTENTS

1. INTRODUCTION

2. METHODS

1.1. Studies selection

1.2. Data extraction and analysis

3. RESULTS

1.1. Trial selection

1.2. Adherence of toxicity reporting to the TRIO recommendations

1.3. Adherence of toxicity reporting to the CONSORT recommendations

1.4. Other clinically-relevant ICIs toxicity reporting items

4. DISCUSSION

ABSTRACT

Introduction. This study aimed at exploring adverse events (AEs) reporting in cancer trials involving immune checkpoint inhibitors (ICIs).

Method. A systematic review on how ICIs phase 3 trials follow TRIO and 2004 CONSORT harms extension recommendations referring to toxicity was performed by two independent reviewers.

Results. Among 46 trials included, 74% did not present separately grade 3 and grade 4 AEs. Timing of onset and duration were reported in 30% and 28%, respectively. AEs occurring in <10% of patients was only reported in 35% of studies. Patient-related outcomes (PROs) were analyzed in only 17% of reports. Eight articles qualified the toxicity profile as “manageable”, “tolerable”, “well tolerated” or “favorable” despite reporting a rate of grade 3-4 greater than 33%.

Conclusion. Reporting toxicity results is crucial. However, toxicity reporting is highly incomplete in clinical trials. Guidelines, new metrics and incorporation of PROs are needed for a comprehensive knowledge of toxicity profile.

KEY WORDS quality reporting, immune toxicity, adverse event, immune checkpoint inhibitors, systematic review

LIST OF ABBREVIATIONS

AEs, adverse events; AEL, adverse effect load; CTLA-4, Cytotoxic T-lymphocyte antigen-4 monoclonal; CONSORT, Consolidated Standards of Reporting Trials; CTCAE, Common Terminology Criteria for Adverse Events; ICIs, immune checkpoint inhibitors; IQR, interquartile range; irAEs, immune-related adverse events; PD-1, programmed death-1 receptor ; PD-L1, programmed death-1 ligand; PROs, patient-reported outcomes; QoL, quality of life; RCTs, randomized clinical trials; ToxT, toxicity over time; TRIO, Trial Reporting in Immuno-Oncology

1. INTRODUCTION

Over the past decade, immune checkpoint inhibitors (ICIs) have greatly improved clinical outcomes of cancer patients[1]. Cytotoxic T-lymphocyte antigen-4 monoclonal (CTLA-4), programmed death-1 receptor (PD-1), and programmed death-1 ligand (PD-L1) antibodies used alone or in combination have been approved by the US Food and Drug Administration and the European Medicines Agency in various tumor types[2-6]. Treatment-induced immune-related adverse events (irAEs) have been observed in patients treated with ICIs and their presentation is heterogeneous and unpredictable[7-10].

The major changes in daily administration of cancer treatments have been based on the published results of prospective randomized clinical trials (RCTs) and thus, on adverse events (AEs) reporting. The safety results are critical for a risk-benefit assessment that is necessary for clinicians and regulators to evaluate the clinical relevance of new treatments. Therefore, standards such as the Consolidated Standards of Reporting Trials (CONSORT) statement and the Trial Reporting in Immuno-Oncology (TRIO) have been adopted in order to increase the quality of harm-reporting in RCTs. In May 2003, the CONSORT members generated an extension of 10 new recommendations upon general harms reporting[11]. **TRIO standards were developed by a working group from American Society of Clinical Oncology (ASCO) and the Society for Immunotherapy of Cancer (SITC) of practicing medical oncologists, immunologists, clinical researchers, biostatisticians, and representatives of industry and government. In 2018, they published 12 specific reporting recommendations to improve the interpretation and comparison of efficacy and toxicity end points and the combination and sequencing of treatments in ICIs clinical trials[12]. These recommendations are based on expert consensus and will likely need regular revision, as a better understanding of ICIs will be acquired with the on-going development of this field. However, harm reporting remains challenging to ensure the safety of immunotherapy[13].**

The aim of this systematic review was to assess the quality of AEs reporting in phase 3 cancer RCTs of ICIs and how validated reporting guidelines might be improved.

2. METHODS

1.1. Studies selection

The databases of PubMed, EMBASE, and Cochrane Library were searched by using the following terms: “immunotherapy”, “nivolumab”, “ipilimumab”, “pembrolizumab”, “atezolizumab”, “tremelimumab”, “avelumab”, “durvalumab”, “PD-1”, “PD-L1”, “CTLA-4”, “immune therapy”, “immune checkpoint”, “immune checkpoint inhibitor”, “immune checkpoint blockade”, “combination immunotherapy”, “carcinoma”, “sarcoma”, “neoplasm”, “tumor”, “cancer”, “malignancy”, “randomized clinical trials”, which were combined using “or” or “and” for relevant studies published on or before September 2019. Two reviewers independently screened articles for eligibility and extracted data from the included studies. All published phase 3 RCTs including patients with various tumors with at least one arm containing an ICI were included. Exclusion criteria were as follows: reviews, abstracts, ongoing clinical trials, studies lacking necessary data, phase 2 trials, subgroup studies, studies not assessing ICIs, and secondary reports on previously published trials.

1.2. Data extraction and analysis

The TRIO and 2004 CONSORT harms extension items referring to the toxicity reporting were identified[11, 12]. Derived from these recommendations, we adapted objective criteria to assess variations in adherence to the TRIO and CONSORT guidelines and the quality of toxicity reporting (Supplement 1). The TRIO Items 10 and 11 were subdivided as proposed by the TRIO statement. Spin, defined as a « specific intentional or unintentional reporting that fails to faithfully reflect the nature and range of findings and that could affect the impression of the results produce in readers », was assessed in the harm qualification profile in discussion[13].

Descriptive statistics were used. Univariate or multivariate analyses were not performed.

3. RESULTS

1.1. Trial selection

The search provided a total of 171 publications (Figure 1). After excluding duplicates, 163 remained. Of these, 74 were discarded because these presented abstracts, letters to the editor, posters, and ongoing studies and thus 89 studies were assessed for eligibility. By full-text assessing, 42 reports were further excluded as these

presented phase II trials, abstracts, reports of subgroup analyses or actualization of previously published trials, and trials not assessing ICIs. As a result, 47 reports of phase 3 trials (with one full-text article on toxicity update of the original source) were included in the analysis. For two and three articles supplementary appendix and the protocol, respectively were not available. A total of 30,615 patients were randomized in all included trials. The trials characteristics were as follows: tumor site was lung and melanoma in 16 (35%) and 12 (26%) of the cases, respectively; the median sample size was 666 (interquartile (IQR) range, 437-880), the median duration of follow-up was 12.5 months (IQR, 8.9-17.8). All studies were funded by the industry. The detailed description of the included studies is presented in Table 1.

Table 1. Characteristics of included studies

Study characteristics		Total (n = 46) n (%)
Tumor site	Lung	16 (35)
	Melanoma	12 (26)
	Urinary tract	9 (19)
	Digestive	4 (9)
	Other (head and neck squamous cell cancer, multiple myeloma, breast cancer)	5 (11)
Source of trial funding	Industry	46 (100)
	Non-industry	0 (0)
Number of arms per study	2	41 (89)
	> 2	5 (11)
Experimental arms (n = 57)	Monotherapy of immunotherapy:	34 (60)
	<i>Anti CTLA-4</i>	10
	<i>Anti PD-1/PDL-1</i>	24
	Combination of immunotherapy with:	23 (40)
	<i>Other immunotherapy agent</i>	4
	<i>Standard chemotherapy</i>	10
	<i>Targeted therapy</i>	5
<i>other</i>	4	
Cancer stage	Early stage	3 (6)
	Advanced	43 (94)
Median sample size (IQR)		666 (437-880)
Median duration of follow-up (months) (IQR)		12.5 (8.9-17.8)*

*Not assessed for 8 articles

Abbreviations: CTLA-4, cytotoxic T-lymphocyte associated protein 4; IQR, interquartile range; PD-1, programmed death-1; PDL-1, programmed death ligand-1

1.2. Adherence of toxicity reporting to the TRIO recommendations

Table 2 summarizes the TRIO standards items and how these were reported in the included studies. All articles referred to the Common Terminology Criteria for Adverse Events (CTCAE) to report the nature and severity of AEs. **Actually, beyond the non-standardized definition, terms of AE were not homogeneous through the trials. For example fatigue / asthenia, hepatitis / increased ASAT, ALAT or immune diarrhea / colitis could be used by authors to define the same immune-related adverse event.** The incidence and nature of grade 1 and grade 2 AEs were not presented separately by any of the articles and of grade 3 and grade 4 were presented jointly by 74% of them. **Grade 3, 4 and 5 were presented grouped in 19 publications (41%), even though it pools a wide range of severity. More particularly, grade 4 AEs are defined as life-threatening toxicities and require urgent intervention whereas grade 3 gathers severe or medically significant but not immediately life threatening AEs. In regards to grade 4 AEs, 13 trials (28%) presented them solely but no article had more detailed information for grade 4 events nor management.** Grade 5 AEs were reported by all studies included and their nature was specified in 100% of them. Only two articles detailed patient characteristics and the events for all the treatment-related deaths in supplement. **However information about grade 5 AEs management was provided by only one publication.** The timing of onset, duration, and resolution were reported in 30%, 28%, and 26% of the studies, respectively. When mentioned, these metrics could be reported solely for one or a few AEs and these data were not available for all harms. The use of high dose steroids to manage AEs was reported in 24% of the analyzed articles.

Table 2. The quality of immune-related adverse events reporting according to the TRIO guideline

TRIO recommended items	Criteria used to collect the data	Articles reporting the criteria	
		(n = 46)	Specified solely in supplementary data n = %
7*. “Differentiate between the clinical diagnoses of IO toxicity and the specific symptoms that led to the diagnoses”	IrAEs and non-specific adverse events, are:		
	- presented separately	39 (85)	22 (48)
	- presented together	1 (2)	
	- irAEs are not presented	6 (13)	
8. “If the prespecified clinical diagnoses used in data collection belong to categories such as ‘irAEs’ or ‘adverse events of special interest’, report how these terms are defined and why these categories were selected for trial reporting”	A definition of irAEs is provided	27 (59)	
9. “Report all toxicity by specific grade”	Grade 1	0 (0)	
	Grade 2	0 (0)	
	pooled Grade 1-2	12 (26)	
	Grade 3	12 (26)	
	Grade 4	13 (28)	
	pooled Grade 3-4	29 (63)	
	Grade >=3 (3, 4 and 5 pooled)	19 (41)	
	Grade 5	46 (100)	
10. “Report clinical interventions used to manage IO toxicity”	“Dose Discontinuation”	44 (96)	
	“Use of High-Dose Steroids”	11 (24)	3 (7)
	“Duration of High-Dose Steroid Use”	0 (0)	
	“Duration of Dose Tapering”	0 (0)	
	“Administration of Additional Immune- Suppressing Agents”	6 (13)	
	“Emergency Center Visit/ Hospitalization”	0 (0)	
11. “Report time of onset and duration of IO toxicity”	“Timing of Toxicity Onset”	14 (30)	3 (7)
	“Time to Resolution of Toxicity”	13 (28)	5 (11)
	“Percent of patients with unresolved toxicity”	12 (26)	5 (11)

*TRIO items detailed here are dedicated to toxicity reporting standards

Abbreviations: AE, adverse event; IO, immunotherapy; irAEs, immune-related adverse events; TRIO, trial reporting in immune-oncology

1.3. Adherence of the toxicity reporting to the CONSORT standards

Table 3 shows the CONSORT standards items related to the toxicity reporting and how these were reported in the selected articles. The mode of data collection was reported in 22% of the trials. The description of the plan of presenting information on harm was never reported in the method section. The time of follow-up of AEs was specified in 25 articles (54%); the median time of follow-up for all and serious AEs after the last dose of treatment was 60 (Interquartile range (IQR) 30-80) and 90 days (IQR 90-90), respectively. The duration, co-occurrence, and recurrences of AEs were presented in 26%, 0% and 4% of included trials, respectively. This information was not always exhaustive and was limited to one or a few AEs.

Table 3. The quality of immune-related adverse events reporting according to the CONSORT guideline

CONSORT recommended items	Criteria used to collect the data	Articles reporting the criteria	
		n (%)	Specified solely in supplementary data n (%)
TITLE AND ABSTRACT			
1.* “If the study collected data on harms and benefits, the title or abstract should so state”	Abstract mentions the safety profile of immunotherapy	46 (100)	
	Conclusion of abstract mentions the safety profile of immunotherapy	28 (61)	
INTRODUCTION			
2. “If addresses both harms and benefits, the introduction should so state”	Introduction mentions the safety profile of immunotherapy	26 (57)	
	Introduction provides argumentation of the safety profile	6 (13)	
METHODS			
6. “List addressed adverse events with definitions for each”	A definition of irAEs is provided	27 (59)	
	The reference to the CTCAE for the evaluation of adverse events is specified	46 (100)	
6. “Clarify how harms-related information was collected”	How often AEs are assessed is specified	15 (33)	
	The mode of data collection is specified	10 (22)	
	Follow-up for AEs reporting is specified	25 (54)	1 (2)
	Rules for treatment discontinuation and/or interruption and/or dose modification are specified	43 (93)	
12. “Describe plans for presenting and analyzing information on harms”	Methods mention the cut-off above which the incidence of AEs will be presented in results	0 (0)	
	Methods specifies how the recurrence of AEs will be recorded	0 (0)	
	Methods specifies that the timing of onset and duration of AEs will be recorded	0 (0)	
	Methods specifies the statistical tool used for the safety analysis	4 (9)	
RESULTS			
13. “Describe for each arm the participant withdrawals that are due to harms and their experiences with the allocated treatment”	The number of discontinuation of study treatment because of AEs is specified for each study arm	23 (50)	2 (4)
	The number of discontinuation of study treatment because of treatment-related AEs is specified for each study arm	29 (63)	
16. “Provide the denominators for analyses on harms”	The number of patients included in the safety analysis is specified	46 (100)	

CONSORT recommended items	Criteria used to collect the data	Articles reporting the criteria	
		n (%)	Specified solely in supplementary data n (%)
17-18-19. “Present the absolute risk per arm and per adverse event type, grade, and seriousness, and present appropriate metrics for recurrent events, continuous variables, and scale variables, whenever pertinent”	The AEs incidence is reported by grade	Detailed in Table 2	
	The duration of AEs is specified	13 (26)	5 (11)
	The co-occurrence of AEs within the same patient is specified	0 (0)	
	Recurrences of a same AEs after resolution is specified	2 (4)	
17-18-19. “Describe any subgroup analyses and exploratory analyses for harms”	The article analyses the safety profile in a specific subgroup of patients	4 (9)	
DISCUSSION			
20-21-22. “Provide a balanced discussion of benefits and harms with emphasis on study limitations, generalizability, and other sources of information on harms”	The safety profile is mentioned in discussion	42 (91)	
	The safety profile is mentioned in conclusions	26 (55)**	
	The discussion provides argumentation of the safety profile	21 (46)	

*The CONSORT items detailed here are dedicated to harms reporting standards

**Not assessed for two articles

Abbreviations: AEs, adverse events; CONSORT, Consolidated Standards of Reporting Trials ; CTCAE, Common Terminology Criteria for Adverse Events; irAEs, immune-related adverse events; NA, not assessed

1.4. Other clinically-relevant ICIs toxicity reporting items

Table 4 shows other than the TRIO nor CONSORT recommendations criteria and how these were reported. The collection of patient-reported outcomes (PROs) and/or quality of life (QoL) analyses were planned and reported **in the main publication** in 93% and 17% of the trials, respectively. **Trials assessed the quality of life with one to four different questionnaires. The European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire Core 30 (EORTC QLQ-C30) was used in 31 articles (74%) and the EQ-5D (3L) in 16 articles (38%). An organ-specific questionnaire was used in 16 articles (38%).** For 25 articles (54%), data on AEs occurring in less than 10% of the patients was not complete. The highest cut-off under which AEs were not presented was 15%. Three articles (7%) presented AEs that occurred with a difference of 5% or more between the trial arms. Four articles (9%) presented a subgroup analysis of toxicity. While age was never an exclusion criterion, none of the studies performed subgroup analysis of older patients. Some of the trials included patients with a history of auto-immune disease, however these were never examined in subgroup analysis.

Furthermore, 30 trials compared immunotherapy with chemotherapy. The use of supportive care could have reduced the occurrence of harms in the standard arm. When available, protocols mentioned supportive care prophylaxis at discretion of the investigator or according to guidelines in 18 publications (60%). Primary prophylaxis with Granocyte Colony-Stimulating Factor was forbidden in 6 studies (20%).

The presence of spin in the qualification of the harm profile in the discussion section was assessed. In 19 studies (41%) safety was qualified as “manageable”, “tolerable”, “well tolerated,” or “favorable”, while the rate of grade ≥ 3 AEs or grade 3-4 AEs was ranging from 7% to 51%. In 31 studies (67%), the toxicity profile was defined as “consistent”, “expected”, or “no new safety signal”, while the grade ≥ 3 AEs or grade 3-4 AEs rate of varied from 8% to 81%. One study (2%) used the term “substantial” and reported 54% of grade 3-4 AEs. Finally, seven studies (15%) did not give any description of the safety profile.

One article presented a post-hoc analysis of safety data of the original trial included in the analysis. All irAEs of grade 5 were detailed and additional information on definition, management, and duration of irAEs and concomitant medications was provided.

Table 4. Authors' proposition for clinically-relevant additional toxicity criteria and evaluation of their reporting quality

Additional reporting items	Criteria used to collect the data	Articles reporting the criteria	
		n (%)	Specified solely in supplementary data n (%)
PROs and/or QoL were planned and analyzed	Collection of PROs and/or QoL was planned	43* (93)	
	PROs and/or QoL were analyzed	8 (17)	1 (2)
	<i>When PROs and/or QoL were analyzed, number of studies, which specify how many patients filed in the related questionnaire</i>		<i>7**</i>
Report duration of experimental treatment		42 (91)	1 (2)
Report the number of dose delay and/or modification of experimental treatment		17 (37)	4 (9)
Cut-off for the incidence of AEs	Lowest cut-off of the incidence above which AEs are reported:		
	- <10% of patients	16 (35)	
	- ≥10% of patients	25 (54)	
	- Difference > 5% between the arms	3 (7)	
	- Not presented	2 (4)	
The presence of spin*** in qualification of the harm profile in discussion	"manageable", "tolerable", "well tolerated", "favorable" <i>The number of articles with grade ≥3 AEs or grade 3-4 AEs > 33%</i>	18 (39)	8
	"consistent", "expected", "no new safety signal" <i>The number of articles with grade ≥3 AEs or grade 3-4 AEs > 33%</i>	20 (44)	12
	"substantial" <i>The number of articles with grade ≥3 AEs or grade 3-4 AEs > 33%</i>	1 (2)	1
	No interpretation provided about the safety profile	7 (15)	

*Protocol not available for two articles

**Calculable for two articles, not available for one article

***Definition: Spin has been defined as a specific intentional or unintentional reporting that fails to faithfully reflect the nature and range of findings and that could affect the impression the results produce in readers.

Abbreviations: AE, adverse events; PROs, patient-reported outcomes; QoL, quality of life

4. DISCUSSION

Our systematic review shows that the quality of checkpoint inhibitor toxicity reporting in phase 3 cancer trials does not reach the expectation of the CONSORT and TRIO international recommendations and that harms are highly underreported.

The quality of toxicity reporting in oncology trials has already been scored as suboptimal in three systematic reviews[14-16]. Sivendran *et al.* showed that 96% of the publications reported only AEs occurring above a threshold rate of severity, 37% did not specify the criteria used to select which AEs were reported, and 88% grouped together AEs of varying severity[14]. Chen *et al.* found a complete reporting of specific characteristics of irAEs including onset, management, and reversibility in 14%, 8%, and 6% of all analyzed studies, respectively[15].

Arnaud-Coffin *et al.* performed a systematic review of AEs assessment in phase 2 and 3 ICI RCTs, based on the CONSORT recommendations and demonstrated a similar estimation of the quality of toxicity reporting as the previous studies[17]. The results of our review extend those findings, with a wider range of the toxicity data collection including the addition of the complete TRIO recommendations and a proposal for new criteria.

Our findings stand up to comparison to other studies. This could be explained, at least in part, by the fact that the authors often face the constraint of space and word limitations for editorial issues. The TRIO recommendations were developed to mature and homogenize the toxicity reporting specifically for ICIs in RCTs, however it is likely that RCT investigators are not yet aware of their content. **Thirteen trials were published after 2018, October the 19th – date of TRIO criteria publication- and there was no improvement in quality reporting according to the TRIO recommendations between trials published before and after this date.**

In the 38 trials that had a QLQ or a PRO planned and did not present the results in the main publication, 14 had a later publication dedicated to the analysis of health-related quality-of-life, even if these data are probably available at the time of the main publication. In the 24 articles that did not, some were planned to be published. However, presenting these data separately denies both authors and readers the opportunity to formulate the balance between benefits and harm; all the more so at metastatic stage, when quality of life is a critical point.

Boutron *et al.* raised awareness among the scientific community on how authors could “spin” their findings, meaning intentionally or unintentionally misreporting and misinterpreting their results, which in turn can lead to a

distortion of the interpretation by the reader[13]. **Sacks *et al.* underlined the necessity of choosing an accurate language to describe the toxicity profile and the danger of the “spin” in our understanding of treatments true impact on the lives of patients [18]. Examples are given of very different interpretations of “generally well tolerated” or “manageable toxicity” that can be reasonably made from different point of views.** The use of subjective terms such as “manageable” or “well tolerated” to qualify the toxicity in analyzed studies highlights the problem with a wide range of safety profile, underlying the subjectivity and heterogeneity of their meaning. Moreover, authors of included studies could use such terms to qualify the safety profile of a drug whereas the results showed 50% of grade 3-4 AEs. A threshold of 33% was selected to assess the overall tolerability of grade ≥ 3 AEs rates as this value is commonly used as a maximum acceptable frequency for dose-limiting toxicity in early phase clinical trials[19, 20].

The strengths of our study are the exhaustive analysis of phase 3 trials and the double-blinded review. A comprehensive review of articles, including supplementary data and related articles, is warranted.

A limit of our study was the subjectivity let by the TRIO and CONSORT guidelines to interpret some of the items. Therefore, we applied the greatest standardization possible; two medical oncologists reviewed data independently and any disagreement was resolved by two other senior physicians and methodologists.

Based on our results, we suggest some improvements in reporting harms in ICI clinical trials. We recommend that clinical trials report AEs of low incidence, especially severe AEs, irAEs, and long lasting AEs as in such a way irAEs of low incidence and of potentially severe or fatal outcome can be captured. As an example, myocarditis occurs rarely (1.14%) and have a mortality rate near to 40%[21, 22]. We also recommend reporting AEs later than a month after the last dose of treatment, especially for severe AEs, irAEs, long-lasting AEs, and recurrences. Indeed, a delayed occurrence of irAEs cannot be ruled out even after treatment discontinuation[23, 24]. Reporting of all grades of severity separately is a major need. An effort remains to be done to perform and report patients’ subjective evaluation of treatment toxicity as their role in making a more informed decision on a treatment choice have been recognized by patients and clinicians[25-29]. Unfortunately, these are still poorly reported in oncology RCTs[30]. Studies should anticipate subgroup analyses of safety profile in groups of potentially more vulnerable patients. The data on safety of ICIs in older adults with cancer are limited and mainly extracted from phase 3 trials.

Several of the CONSORT items were not or very rarely reported in the trials or were presented without any additional detail. Although expected in methods, these were often detailed in the trial protocols. We consider that incorporating in the article a significant amount of protocol data might prevent from the full article reading and flood more essential information.

New tools to report toxicity data are needed. Thanarajasingam *et al.* underlined the need of longitudinal AEs assessment to capture the evolution of toxicity over time[31]. The maximum severity of an AE is not adequate to assess long-lasting low grade AEs that will impact patient's quality of life. New metrics to present toxicity data are being evaluated to give an overview of the different aspects of tolerance, summarized in an easily readable figure. Thanarajasingam *et al.* proposed a toxicity over time (ToxT) assessment that analyses the Area Under the Curve for each Aes[32]. More recently, adverse effect load (AEL) has been tested as new metric[33]. AEL is a longitudinal score of worst grade possible excluding death over the entire treatment for an AE. Gresham *et al.* introduced a toxicity index[34] This metric accounts for the frequency and severity of harms to provide a precise description of the overall burden of multiple toxicities. Finally, a more accurate and longer toxicity reporting in phase 3 trials could provide better understanding of irAEs specificity and improved medical decision-making. Editors also need to ensure a high quality of the toxicity reporting before accepting a manuscript for publication.

Conclusions

Our study presents an exhaustive double-blinded review of phase 3 trials of ICI. We revealed the high rate of toxicity under-reporting in phase 3 cancer trials of ICI. These results should encourage the oncology community, industry, and editors to follow uniform guidelines in order to improve the exhaustivity of toxicity reporting, including the patient perspective. A comprehensive review of articles, including supplementary data and related articles, is warranted and the development of tools to present a clear overview of toxicity, including the dimension over time, is needed.

References

1. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer*. 2012;12(4):252-264. doi:10.1038/nrc3239
2. 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. doi:10.1056/NEJMoa1003466
3. Motzer RJ, Escudier B, McDermott DF, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med*. 2015;373(19):1803-1813. doi:10.1056/NEJMoa1510665
4. Cohen EEW, Soulières 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. doi:10.1016/S0140-6736(18)31999-8
5. 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. doi:10.1056/NEJMoa1716948
6. Antonia SJ, Villegas A, Daniel D, et al. Durvalumab after chemoradiotherapy in stage III non–small-cell lung cancer. *N Engl J Med*. 2017;377(20):1919-1929. doi:10.1056/NEJMoa1709937
7. Champiat S, Lambotte O, Barreau E, et al. Management of immune checkpoint blockade dysimmune toxicities: a collaborative position paper. *Ann Oncol*. 2016;27(4):559-574. doi:10.1093/annonc/mdv623
8. Weber JS, Hodi FS, Wolchok JD, et al. Safety Profile of Nivolumab Monotherapy: A Pooled Analysis of Patients With Advanced Melanoma. *J Clin Oncol*. 2017;35(7):785-792. doi:10.1200/JCO.2015.66.1389
9. Sznol M, Ferrucci PF, Hogg D, et al. Pooled Analysis Safety Profile of Nivolumab and Ipilimumab Combination Therapy in Patients With Advanced Melanoma. *J Clin Oncol*. 2017;35(34):3815-3822. doi:10.1200/JCO.2016.72.1167
10. Raschi E, Mazzarella A, Antonazzo IC, et al. Toxicities with Immune Checkpoint Inhibitors: Emerging Priorities From Disproportionality Analysis of the FDA Adverse Event Reporting System. *Target Oncol*. 2019;14(2):205-221. doi:10.1007/s11523-019-00632-w
11. Ioannidis JPA, Evans SJW, Gøtzsche PC, et al. Better reporting of harms in randomized trials: An extension of the CONSORT statement. *Ann Intern Med*. 2004;141(10):781-788. doi:10.7326/0003-4819-141-10-200411160-00009

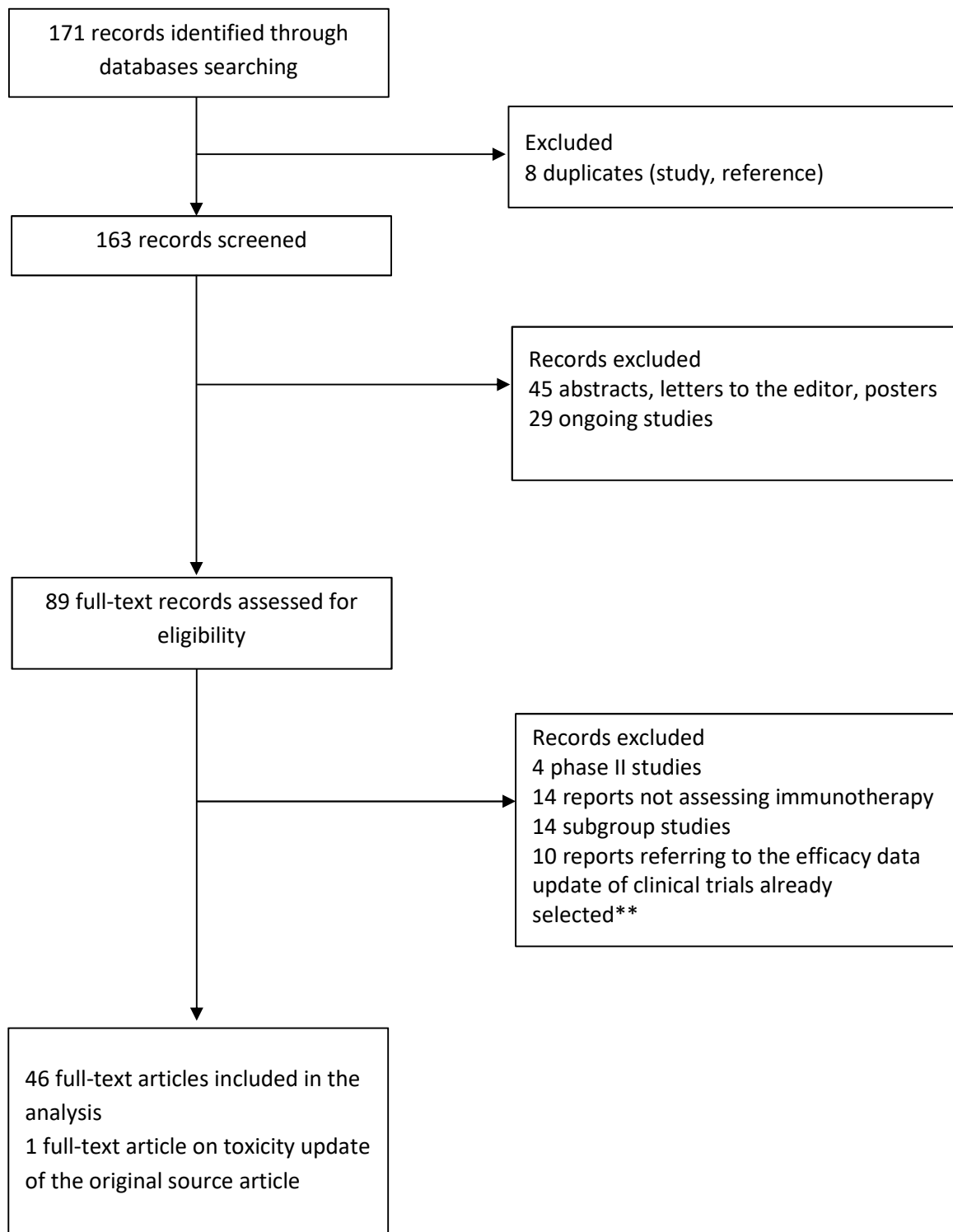
12. Tsimberidou AM, Levit LA, Schilsky RL, et al. Trial Reporting in Immuno-Oncology (TRIO): An American Society of Clinical Oncology-Society for Immunotherapy of Cancer Statement. *J Immunother Cancer*. 2018;6(1):108. doi:10.1186/s40425-018-0426-7
13. Boutron I, Ravaud P. Misrepresentation and distortion of research in biomedical literature. *Proc Natl Acad Sci U S A*. 2018;115(11):2613-2619. doi:10.1073/pnas.1710755115
14. Sivendran S, Latif A, McBride RB, et al. Adverse event reporting in cancer clinical trial publications. *J Clin Oncol*. 2014;32(2):83-89. doi:10.1200/JCO.2013.52.2219
15. Chen TW, Razak AR, Bedard PL, Siu LL, Hansen AR. A systematic review of immune-related adverse event reporting in clinical trials of immune checkpoint inhibitors. *Ann Oncol*. 2015;26(9):1824-1829. doi:10.1093/annonc/mdv182
16. Péron J, Maillet D, Gan HK, Chen EX, You B. Adherence to CONSORT adverse event reporting guidelines in randomized clinical trials evaluating systemic cancer therapy: A systematic review. *J Clin Oncol*. 2013;31(31):3957-3963. doi:10.1200/JCO.2013.49.3981
17. Arnaud-Coffin P, Maillet D, Gan HK, et al. A systematic review of adverse events in randomized trials assessing immune checkpoint inhibitors. *Int J Cancer*. 2019;145(3):639-648. doi:10.1002/ijc.32132
18. Sacks CA, Miller PW, Longo DL. Talking about toxicity — “What we’ve got here is a failure to communicate”. *N Engl J Med*. 2019;381(15):1406-1408. doi:10.1056/NEJMp1908310
19. Paoletti X, Ezzalfani M, Le Tourneau C. Statistical controversies in clinical research: Requiem for the 3 + 3 design for phase I trials. *Ann Oncol*. 2015;26(9):1808-1812. doi:10.1093/annonc/mdv266
20. Schwandt A, Harris PJ, Hunsberger S, et al. The role of age on dose-limiting toxicities in Phase I dose-escalation trials. *Clin Cancer Res*. 2014;20(18):4768-4775. doi:10.1158/1078-0432.CCR-14-0866
21. Mahmood SS, Fradley MG, Cohen J V, et al. Myocarditis in Patients Treated With Immune Checkpoint Inhibitors. *J Am Coll Cardiol*. 2018;71(16):1755-1764. doi:10.1016/j.jacc.2018.02.037.Myocarditis
22. Wang DY, Salem JE, Cohen J V., et al. Fatal Toxic Effects Associated With Immune Checkpoint Inhibitors: A Systematic Review and Meta-analysis. *JAMA Oncol*. 2018;4(12):1721-1728. doi:10.1001/jamaoncol.2018.3923
23. Dolladille C, Ederhy S, Sassier M, et al. Immune Checkpoint Inhibitor Rechallenge After Immune-Related Adverse Events in Patients With Cancer. *JAMA Oncol*. April 2020. doi:10.1001/jamaoncol.2020.0726

24. Michot JM, Bigenwald C, Champiat S, Collins M. ScienceDirect Immune-related adverse events with immune checkpoint blockade : a comprehensive review. *Eur J Cancer*. 2016;54:139-148. doi:10.1016/j.ejca.2015.11.016
25. Basch E, Iasonos A, McDonough T, et al. Patient versus clinician symptom reporting using the National Cancer Institute Common Terminology Criteria for Adverse Events: results of a questionnaire-based study. *Lancet Oncol*. 2006;7(11):903-909. doi:10.1016/S1470-2045(06)70910-X
26. Di Maio M, Gallo C, Leigh NB, et al. Symptomatic toxicities experienced during anticancer treatment: Agreement between patient and physician reporting in three randomized trials. *J Clin Oncol*. 2015;33(8):910-915. doi:10.1200/JCO.2014.57.9334
27. Basch E, Jia X, Heller G, et al. Adverse symptom event reporting by patients vs clinicians: Relationships with clinical outcomes. *J Natl Cancer Inst*. 2009;101(23):1624-1632. doi:10.1093/jnci/djp386
28. Atkinson TM, Rogak LJ, Heon N, et al. Exploring differences in adverse symptom event grading thresholds between clinicians and patients in the clinical trial setting. *J Cancer Res Clin Oncol*. 2017;143(4):735-743. doi:10.1007/s00432-016-2335-9
29. Cirillo M, Venturini M, Ciccarelli L, Coati F, Bortolami O, Verlatto G. Clinician versus nurse symptom reporting using the National Cancer Institute - Common terminology criteria for adverse events during chemotherapy: Results of a comparison based on patient's self-reported questionnaire. *Ann Oncol*. 2009;20(12):1929-1935. doi:10.1093/annonc/mdp287
30. Kyte D, Retzer A, Ahmed K, et al. Systematic Evaluation of Patient-Reported Outcome Protocol Content and Reporting in Cancer Trials. *J Natl Cancer Inst*. 2019;111(11):1170-1178. doi:10.1093/jnci/djz038
31. Thanarajasingam G, Hubbard JM, Sloan JA, Grothey A. The Imperative for a New Approach to Toxicity Analysis in Oncology Clinical Trials. *J Natl Cancer Inst*. 2015;107(10):1-4. doi:10.1093/jnci/djv216
32. Thanarajasingam G, Atherton PJ, Novotny PJ, Loprinzi CL, Sloan JA, Grothey A. Longitudinal adverse event assessment in oncology clinical trials: the Toxicity over Time (ToxT) analysis of Alliance trials NCCTG N9741 and 979254. *Lancet Oncol*. 2016;17(5):663-670. doi:10.1016/S1470-2045(16)00038-3
33. Tournigand C, Lopes G, Olsword CL, et al. How to improve toxicity evaluation in clinical trials? Testing new metrics from irinotecan or oxaliplatin-based treatments in metastatic colorectal cancer (mCRC): A pooled

analysis from 2,349 patients in ARCAD database. *J Clin Oncol.* 2020;38(4_suppl):89-89.

doi:10.1200/jco.2020.38.4_suppl.89

34. Gresham G, Diniz MA, Razaee ZS, et al. Evaluating Treatment Tolerability in Cancer Clinical Trials using the Toxicity Index. *J Natl Cancer Inst.* February 2020. doi:10.1093/jnci/djaa028



**No supplementary data on toxicity

Figure 1. Flow chart of the study selection