Patient-reported outcome instruments used in immune-checkpoint inhibitor clinical trials in oncology: a systematic review

Sara Colomer-Lahiguera[1,2*] , Denise Bryant-Lukosius[3,4,5], Sarah Rietkoetter[3,4], Lorraine Martelli[3,6], Karin Ribi[1,7], Donna Fitzpatrick-Lewis[8], Diana Sherifali[9], Angela Orcurto[2], Rosalyn Juergens[4,5] and Manuela Eicher[1,2]

Abstract

Context: Immune-checkpoint inhibitors (ICI) have shown significant benefits for overall survival across various cancer types. Patient-reported outcomes (PROs) are assessed in clinical trials as a measure of efficacy. However, it remains unclear to what extent current PRO instruments capture symptoms specific to ICI toxicities. We conducted a systematic review to identify the use and content validity of PRO instruments in ICI clinical trials in oncology.

Methods: Literature was retrieved from PubMed, Embase, PsycINFO, Medline and CINAHL databases. Articles presenting ICI clinical trials' PRO results, clinical trial study protocols, and conference abstracts stating the use of PRO measures were assessed. We evaluated the validity of identified instruments by comparing their symptom-related content with the adverse events reported in each ICI clinical trial.

Results: From database inception until January 2020, we identified 191 ICI clinical trials stating the use of PRO measures of which 26 published PRO results. The cancer-specific EORTC QLQ-C30 and the generic EQ-5D questionnaires were the most widely used instruments, often in combination with disease-specific PROs. Instruments used to report PRO symptom-related toxicities covered 45% of the most frequently reported AEs, whereas 23% of AEs were partially covered and 29% were not covered at all. Of non-covered AEs, 59% referred to the dermatologic system. Partially covered AEs related to endocrine and specific types of pain.

Conclusion: Despite the high frequency of symptom-related toxicities related to ICI, these events are only partially covered (or not addressed) by current PRO instruments, even when combined. Further research is needed to develop new strategies to tailor PRO instruments to specific ICI toxicities.

Keywords: Adverse events, Clinical trial, Immune-checkpoint inhibitors, Oncology, Patient-reported outcomes, Symptoms

Introduction

Immune-checkpoint inhibitors (ICIs) have shown significant clinical benefit by improving treatment response and survival rates for different cancer patient populations [1]. ICIs are associated with specific toxicities related to inflammatory or autoimmune responses - known as immune-related adverse events (irAEs) that affect different organs and systems [2]. These irAEs result from broad and unspecific immune over-activation [3]. In most cases, toxicities are reversible with corticosteroid or additional immunosuppressant treatment but some rare and potentially life-threatening irAEs have been reported [3]. Early recognition of irAEs and timely management is essential to reduce morbidity and mortality [4].

In clinical trials (CTs), the Common Terminology Criteria for Adverse Events (CTCAE) is the most widely

* Correspondence: sara.colomer-lahiguera@chuv.ch
[1]Institute of Higher Education and Research in Healthcare (IUFRS), Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland
[2]Department of Oncology, Lausanne University Hospital (CHUV), Lausanne, Switzerland
Full list of author information is available at the end of the article

used method to report toxicities [5]. Of these, 10% focus on symptomatic adverse events (AEs) amenable to patient self-report [6]. Thus, clinicians are often required to evaluate subjective patient experiences – resulting in potential underreporting of AEs [7, 8].

Patient-reported outcomes (PROs) are defined as "any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else" [9]. Several studies demonstrate that compared to clinician report, PROs are more concordant with patient overall health status and result in earlier detection of symptom occurrence and severity [10, 11]. Thus, patient-reported symptom assessment that incorporates items relevant to ICIs may facilitate early AE detection and may provide complementary data to support decision-making.

The overall aim of the study was to identify and categorize the types and combination of PRO instruments used in CTs involving cancer patients receiving ICI therapy (ICI-CTs), and to assess their frequency of use. To further examine the content validity of the PRO instruments, we identified and compared the most frequent AEs reported in ICI-CTs with published PRO results.

Methods

Literature search

This systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement [12] and the protocol was prospectively registered with PROSPERO (registration number: CRD42018090912). A literature search was performed in PubMed, Embase, PsycINFO, Medline and CINAHL databases using search terms adapted to the respective database. We used broad parameters employing controlled vocabulary (MeSH/Emtree preferred and related concepts) as well as specific keywords related to three constructs: (1) ICI terms (e.g. FDA-approved ICI), (2) PRO terms and (3) oncology and immunotherapy terms (Table S1 in the Supplementary Information). To ensure a comprehensive search strategy, we used the CT identification number from selected conference abstracts and CT study protocols to find additional CTs with published PRO results. The search was conducted from database inception to June 26, 2017 and updated on January 22, 2020. Articles written in English, French, Spanish, and German were included.

Selection criteria

Four pairs of reviewers (SCL, SR, LM, DBL, ME, NK, AC, AA) screened titles and abstracts following a predefined list of inclusion/exclusion criteria (Table S2). A Cohen's kappa test was performed in order to ensure consistency between reviewers (κ = 0.81). Full texts of publications and conference abstracts were obtained and reviewed independently by two pairs of reviewers (SCL, SR, KR, LM) to identify articles for final analysis. Disagreements were resolved by discussion to reach consensus.

Types of publications included

Three types of publications were included (Fig. 1). Full-text articles with PRO results served to evaluate the reporting of PRO in ICI-CTs. CT study protocols and conference abstracts were used to assess the use and combination of PRO instruments in ICI-CTs and to identify additional full-text articles reporting on PRO results in ICI-CTs.
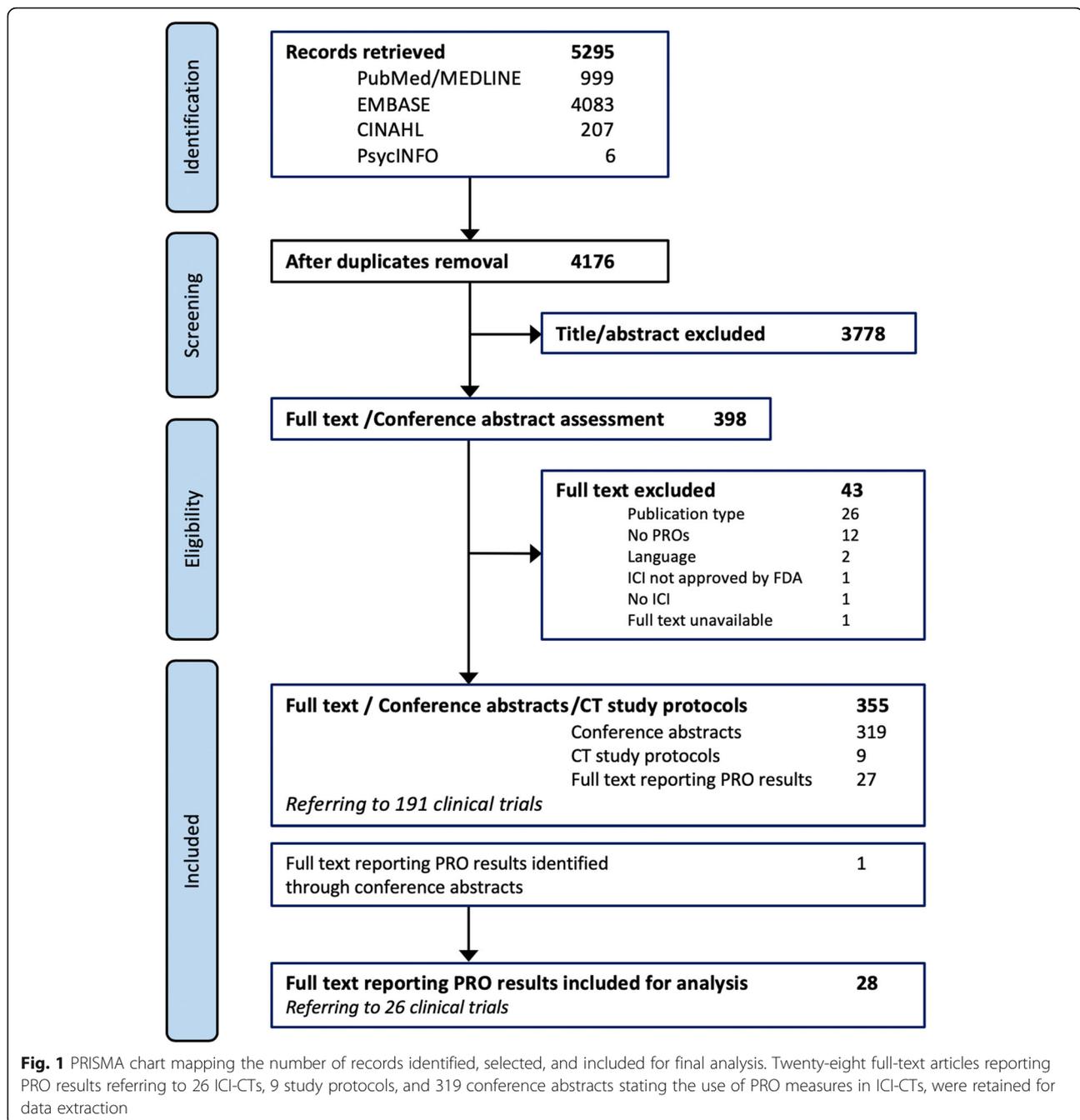
Data extraction

Reviewers (SR, LM, KR, DFL) completed a structured, online form using Distiller SR, to extract data from each full-text article including study and ICI-CT characteristics, analysis, results, study conclusions, and information about the PRO instruments used in the study (Table S3). Second reviewers (SCL and DS) verified the extracted data. For CT study protocols and conference abstracts, the following data were extracted: first author, publication year, study characteristics (name of the clinical trial/study, CT identifier, start year, cancer type, and number of participants), PRO endpoints, and name of the PRO instrument(s). Collected information was completed and verified across sources: clinicaltrials.gov/EudraCT, published CT protocol (when available), and primary CT publication (Table S4).

Categorization, combination and frequency of use of PRO instruments

Identified PRO instruments for all ICI-CTs were grouped into categories based on the focus of the measure (e.g. generic, cancer-specific, disease/tumor-specific, symptom-specific, item bank/single items, or other) and their frequency of use and combination were subsequently analyzed. The *Cochrane checklist for describing and assessing PROs in CTs* [13] was applied to the full-text articles presenting PRO results to assess measurement properties of the PRO instruments and PROs reporting (Table S5).

Comparison of PRO symptom content and most frequent reported AEs

We examined the content validity of PRO instruments used in the full-text articles reporting PRO results by comparing the extent to which the symptom-related content of identified PRO instruments (Table S7) aligned with the ten most frequent AEs reported in their respective CT study publication (Tables S8). A panel of experts including oncologists, advanced practice

Fig. 1 PRISMA chart mapping the number of records identified, selected, and included for final analysis. Twenty-eight full-text articles reporting PRO results referring to 26 ICI-CTs, 9 study protocols, and 319 conference abstracts stating the use of PRO measures in ICI-CTs, were retained for data extraction

oncology nurses, and oncology nursing researchers reviewed and, through consensus, agreed on the differentiation of complex AEs into multi-symptoms that can be reported by patients (Table S9). AEs were considered 'covered' when the item matched all corresponding symptoms contained in the PRO instrument. Conversely, items were considered 'not covered' when PRO instruments did not include the corresponding symptom. Multi-symptom AEs were considered 'partially covered' when at least two of the symptoms pre-defined (Table

S9) were contained in the PRO instrument(s). Coverage of the most frequently reported AEs (for each ICI-CT arm) were calculated and reported as a percentage.

Results

The search yielded 5295 records. After removing duplicates and screening titles and abstracts, 398 publications remained. Subsequent full-text assessment identified 27 articles reporting ICI-CTs PRO results, 319 conference abstracts, and 9 published CT study protocols. An

additional article reporting PRO results was identified through conference abstracts and was included in the final analysis (total articles $n = 28$) (Fig. 1). Detailed review of articles, CT study protocols, and conference abstracts identified 191 ICI-CTs. We found information on PRO instruments for 156 of these ICI-CTs (Figure S1 and Table S4).

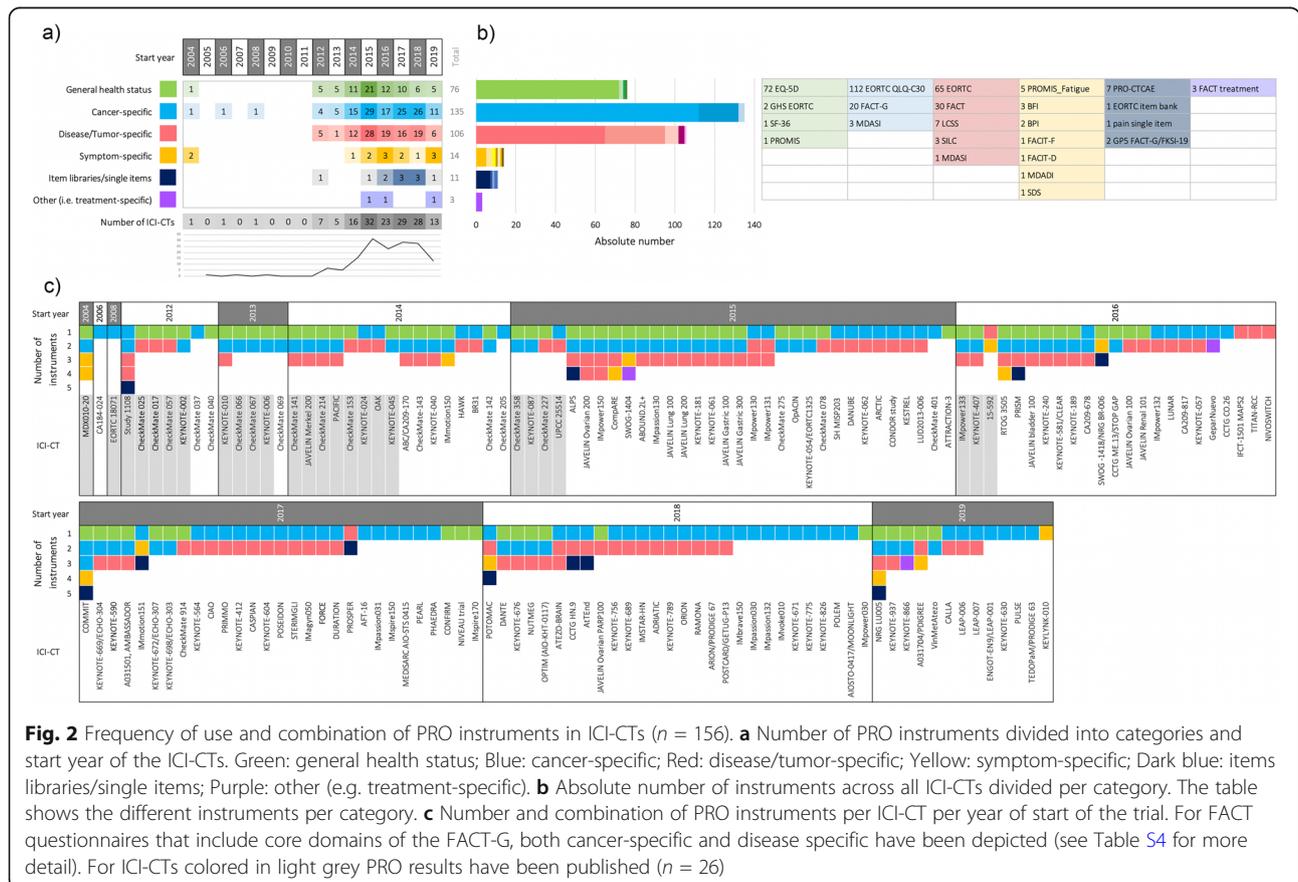### Categorization, combination and frequency of use of PRO instruments in ICI-CTs

The number, categorization, and combination of PRO instruments for each of the 156 ICI-CTs is shown in Fig. 2 and supplementary Table S4. Data review and analysis of the 156 ICI-CTs showed an increase over the years of the number of ICI-CTs stating the use or measurement of PROs in publications or conference abstracts (Fig. 2a). The peak was reached in 2015, with 32 ICI-CTs, followed by a plateau in the consecutive years. In the same period of time an increase of ICI-CTs can be observed (Figure S1).

Of the 156 ICI-CTs, 35 used one instrument, 68 used two instruments, 41 used three instruments, and 12 used four or five instruments (Fig. 2c and Table S4). Among the 68 ICI-CTs using two instruments, 16 paired generic with cancer-specific, 10 paired generic with disease/

tumor-specific, 39 paired cancer-specific and disease/tumor-specific, and the rest ($n = 3$) combined a disease/tumor or cancer-specific with a symptom-specific questionnaire ($n = 1$), a treatment-specific (n = 1), or selected items (n = 1). In studies using three instruments, combining generic, cancer-specific, and disease-specific questionnaires was the most commonly used approach ($n = 30/41$). The other 11 ICI-CTs combined generic or cancer-specific instruments with disease-specific and symptoms or single item from item-libraries. Of the 12 ICI-CTs that used four or five instruments, most of them combined generic, cancer-specific and disease/tumor-specific with either symptom questionnaires or single items selected from item libraries such as the Patient Reported Outcomes Measurement Information System (PROMIS), the PRO version of the CTCAE (PRO-CTCAE), or the European Organisation for Research and treatment of Cancer (EORTC) (Fig. 2c).

The cancer-specific EORTC Quality of Life Questionnaire Core 30 (EORTC QLQ-C30) was the most widely used PRO instrument (112/156, 72%), followed by the generic health EuroQol 5 Dimension Scale (EQ-5D) (72/156, 46%) (Fig. 2b).

Focusing on the evolution of the use of PROs over time, we observe two main periods. Between 2004

Fig. 2 Frequency of use and combination of PRO instruments in ICI-CTs ($n = 156$). **a** Number of PRO instruments divided into categories and start year of the ICI-CTs. Green: general health status; Blue: cancer-specific; Red: disease/tumor-specific; Yellow: symptom-specific; Dark blue: items libraries/single items; Purple: other (e.g. treatment-specific). **b** Absolute number of instruments across all ICI-CTs divided per category. The table shows the different instruments per category. **c** Number and combination of PRO instruments per ICI-CT per year of start of the trial. For FACT questionnaires that include core domains of the FACT-G, both cancer-specific and disease specific have been depicted (see Table S4 for more detail). For ICI-CTs colored in light grey PRO results have been published ($n = 26$)

and 2013, few studies used single instruments with the majority (8/15) of studies using two instruments. Between 2014 and 2019 ($n = 141$), we observe a higher variability in the number of instruments used (Fig. 2a).

Articles reporting on PRO results

From the 28 full-text articles reporting PRO results, one discussed associations between biomarkers and PRO results, 18 reported PRO results alone, and seven presented safety and efficacy data from the main ICI-CT together with PRO results. For two of the latter, additional full-text manuscripts on PRO results alone were published. Thus, the 28 publications referred to 26 individual ICI-CTs on lung carcinoma ($n = 11$) [14–24], melanoma ($n = 6$) [25–30], renal cell carcinoma ($n = 3$) [31–33], urothelial carcinoma ($n = 2$) [34, 35], Hodgkin's lymphoma (n = 1) [36, 37], head and neck carcinoma (n = 1) [38, 39], Merkel cell carcinoma (n = 1) [40], and cervical and vaginal carcinoma (n = 1) [41] (Table S5). With regard to measured PRO concepts, all publications reported on health-related quality of life (HRQoL) or global health status (GHS) as assessed by the cancer-specific EORTC QLQ-C30 (17/26) and/or the generic EQ-5D (15/26). Eleven publications additionally included the measure of disease-related symptoms, and another four added a function assessment. In general, the selection of the PRO instrument(s) was justified based on previous use and validation in similar patient populations - either general cancer patients or disease-specific populations (Table S5). Only one publication [40] re-validated the instrument in the study population and the psychometric results have been published elsewhere [42].

Overall, most of the publications reported differences between groups in terms of "clinically meaningful deterioration", "clinically meaningful improvement", "clinically relevant", "stable", "minimal change", "statistically significant", or "significant difference" (Table S5). The primary reasons for attrition and missing PRO data were disease progression and AEs (15/26) (Table S6).

Content validity of PRO instruments

The use of the Cochrane checklist (Table S5) revealed a lack of coverage of common ICI-related AEs. We extracted the symptom-related content of each PRO instrument used in each of the 26 ICI-CTs identified (Table S7), followed by the extraction of the AEs reported in the main CT publication (Table S8). We then charted the ten most frequent AEs (any grade) reported in the respective ICI-CTs denoting which AEs were covered by PRO instruments employed in the corresponding study (Fig. 3). In total, almost half of the AEs (142/299, 47%) were covered by PRO instrument(s) employed in the ICI-CTs, even when used in combination. Nearly a third of AEs reported in ICI-CTs were not covered by the PRO instrument(s) (86/299, 29%). Among 'not covered'

AEs, 51/86 (59%) were dermatologic in nature (e.g. rash, pruritus, vitiligo, dry skin). Further, multi-symptom AEs were partially covered in 71/299 (24%) of which 34% (24/71) related to a specific type of pain (e.g. headache, arthralgia, myalgia) and 28% (20/71) were endocrine-related (e.g. hypothyroidism, hyperthyroidism, hypophysitis).

Discussion

This systematic review examines the use of PRO instruments in clinical trials of FDA-approved ICI (up to January 2020). Our analysis reveals the cancer-specific EORTC QLQ-C30 and the generic EQ-5D questionnaires are the most widely used PRO instruments (72% and 46%, respectively) - either alone or in combination with disease-specific instruments. Although we observed an increase in the stated use of PRO measurements in ICI-CTs, reaching a peak in 2015, this should be seen in the global context of ICI development and the increase in ICI-CTs over the years. HRQoL or QoL are the most frequent measured concepts and only one third of the articles included additionally disease-specific symptom measures. Furthermore, our results indicate that the used PRO instruments did not include items specifically relating to ICI treatment-related symptoms. These results raise important concerns about the comprehensiveness of current instruments and the measurement of relevant PROs in the context of ICI-CTs. Similar conclusions have been drawn in other recently reported studies of PROs in ICI-CTs [43, 44]. However, the rigorous approach used in our review offers new, more detailed, and compelling insights about the use and content validity of PRO instruments in ICI-CTs. First, we provide a more comprehensive assessment of the aggregate content of PRO instruments in relation to frequent AEs by identifying and categorizing all the PRO instruments used in each ICI-CT (Fig. 3). In contrast, Hall et al. [44] only examined PRO instruments related to symptoms contained in the EORTC questionnaire. Secondly, our study examined the full spectrum of AEs reported in each ICI-CT (Fig. 3) as opposed to King-Kallimanis et al., who limited their examination to eight pre-selected symptoms for anti-PD-1/PD-L1 agents [43].

The use of PROs in CTs has grown in the last decade and PROs are recommended as important endpoints for trials [9, 45]. However, the availability of validated questionnaires and use of the most suitable instruments remains a challenge urging caution when interpreting HRQoL data in ICI-CTs [46–48]. A recent review of studies that led to the approval of ICIs identified several methodological problems with PRO measurement. Specifically, general HRQoL measures might lack sensitivity to provide a comprehensive assessment of the impact that ICI therapies have on patients [49]. For instance, some studies report no differences in HRQoL of ICI-treated patients

ICI-CT	Condition	Treatment arm	Most frequent AEs reported in the main ICI-CT										PRO instruments			
CheckMate 057	NSCLC	Niv_ 3 mg/kg	Fatigue	Nausea	Dec. appetite	Asthenia	Diarrhea	Peripheral edema	Myalgia	Anemia	Alopecia	Neutropenia	EQ-5D		LCSS	
KEYNOTE-407	NSCLC	Pem_200 mg + Chemo	Anemia	Alopecia	Neutropenia	Nausea	Thrombocytopenia	Diarrhea	Dec. appetite	Constipation	Fatigue	Asthenia	EQ-5D	EORTC-QLQ-C30	QLQ-LC13	
CheckMate 017	NSCLC	Niv_ 3 mg/kg	Fatigue	Dec. appetite	Asthenia	Nausea	Diarrhea	Arthralgia	Pyrexia	Pneumonitis	Rash	Muc. inflammation	EQ-5D		LCSS	
CheckMate 227	NSCLC	Niv_3 mg/kg+Ipi 1 mg/kg	Diarrhea	Rash	Fatigue	Pruritus	Dec. appetite	Asthenia	Nausea	Vomiting	Constipation	Anemia	EQ-5D		LCSS	
CheckMate 153	NSCLC	Niv_3 mg/kg	Fatigue	Diarrhea	Nausea	Hypothyroidism	Dec. appetite	Arthralgia	Rash	Pruritus	Inc. TSH		EQ-5D		LCSS	
KEYNOTE-010	NSCLC	Pem_2 mg/kg	Dec. appetite	Fatigue	Nausea	Rash	Hypothyroidism	Diarrhea	Asthenia	Pneumonitis	Stomatitis	Hyperthyroidism	EQ-5D	EORTC-QLQ-C30	QLQ-LC13	
		Pem_10 mg/kg	Dec. appetite	Fatigue	Nausea	Rash	Hypothyroidism	Diarrhea	Hyperthyroidism	Asthenia	Pneumonitis	Anemia				
PACIFIC	NSCLC	Dur_10 mg/kg	Cough	Pneumonitis	Fatigue	Dyspnea	Diarrhea	Pyrexia	Dec. appetite	Nausea	Pneumonia	Arthralgia	EQ-5D	EORTC QLQ C30	QLQ-LC13	
OAK	NSCLC	Ate_1200 mg	Fatigue	Dec. appetite	Cough	Dyspnea	Asthenia	Nausea	Pyrexia	Constipation	Diarrhea	Vomiting		EORTC-QLQ-C30	QLQ-LC13	
KEYNOTE-024	NSCLC	Pem_200 mg	Diarrhea	Fatigue	Pyrexia	Nausea	Dec. appetite	Hypothyroidism	Hyperthyroidism	Pneumonitis	Anemia	Infusion reaction		EORTC-QLQ-C30	QLQ-LC13	
UPCC 25514	NSCLC	Pem_200 mg	Pain	Fatigue[1]	Rash/skin condition	Dyspnea	Cough	Pruritus	Dizziness	Edema	Nausea	Pneumonitis		FACT-G	FACT-L	
IMpower 133	SCLC	Ate_1200 mg	Anemia	Neutropenia	Alopecia	Nausea	Fatigue	Dec. appetite	Neutropenia	Thrombocytopenia	Vomiting	Leukopenia	EQ-5D	EORTC-QLQ-C30	QLQ-LC13	
CheckMate 141	Head & Neck	Niv_2 mg/kg	Fatigue	Nausea	Rash	Dec. appetite	Pruritus	Diarrhea	Anemia	Asthenia	Vomiting	Dry skin	EQ-5D	EORTC-QLQ-C30	QLQ-H&N35	
EORTC-18071	Melanoma	Ipi_10 mg/kg	Diarrhea	Pruritus	Fatigue	Rash	Headache	Weight loss	Nausea	Inc. ALT	Liver function inc.	Hypophysitis		EORTC-QLQ-C30		
CheckMate 066	Melanoma	Niv_3 mg/kg	Fatigue	Pruritus	Nausea	Diarrhea	Rash	Constipation	Vitiligo	Asthenia	Vomiting	Neutropenia	EQ-5D	EORTC-QLQ-C30		
KEYNOTE-006	Melanoma	Ipi_3 mg/kg	Pruritus	Diarrhea	Fatigue	Rash	Nausea	Colitis	Asthenia	Arthralgia	Hyperthyroidism	Hypophysitis	EQ-5D	EORTC-QLQ-C30		
		Pem_10 mg/kg Q2W	Fatigue	Diarrhea	Rash	Pruritus	Asthenia	Hypothyroidism	Nausea	Arthralgia	Vitiligo	Hyperthyroidism				
		Pem_10 mg/kg Q3W	Fatigue	Diarrhea	Pruritus	Rash	Arthralgia	Nausea	Asthenia	Vitiligo	Hypothyroidism	Colitis				
CheckMate 067	Melanoma	Ipi_3 mg/kg	Pruritus	Diarrhea	Rash	Fatigue	Nausea	Dec. appetite	Colitis	Headache	Vomiting	Pyrexia	EQ-5D	EORTC-QLQ-C30		
		Niv + Ipi	Diarrhea	Rash	Fatigue	Pruritus	Nausea	Pyrexia	Dec. appetite	ALT inc.	AST inc.	Vomiting				
		Niv_3 mg/kg	Fatigue	Rash	Diarrhea	Pruritus	Nausea	Dec. appetite	Hypothyroidism	Arthralgia	Headache	Vomiting				
KEYNOTE-002	Melanoma	Pem_2 mg/kg	Fatigue	Pruritus	Rash	Diarrhea	Nausea	Vitiligo	Dry skin	Hypothyroidism	Myalgia	Nausea	EQ-5D	EORTC-QLQ-C30		
		Pem_10 mg/kg	Fatigue	Pruritus	Rash	Diarrhea	Dec. appetite	Nausea	Hypothyroidism	Arthralgia	Maculopap. rash	Vomiting				
MDX010-20	Melanoma	Ipi_3 mg/kg	Fatigue	Nausea	Diarrhea	Dec. appetite	Pruritus	Vomiting	Constipation	Rash	Cough[2]	Abdominal pain	SF-36	EORTC QLQ C30	FACIT	SDS
Javelin Merkel 200	Merkel cell	Ave_10 mg/kg	Fatigue	Infusion reaction	Asthenia	Inc. Lipase	Inc. ALT	Arthralgia	Inc. CPK	Chills[3]	Diarrhea	Dry mouth	EQ-5D	FACT-G	FACT-M	
KEYNOTE-087	Hodgkin Lymp	Pem_200 mg	Hypothyroidism	Pyrexia	Fatigue	Rash	Diarrhea	Headache	Cough	Nausea	Neutropenia	Arthralgia	EQ-5D	EORTC-QLQ-C30		
Study 1108	Urothelial	Dur_10 mg/kg	Fatigue	Dec. appetite	Diarrhea	Rash	Nausea	Arthralgia	Pyrexia	Pruritus	Hypothyroidism	Inc. ALT	EORTC QLQ C30	FACT-G	FACT-BL	Pain
KEYNOTE-045	Urothelial	Pem_200 mg	Pruritus	Fatigue	Nausea	Diarrhea	Dec. appetite	Hypothyroidism	Asthenia	Pneumonitis	Anemia	Colitis		EORTC QLQ C30		
CheckMate 025	Renal-cell	Niv_3 mg/kg	Fatigue	Nausea	Pruritus	Diarrhea	Dec. appetite	Rash	Cough	Anemia	Dyspnea	Peripheral edema	EQ-5D		FKSI-DRS	
15-592	Renal-cell	Ate_1200 mg	Fatigue	Proteinuria	Musculosketal pain	Diarrhea	Rash/skin condition	Hypertension	Pruritus	Thyroid disorder	Hepatitis	Fever		FKSI-19	BFI	
CheckMate 214	Renal-cell	Niv_3 mg/kg+Ipi_1 mg/kg	Fatigue	Pruritus	Diarrhea	Rash	Nausea	Inc. lipase level	Hypothyroidism	Dec. appetite[4]	Asthenia	Vomiting	EQ-5D	FACT-G	FKSI-19	
CheckMate 358	Cervical	Niv_240 mg	Diarrhea	Fatigue	Pneumonitis	Abdominal pain	Stomatitis	Dry eye	Arthralgia	Dec. appetite			EQ-5D	EORTC-QLQ-C30		
	Vaginal/Vulvar	Niv_240 mg	Dec. appetite	Fatigue	Abdominal pain	Arthralgia										

Fig. 3 Comparison of PRO instruments' symptom-related content and AEs reported in ICI-CTs with published PRO results (*n* = 26). The most frequently reported AEs (any grade) are shown for the 26 ICI-CTs identified. AE frequency (most common to least common) is depicted from left to right respectively. Symptom-related content from each PRO instrument was compared to the AEs in the corresponding trial arm. The PRO instruments used in each ICI-CT is shown. For a detailed PRO instrument content, see Table S8. Green: AEs covered; Yellow: AEs partially covered; Light yellow: AEs partially covered related to specific types of pain; Red: AEs not covered. Laboratory tests results (grey) were not included in calculations. "Asthenia" was considered covered by "fatigue" following the National Cancer Institute toxicity grading scale version 3 that included asthenia, lethargy, and malaise under the umbrella of "fatigue". (1) "Fatigue" was considered partially covered because the term included in the FACT-G was "lack of energy"; (2) While the study protocol announced the use of SF-36, EORTC QLQ-C30, FACIT-Fatigue, and SDS questionnaires, only results from the EORTC QLQ-C30 were reported in the full-text publication. "Cough" would be covered by the SDS instrument; (3) "Chills" was partially covered as related to "fever", included in the FACT-G; (4) "Dec. appetite" was considered as partially covered by the term "good appetite" in the FKSI-19. Ipi, Ipilimumab; Niv, Nivolumab; Pem, Pembrolizumab; Ate, Atezolizumab; Dur, Durvalumab; Inc., increased; Dec, decreased

despite relatively frequent grade 3–4 toxicity. Broadly, ICI-CTs results suggest a stable or overall increased HRQoL yet specific ICI treatment-related symptoms and AEs are not adequately assessed by existing PRO tools. In our study, we identified that dermatologic, endocrine, and specific types of pain are among the most reported ICI-related problems regardless of cancer type/diagnosis. Despite the frequency of such events, these AEs are only partially covered if at all by currently applied PRO instruments.

Studies included in this analysis often employed a combination of different PRO instruments (e.g. generic and cancer-specific, cancer-specific with disease-specific and/or symptom-specific questionnaires). The advantage of generic instruments is their broad applicability for evaluating general aspects of HRQoL independent of the specific diagnosis. However, generic instruments tend to lack responsiveness for detecting deterioration in specific health contexts. Disease- or symptom-specific instruments may be more acceptable and responsive to detecting changes but have limited generalizability for identifying

unanticipated events across different populations [13]. In the present review, studies employed PRO instruments based on the intended measure (e.g. to detect changes in HRQoL, GHS, specific disease-related symptoms). Closely examining ICI-related toxicities revealed a unique profile - independent of cancer diagnosis [50–52]. This observation challenges the practice of choosing PRO instruments based on the population rather than treatment type. In other words, both cancer-specific tools and generic HRQoL questionnaires designed for traditional chemotherapies may not capture symptoms and toxicities specific to ICI therapies. Following this line of thought, Kluetz and colleagues have suggested the individual measurement of symptomatic adverse events, physical function, and disease-related symptoms. These three concepts are highly relevant to HRQoL and may better reflect the therapy's effect on the patient and the patient's disease [53]. An alternative approach to the traditional static instruments is the use of item libraries or computerized adaptive testing with content and symptom-items adapted for ICI treatments such as the

National Cancer Institute PRO-CTCAE, PROMIS or the EORTC item library [6, 54–56]. Recently, efforts have been undertaken to develop such ICI specific modules [57].

Our study highlights the general underreporting of PRO results. Of 3231 ICI-CTs registered in the period between 2004 and 2019, only 191 (6%) had a published statement (study protocol, conference abstract) about the measurement of PROs, of which only 26 had a full-text article on PRO results. This finding aligns with the results from a recent systematic evaluation of PROs in cancer trials by Kyte et al. [58]. Importantly, PROs inform and facilitate treatment decision-making in clinical practice. As such, systematic PRO collection, analysis and reporting are highly relevant and should be considered a priority. This delay in PRO reporting is at odds with recommendations from both the European and American oncology societies (ESMO and ASCO) to incorporate such measures in clinical trials as read-outs for treatment benefit evaluation [59, 60].

Conclusion

Our systematic review on the use of PROs in ICI-CTs identifies variability in the use of instruments and gaps in existing instruments for addressing symptomatic toxicities that are ICI treatment-related. Further work is needed to develop novel strategies that permit the more comprehensive and systematic collection of PRO data specific to ICI treatments to inform and support drug labelling and clinical decision-making.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s41687-020-00210-z.

Additional file 1: Table S1. List with the search terms. **Table S2.** List with the inclusion / exclusion criteria. **Table S3.** Data extraction. **Figure S1.** ICI-CT registered in clinicaltrials.gov (2004–2019). **Table S4.** List of identified clinical trials stating the use of PROs. **Table S5.** *Cochrane checklist for describing and assessing PROs in CTs.* **Table S6.** PRO instruments completion rates. **Table S7.** Symptom-related content of PRO instruments. **Table S8.** AEs reported in selected ICI-CT. **Table S9.** AEs differentiated into multi-symptom.

Additional file 2: Figure S1. Frequency of use and combination of PRO instruments in ICI-CTs (*n* = 156). **a** Number of PRO instruments divided into categories and start year of the ICI-CTs. Green: general health status; Blue: cancer-specific; Red: disease/tumor-specific; Yellow: symptom-specific; Dark blue: items libraries/single items; Purple: other (e.g. treatment-specific). **b** Absolute number of instruments across all ICI-CTs divided per category. The table shows the different instruments per category. **c** Number and combination of PRO instruments per ICI-CT per year of start of the trial. For FACT questionnaires that include core domains of the FACT-G, both cancer-specific and disease specific have been depicted (see Table S4 for more detail). For ICI-CTs colored in light grey PRO results have been published (*n* = 26)

Additional file 3: Figure S2. Comparison of PRO instruments' symptom-related content and AEs reported in ICI-CTs with published PRO results (*n* = 26). The most frequently reported AEs (any grade) are shown for the 26 ICI-CTs identified. AE frequency (most common to least common) is depicted from left to right respectively. Symptom-related

content from each PRO instrument was compared to the AEs in the corresponding trial arm. The PRO instruments used in each ICI-CT is shown. For a detailed PRO instrument content, see Table S8. Green: AEs covered; Yellow: AEs partially covered; Light yellow: AEs partially covered related to specific types of pain; Red: AEs not covered. Laboratory tests results (grey) were not included in calculations. "Asthenia" was considered covered by "fatigue" following the National Cancer Institute toxicity grading scale version 3 that included asthenia, lethargy, and malaise under the umbrella of "fatigue". (1) "Fatigue" was considered partially covered because the term included in the FACT-G was "lack of energy"; (2) While the study protocol announced the use of SF-36, EORTC QLQ-C30, FACIT-Fatigue, and SDS questionnaires, only results from the EORTC QLQ-C30 were reported in the full-text publication. "Cough" would be covered by the SDS instrument; (3) "Chills" was partially covered as related to "fever", included in the FACT-G; (4) "Dec. appetite" was considered as partially covered by the term "good appetite" in the FKSI-19. Ipi, Ipilimumab; Niv, Nivolumab; Pem, Pembrolizumab; Ate, Atezolizumab; Dur, Durvalumab; Inc., increased; Dec, decreased

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

SCL and ME received honoraria from Vifor Pharma and research grants from Roche and Bristol Meyers Squibb; RJ serves on advisory boards for Novartis,

Author details
[1]Institute of Higher Education and Research in Healthcare (IUFRS), Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland. [2]Department of Oncology, Lausanne University Hospital (CHUV), Lausanne, Switzerland. [3]School of Nursing, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada. [4]Juravinski Hospital and Cancer Centre at Hamilton Health Sciences, Hamilton, ON, Canada. [5]Department of Oncology, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada. [6]Trillium Health Partners, Mississauga, ON, Canada. [7]International Breast Cancer Study Group (IBCSG), Coordinating Center, Bern, Switzerland. [8]McMaster Evidence and Review Synthesis Team, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada. [9]Population Health Research Institute, Hamilton Health Sciences, McMaster Evidence and Review Synthesis Team, School of Nursing, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada.

References
1. Ribas, A., & Wolchok, J. D. (2018). Cancer immunotherapy using checkpoint blockade. *Science, 359*(6382), 1350–1355.
2. Marrone, K. A., Ying, W., & Naidoo, J. (2016). Immune-related adverse events from immune checkpoint inhibitors. *Clinical Pharmacology and Therapeutics, 100*(3), 242–251.
3. Postow, M. A., Sidlow, R., & Hellmann, M. D. (2018). Immune-related adverse events associated with immune checkpoint blockade. *The New England Journal of Medicine, 378*(2), 158–168.
4. Friedman, C. F., Proverbs-Singh, T. A., & Postow, M. A. (2016). Treatment of the immune-related adverse effects of immune checkpoint inhibitors: A review. *JAMA Oncology, 2*(10), 1346–1353.
5. U.S. Department of Health and Human Services N.I.H., National Cancer Institute (2017) Common terminology criteria for adverse events (CTCAE) version 5.0. p https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_8.5x11.pdf.
6. Basch, E., Reeve, B. B., Mitchell, S. A., Clauser, S. B., Minasian, L. M., Dueck, A. C., et al. (2014). Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). *Journal of the National Cancer Institute, 106*(9), dju244.
7. Fromme, E. K., Eilers, K. M., Mori, M., Hsieh, Y. C., & Beer, T. M. (2004). How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the quality-of-life questionnaire C30. *Journal of Clinical Oncology, 22*(17), 3485–3490.
8. Di Maio, M., Basch, E., Bryce, J., & Perrone, F. (2016). Patient-reported outcomes in the evaluation of toxicity of anticancer treatments. *Nature Reviews. Clinical Oncology, 13*(5), 319–325.
9. U.S. Department of Health and Human Services Food and Drug Administration F.D.A. (2009) Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims.
10. Atkinson, T. M., Ryan, S. J., Bennett, A. V., Stover, A. M., Saracino, R. M., Rogak, L. J., et al. (2016). The association between clinician-based common terminology criteria for adverse events (CTCAE) and patient-reported outcomes (PRO): a systematic review. *Support Care Cancer, 24*(8), 3669–3676.
11. Basch, E. (2017). Patient-reported outcomes - harnessing patients' voices to improve clinical care. *The New England Journal of Medicine, 376*(2), 105–108.
12. Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., et al. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ, 339*, b2700.
13. Donald L Patrick, Gordon H Guyatt, Catherine Acquadro, & on behalf of the Cochrane Patient Reported Outcomes Methods Group (2011) Chapter 17: patient-reported outcomes. *Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]*, (The Cochrane Collaboration).
14. Brahmer, J. R., Rodriguez-Abreu, D., Robinson, A. G., Hui, R., Csoszi, T., Fulop, A., et al. (2017). Health-related quality-of-life results for pembrolizumab versus chemotherapy in advanced, PD-L1-positive NSCLC (KEYNOTE-024): A multicentre, international, randomised, open-label phase 3 trial. *The Lancet Oncology, 18*(12), 1600–1609.
15. Mazieres, J., Kowalski, D., Luft, A., Vicente, D., Tafreshi, A., Gumus, M., et al. (2020). Health-related quality of life with carboplatin-paclitaxel or nab-paclitaxel with or without pembrolizumab in patients with metastatic squamous non-small-cell lung cancer. *Journal of Clinical Oncology, 38*(3), 271–280.
16. Reck, M., Brahmer, J., Bennett, B., Taylor, F., Penrod, J. R., DeRosa, M., et al. (2018). Evaluation of health-related quality of life and symptoms in patients with advanced non-squamous non-small cell lung cancer treated with nivolumab or docetaxel in CheckMate 057. *European Journal of Cancer, 102*, 23–30.
17. Reck, M., Taylor, F., Penrod, J. R., DeRosa, M., Morrissey, L., Dastani, H., et al. (2018). Impact of nivolumab versus docetaxel on health-related quality of life and symptoms in patients with advanced squamous non-small cell lung cancer: results from the CheckMate 017 study. *Journal of Thoracic Oncology, 13*(2), 194–204.
18. Reck, M., Schenker, M., Lee, K. H., Provencio, M., Nishio, M., Lesniewski-Kmak, K., et al. (2019). Nivolumab plus ipilimumab versus chemotherapy as first-line treatment in advanced non-small-cell lung cancer with high tumour mutational burden: patient-reported outcomes results from the randomised, open-label, phase III CheckMate 227 trial. *European Journal of Cancer, 116*, 137–147.
19. Spigel, D. R., McCleod, M., Jotte, R. M., Einhorn, L., Horn, L., Waterhouse, D. M., et al. (2019). Safety, efficacy, and patient-reported health-related quality of life and symptom burden with nivolumab in patients with advanced non-small cell lung cancer, including patients aged 70 years or older or with poor performance status (CheckMate 153). *Journal of Thoracic Oncology, 14*(9), 1628–1639.
20. Barlesi, F., Garon, E. B., Kim, D. W., Felip, E., Han, J. Y., Kim, J. H., et al. (2019). Health-related quality of life in KEYNOTE-010: A phase II/III study of pembrolizumab versus docetaxel in patients with previously treated advanced, programmed death ligand 1-expressing NSCLC. *Journal of Thoracic Oncology, 14*(5), 793–801.
21. Hui, R., Ozguroglu, M., Villegas, A., Daniel, D., Vicente, D., Murakami, S., et al. (2019). Patient-reported outcomes with durvalumab after chemoradiotherapy in stage III, unresectable non-small-cell lung cancer (PACIFIC): a randomised, controlled, phase 3 study. *The Lancet Oncology, 20*(12), 1670–1680.
22. Bordoni, R., Ciardiello, F., von Pawel, J., Cortinovis, D., Karagiannis, T., Ballinger, M., et al. (2018). Patient-reported outcomes in OAK: a phase III study of Atezolizumab versus docetaxel in advanced non-small-cell lung cancer. *Clinical Lung Cancer, 19*(5), 441–449 e444.
23. Bauml, J. M., Mick, R., Ciunci, C., Aggarwal, C., Davis, C., Evans, T., et al. (2019). Pembrolizumab after completion of locally ablative therapy for oligometastatic non-small cell lung cancer: a phase 2 trial. *JAMA Oncology, 5*, 1283–1290.
24. Mansfield, A. S., Kazarnowicz, A., Karaseva, N., Sanchez, A., De Boer, R., Andric, Z., et al. (2020). Safety and patient-reported outcomes of atezolizumab, carboplatin, and etoposide in extensive-stage small-cell lung cancer (IMpower133): a randomized phase I/III trial. *Annals of Oncology, 31*(2), 310–317.
25. Schadendorf, D., Larkin, J., Wolchok, J., Hodi, F. S., Chiarion-Sileni, V., Gonzalez, R., et al. (2017). Health-related quality of life results from the phase III CheckMate 067 study. *European Journal of Cancer, 82*, 80–91.
26. Schadendorf, D., Dummer, R., Hauschild, A., Robert, C., Hamid, O., Daud, A., et al. (2016). Health-related quality of life in the randomised KEYNOTE-002 study of pembrolizumab versus chemotherapy in patients with ipilimumab-refractory melanoma. *European Journal of Cancer, 67*, 46–54.
27. Coens, C., Suciu, S., Chiarion-Sileni, V., Grob, J. J., Dummer, R., Wolchok, J. D., et al. (2017). Health-related quality of life with adjuvant ipilimumab versus placebo after complete resection of high-risk stage III melanoma (EORTC 18071): secondary outcomes of a multinational, randomised, double-blind, phase 3 trial. *The Lancet Oncology, 18*, 393–403.
28. Revicki, D. A., van den Eertwegh, A. J., Lorigan, P., Lebbe, C., Linette, G., Ottensmeier, C. H., et al. (2012). Health related quality of life outcomes for unresectable stage III or IV melanoma patients receiving ipilimumab treatment. *Health and Quality of Life Outcomes, 10*, 66.
29. Petrella, T. M., Robert, C., Richtig, E., Miller Jr., W. H., Masucci, G. V., Walpole, E., et al. (2017). Patient-reported outcomes in KEYNOTE-006, a randomised study of pembrolizumab versus ipilimumab in patients with advanced melanoma. *European Journal of Cancer, 86*, 115–124.
30. Long, G. V., Atkinson, V., Ascierto, P. A., Robert, C., Hassel, J. C., Rutkowski, P., et al. (2016). Effect of nivolumab on health-related quality of life in patients with treatment-naive advanced melanoma: results from the phase III CheckMate 066 study. *Annals of Oncology, 27*(10), 1940–1946.
31. Cella, D., Grunwald, V., Nathan, P., Doan, J., Dastani, H., Taylor, F., et al. (2016). Quality of life in patients with advanced renal cell carcinoma given nivolumab versus everolimus in CheckMate 025: a randomised, open-label, phase 3 trial. *The Lancet Oncology, 17*(7), 994–1003.

32. Cella, D., Grunwald, V., Escudier, B., Hammers, H. J., George, S., Nathan, P., et al. (2019). Patient-reported outcomes of patients with advanced renal cell carcinoma treated with nivolumab plus ipilimumab versus sunitinib (CheckMate 214): a randomised, phase 3 trial. *The Lancet Oncology, 20*(2), 297–310.

33. McGregor, B. A., McKay, R. R., Braun, D. A., Werner, L., Gray, K., Flaifel, A., et al. (2020). Results of a multicenter phase II study of atezolizumab and bevacizumab for patients with metastatic renal cell carcinoma with variant histology and/or sarcomatoid features. *Journal of Clinical Oncology, 38*(1), 63–70.

34. O'Donnell, P. H., Arkenau, H. T., Sridhar, S. S., Ong, M., Drakaki, A., Spira, A. I., et al. (2020). Patient-reported outcomes and inflammatory biomarkers in patients with locally advanced/metastatic urothelial carcinoma treated with durvalumab in phase 1/2 dose-escalation study 1108. *Cancer, 126*(2), 432–443.

35. Vaughn, D. J., Bellmunt, J., Fradet, Y., Lee, J. L., Fong, L., Vogelzang, N. J., et al. (2018). Health-related quality-of-life analysis from KEYNOTE-045: a phase III study of pembrolizumab versus chemotherapy for previously treated advanced urothelial cancer. *Journal of Clinical Oncology, 36*(16), 1579–1587.

36. Chen, R., Zinzani, P. L., Fanale, M. A., Armand, P., Johnson, N. A., Brice, P., et al. (2017). Phase II study of the efficacy and safety of pembrolizumab for relapsed/refractory classic Hodgkin lymphoma. *Journal of Clinical Oncology, 35*(19), 2125–2132.

37. von Tresckow, B., Fanale, M., Ardeshna, K. M., Chen, R., Meissner, J., Morschhauser, F., et al. (2019). Patient-reported outcomes in KEYNOTE-087, a phase 2 study of pembrolizumab in patients with classical Hodgkin lymphoma. *Leukemia & Lymphoma, 60*(11), 2705–2711.

38. Harrington, K. J., Ferris, R. L., Blumenschein Jr., G., Colevas, A. D., Fayette, J., Licitra, L., et al. (2017). Nivolumab versus standard, single-agent therapy of investigator's choice in recurrent or metastatic squamous cell carcinoma of the head and neck (CheckMate 141): health-related quality-of-life results from a randomised, phase 3 trial. *The Lancet Oncology, 18*, 1104–1115.

39. Ferris, R. L., Blumenschein Jr., G., Fayette, J., Guigay, J., Colevas, A. D., Licitra, L., et al. (2016). Nivolumab for recurrent squamous-cell carcinoma of the head and neck. *The New England Journal of Medicine, 375*(19), 1856–1867.

40. Kaufman, H. L., Hunger, M., Hennessy, M., Schlichting, M., & Bharmal, M. (2018). Nonprogression with avelumab treatment associated with gains in quality of life in metastatic Merkel cell carcinoma. *Future Oncology, 14*(3), 255–266.

41. Naumann, R. W., Hollebecque, A., Meyer, T., Devlin, M. J., Oaknin, A., Kerger, J., et al. (2019). Safety and efficacy of nivolumab monotherapy in recurrent or metastatic cervical, vaginal, or vulvar carcinoma: results from the phase I/II CheckMate 358 trial. *Journal of Clinical Oncology, 37*(31), 2825–2834.

42. Bharmal, M., Fofana, F., Barbosa, C. D., Williams, P., Mahnke, L., Marrel, A., et al. (2017). Psychometric properties of the FACT-M questionnaire in patients with Merkel cell carcinoma. *Health and Quality of Life Outcomes, 15*(1), 247.

43. King-Kallimanis, B. L., Howie, L. J., Roydhouse, J. K., Singh, H., Theoret, M. R., Blumenthal, G. M., et al. (2019). Patient reported outcomes in anti-PD-1/PD-L1 inhibitor immunotherapy registration trials: FDA analysis of data submitted and future directions. *Clinical Trials, 16*(3), 322–326.

44. Hall, E. T., Singhal, S., Dickerson, J., Gabster, B., Wong, H.-N., Aslakson, R. A., et al. (2019). Patient-reported outcomes for cancer patients receiving checkpoint inhibitors: opportunities for palliative care - a systematic review. *Journal of Pain and Symptom Management, 58*(1), 137–156.e131.

45. European Medicines Agency E.M.A. (2016) Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man. The use of patient-reported outcome (PRO) measures in oncology studies.

46. Brundage, M., & Hanna, T. (2017). Adjuvant ipilimumab for stage III melanoma: the patient voice. *The Lancet Oncology, 18*(3), 282–284.

47. Lorigan, P., & Green, A. C. (2017). Immunotherapy: does adjuvant ipilimumab have little adverse effect on quality of life? *Nature Reviews. Clinical Oncology, 14*(7), 395–396.

48. Singer, S. (2017). Quality of life after nivolumab treatment for head and neck cancer. *The Lancet Oncology, 18*(8), 993–994.

49. Das, S., & Horn, L. (2017). Quality matters: Immunotherapy and the evolving landscape of advanced cancer care. *Expert Review of Quality of Life in Cancer Care, 2*(5), 235–244.

50. De Velasco, G., Je, Y., Bossé, D., Awad, M. M., Ott, P. A., Moreira, R. B., et al. (2017). Comprehensive meta-analysis of key immune-related adverse events from CTLA-4 and PD-1/PD-L1 inhibitors in cancer patients. *Cancer Immunology Research, 5*(4), 312–318.

51. El Osta, B., Hu, F., Sadek, R., Chintalapally, R., & Tang, S. C. (2017). Not all immune-checkpoint inhibitors are created equal: meta-analysis and systematic review of immune-related adverse events in cancer trials. *Critical Reviews in Oncology/Hematology, 119*, 1–12.

52. Khoja, L., Day, D., Wei-Wu Chen, T., Siu, L. L., & Hansen, A. R. (2017). Tumour- and class-specific patterns of immune-related adverse events of immune checkpoint inhibitors: a systematic review. *Annals of Oncology, 28*(10), 2377–2385.

53. Kluetz, P. G., Slagle, A., Papadopoulos, E. J., Johnson, L. L., Donoghue, M., Kwitkowski, V. E., et al. (2016). Focusing on core patient-reported outcomes in cancer clinical trials: symptomatic adverse events, physical function, and disease-related symptoms. *Clinical Cancer Research, 22*(7), 1553–1558.

54. Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., et al. (2010). The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *Journal of Clinical Epidemiology, 63*(11), 1179–1194.

55. Kulis, D., Bottomley, A., Whittaker, C., van de Poll-Franse, L. V., Darlington, A., Holzner, B., et al. (2017). The use of the Eortc item library to supplement Eortc quality of life instruments. *Value in Health, 20*(9), A775.

56. Petersen, M. A., Groenvold, M., Aaronson, N. K., Chie, W. C., Conroy, T., Costantini, A., et al. (2011). Development of computerized adaptive testing (CAT) for the EORTC QLQ-C30 physical functioning dimension. *Quality of Life Research, 20*(4), 479–490.

57. Hansen, A. R., Ala-Leppilampi, K., McKillop, C., Siu, L. L., Bedard, P. L., Abdul Razak, A. R., et al. (2020). Development of the functional assessment of cancer therapy-immune checkpoint modulator (FACT-ICM): a toxicity subscale to measure quality of life in patients with cancer who are treated with ICMs. *Cancer, 126*(7), 1550–1558.

58. Kyte, D., Retzer, A., Ahmed, K., Keeley, T., Armes, J., Brown, J. M., et al. (2019). Systematic evaluation of patient-reported outcome protocol content and reporting in cancer trials. *Journal of the National Cancer Institute, 111*, 1170–1178.

59. Cherny, N. I., Sullivan, R., Dafni, U., Kerst, J. M., Sobrero, A., Zielinski, C., et al. (2015). A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: The European Society for Medical Oncology magnitude of clinical benefit scale (ESMO-MCBS). *Annals of Oncology, 26*(8), 1547–1573.

60. Schnipper, L. E., Davidson, N. E., Wollins, D. S., Blayney, D. W., Dicker, A. P., Ganz, P. A., et al. (2016). Updating the American Society of Clinical Oncology value framework: revisions and reflections in response to comments received. *Journal of Clinical Oncology, 34*(24), 2925–2934.

Publisher's Note