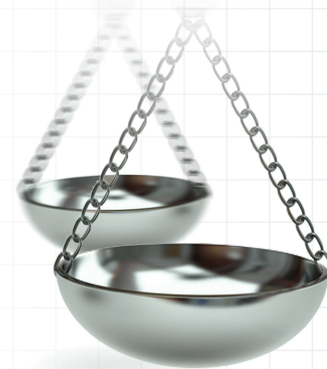




Clinician-reported symptomatic adverse events in cancer trials: are they concordant with patient-reported outcomes?

Journal of **Comparative Effectiveness Research**

Francesco Sparano^{*1}, Neil K Aaronson², Francesco Cottone¹, Alfonso Piciocchi¹, Edoardo La Sala¹, Amelie Anota³, Nina Deliu¹, Jacobien M Kieffer² & Fabio Efficace¹

¹Italian Group for Adult Hematologic Diseases (GIMEMA), Data Center & Health Outcomes Research Unit, Rome, Italy

²Department of Psychosocial Research, Division of Psychosocial Research & Epidemiology, The Netherlands Cancer Institute, The Netherlands

³Quality of Life in Oncology French National Platform, Methodological & Quality of Life in Oncology Unit (INSERM UMR 1098), University Hospital of Besançon, Besançon, France

*Author for correspondence: Tel.: +39 06 441 639 837; f.sparano@gimema.it

Aim: We investigate the concordance, in terms of favoring the same treatment arm, between clinician-reported symptomatic adverse events (AEs) and information obtained via patient-reported outcomes (PRO) measures in cancer randomized controlled trials (RCTs). **Methods:** We conducted a systematic literature search to identify all RCTs conducted in breast, colorectal, lung and prostate cancer, published between 2004 and 2017. **Results:** We identified 207 RCTs. In the majority of RCTs ($n = 133$, 64.2%) a discordance between PROs and AEs was found. In 104 studies (50.2%), PRO data favored the experimental arm when AEs did not, while the opposite situation was found in 29 trials (14.0%). **Conclusion:** Frequently, information obtained via PRO measures and clinician-reported AEs do not favor the same treatment arm in RCT settings.

First draft submitted: 7 September 2018; Accepted for publication: 11 January 2019; Published online: 6 March 2019

Keywords: adverse events • cancer • patient-reported outcomes • randomized controlled trials • symptoms

Although cancer treatments are continuously improving, they all impact patient's well-being and daily functioning [1]. Even modern anti-cancer targeted therapies, such as small molecule inhibitors, generate side effects and adverse events (AEs) that need to be properly managed in clinical practice [2].

Appropriate documentation and reporting of AEs is particularly relevant in randomized controlled trials (RCTs), where new anti-cancer drugs are compared with standards of care with the ultimate goal of demonstrating better clinical efficacy in patients treated with the experimental treatment [3]. Indeed, reporting the incidence and severity of AEs between treatment arms (experimental versus control) is critical to guarantee patient safety and to better understand the overall clinical value of the new drug being tested. Traditionally, ratings of AEs in RCTs have relied upon standardized coding systems, such as the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) [4]. The majority of AEs in this classification can be measured with lab tests (e.g., anemia, neutropenia or leukocytopenia); however, some AEs are evaluated via physicians' judgment and refer to symptomatic toxicities (e.g., fatigue, pain or headache).

Over the last decade, empirical data have called into question the ability of physicians to accurately capture their patients' symptom burden [3,5–9]. For this reason, the NCI recently developed a patient-reported outcome (PRO) version of the CTCAE, which includes a list of 124 items representing 78 symptomatic AEs that are to be evaluated by patients themselves [3,10].

In an effort to better establish the potential benefits (if any) of new anticancer drugs from the patient's perspective, the number of RCTs that have included PRO measures of symptom burden and health-related quality of life (HRQOL) has increased substantially over the last decades [11]. PRO measures are typically multi-dimensional questionnaires to be completed by patients themselves, and cover various domains of the patient's experience, including symptoms, functional abilities or psychosocial concerns. They are intended to complement other more

traditional RCT outcomes, such as progression-free survival and clinical response, in order to more comprehensively evaluate the impact of therapy from the patient's standpoint. The extent to which RCT outcomes are consistent in supporting a new drug being tested is an important criterion for making overall treatment recommendations. However, there is little knowledge on the extent of concordance between clinical trial outcomes based on PRO measures and those based on more traditional clinician-rated AEs.

The primary objective of this study was to systematically investigate the concordance (i.e., favoring the same treatment arm), between clinician-reported symptomatic AEs and information obtained via PRO measures in RCT settings. A secondary objective was to examine the concordance between patient- and clinician-reported individual symptomatic AEs.

Methods

Literature search

We conducted a systematic literature search in PubMed/MEDLINE, the Cochrane library, PsycINFO and PsycARTICLES to identify all cancer RCTs that have included at least one PRO, either as a primary or secondary study endpoint. We selected all RCTs published between January 2004 and April 2017 in breast, colorectal, lung and prostate cancer that enrolled at least 50 patients. The search included studies published since 2004 as previous work had already evaluated cancer RCTs, including PRO measures, published before that date [12]. We focused on these disease sites as they are considered the 'big four' (i.e., highest incidence) cancer types [13]. Also, it has been estimated that these four cancer types will account for approximately 45% of cancer deaths in the USA in 2018 [14].

Studies assessing prevention or screening programs, complementary or alternative medicine or psychosocial intervention were excluded. The search was restricted to English language articles. If a selected study had multiple publications, we incorporated all relevant papers in the current analysis. Inclusion criteria were based on previously conducted systematic reviews [11,15,16].

Study selection

For the primary study objective, we only selected RCTs that included data on both clinician-reported AEs and PROs (regardless of the PRO instrument used). In order to evaluate the concordance between clinician-reported symptomatic AEs (in the remainder of the text, we simply refer to: AEs) and PROs, we included those studies that reported both frequencies of AEs (regardless of grade severity) and statistical significance of PROs between treatment arms. For the purpose of our analysis, we only considered the AEs included in the PRO-CTCAE list [10,17]. If one of these symptoms was reported using a different wording (e.g. arthralgia used to refer to joint pain) we selected it regardless of the term used (for the full list of synonyms considered see Supplementary Material). Studies reporting only descriptive results from PRO measures, (i.e., not applying a test for statistical significance), or studies reporting only laboratory-based AEs were excluded.

For the secondary objective, focusing on the individual analysis of AEs, we restricted the selection to RCTs assessing PROs with the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (EORTC QLQ-C30) questionnaire [18]; the rationale being that the EORTC QLQ-C30 is one of the very few multidimensional PRO measures that assess a core set of symptoms separately from broader health domains. Also, the way in which the questionnaire's symptoms-related questions are worded, and the response choices used, facilitate comparison with AE [19]. The EORTC QLQ-C30 includes the following core set of major symptoms often experienced by cancer patients: fatigue, pain, nausea/vomiting, dyspnoea, insomnia, appetite loss, constipation and diarrhoea [20]. Details on the selection process are reported according to the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) guidelines [21].

Evaluation of concordance between multidimensional patient-reported outcomes & clinician-reported symptomatic AEs

For our primary study objective, we classified PROs and AEs as better, equal or worse in the experimental with respect to the control arm. PRO analysis was conducted considering all subscales (e.g., functional or psychosocial aspects, symptom scales or global quality of life scales) of the PRO instruments used in each study. Specifically, in each RCT, we determined the number of subscales that reported a statistically significant and/or clinically meaningful difference between arms at any time point during the study period. If more than half of the PRO instrument subscales favored the experimental arm, we defined this as 'Better PRO'. Conversely, if more than half of the PROs favored the control arm, we defined this as 'Worse PRO'. If none of the subscales of the PRO instrument

demonstrated a statistically significant difference favoring one of the two arms, or the number of significant differences was balanced between the study arms, we defined this as 'PRO equivalence'. For the classification of AEs, we used previously reported criteria proposed by Schuurhuizen and colleagues [22]. We determined the number of AEs (any grade) reported at least 5% more frequently in one treatment arm versus the other(s). If, based on these criteria, two or more of the AEs were reported more frequently in the experimental arm than the control arm, then we defined this scenario as 'Worse AEs'. Conversely, if two or more of the AEs were reported more frequently in the control group than in the experimental group(s), then we defined this scenario as 'Better AEs'. If the number of AEs reported in study arms were similar (only one or no differences between arms), then we defined it as 'AEs equivalence'. In those trials where only severe AEs grades were reported, we applied the same rating criteria as described above, but restricted to severe AEs.

Evaluation of the concordance between patient- & clinician-reported individual symptomatic AEs

For our secondary objective, we considered the following eight patient-reported symptoms as assessed by the EORTC QLQ-C30: nausea/vomiting, fatigue, pain, constipation, diarrhoea, dyspnoea, insomnia and appetite loss. For each study, we classified the outcomes as better, worse or equal in the experimental arm with respect to the control arm. In the case of patient-reported symptoms, when a symptom showed a statistically significant difference between arms, we considered it as 'better' or 'worse' if it favored the experimental or control arm, respectively. If no differences were found, this was defined as 'equal'. For clinician-reported AEs, when a single AE was reported at least 5% more frequently in the experimental or control arm, we considered it as 'worse' or 'better,' respectively. If a difference of less than 5% was reported, this was defined as 'equal'. We conducted the same analysis both for any grade AEs and severe grades AEs, whenever possible.

For the purpose of this analysis, patient- reported symptoms and clinician-reported AEs were grouped into three categories. Concordance between patients and clinicians: when both patients and clinicians classified the symptom as equal, or better or worse in the experimental arm, patients favored the experimental arm: situations where the symptom in the experimental arm was reported by the patients to be better than the clinicians' evaluation (i.e., equal PRO-Worse AE, better PRO-worse AE and better PRO-equal AE), clinicians favored the experimental arm: situations where the symptom in the experimental arm was reported by clinicians to be better than the patients' evaluation (i.e., equal PRO-better AE, worse PRO-equal AE and worse PRO-better AE).

Findings are reported using descriptive statistics. All analyses were performed with SAS version 9.4 (SAS Institute Inc., NC, USA).

Results

Overview of studies

We identified 488 RCTs in breast, colorectal, lung and prostate cancer published between January 2004 and April 2017. From this list, we excluded 65 RCTs for which it was not possible to evaluate PROs (e.g., only descriptive PRO data were reported without statistical testing between study arms). Of the remaining 423 RCTs, 216 were excluded for other reasons (Figure 1). The final sample included 207 RCTs, 46 of which were used in the single-symptom analysis with the EORTC QLQ-C30.

Of the 207 RCTs, 69 (33.3%) were in lung cancer, 61 (29.5%) in breast cancer, 41 (19.8%) in colorectal cancer and 36 (17.4%) in prostate cancer. The majority of these trials (n = 150, 72.5%) included patients with metastatic disease. Nearly half of the studies (n = 99, 47.8%) were multinational, and 149 (72.0%) enrolled more than 200 patients. The NCI-CTCAE was used to assess AEs in 154 RCTs (74.4%); 13 RCTs (6.3%) evaluated AEs according to the WHO criteria [23], and in 30 trials (14.5%) the instrument used was not reported. The three most frequently used PRO questionnaires in these 207 RCTs were the EORTC questionnaires, used in 98 RCTs (47.3%), the FACT questionnaires (n = 65, 31.4%) and VAS or LASA questionnaires (n = 17, 8.2%). In 31 (15%) RCTs, PRO was primary endpoint of the study and in 176 (85%) secondary endpoint. With regard to RCT design, we found that 67 (32.4%) were blinded studies, 51 (24.6%) were open-label and for 89 (43%) this aspect was not evaluable. Details are reported in Table 1.

Concordance between patient-reported outcome measures & AEs reporting

The overall concordance analysis was therefore performed on 207 RCTs. Concordance between PROs and AEs was observed in slightly more than one-third of studies (n = 74, 35.8%). In 13 RCTs (6.3%), both AEs and PROs

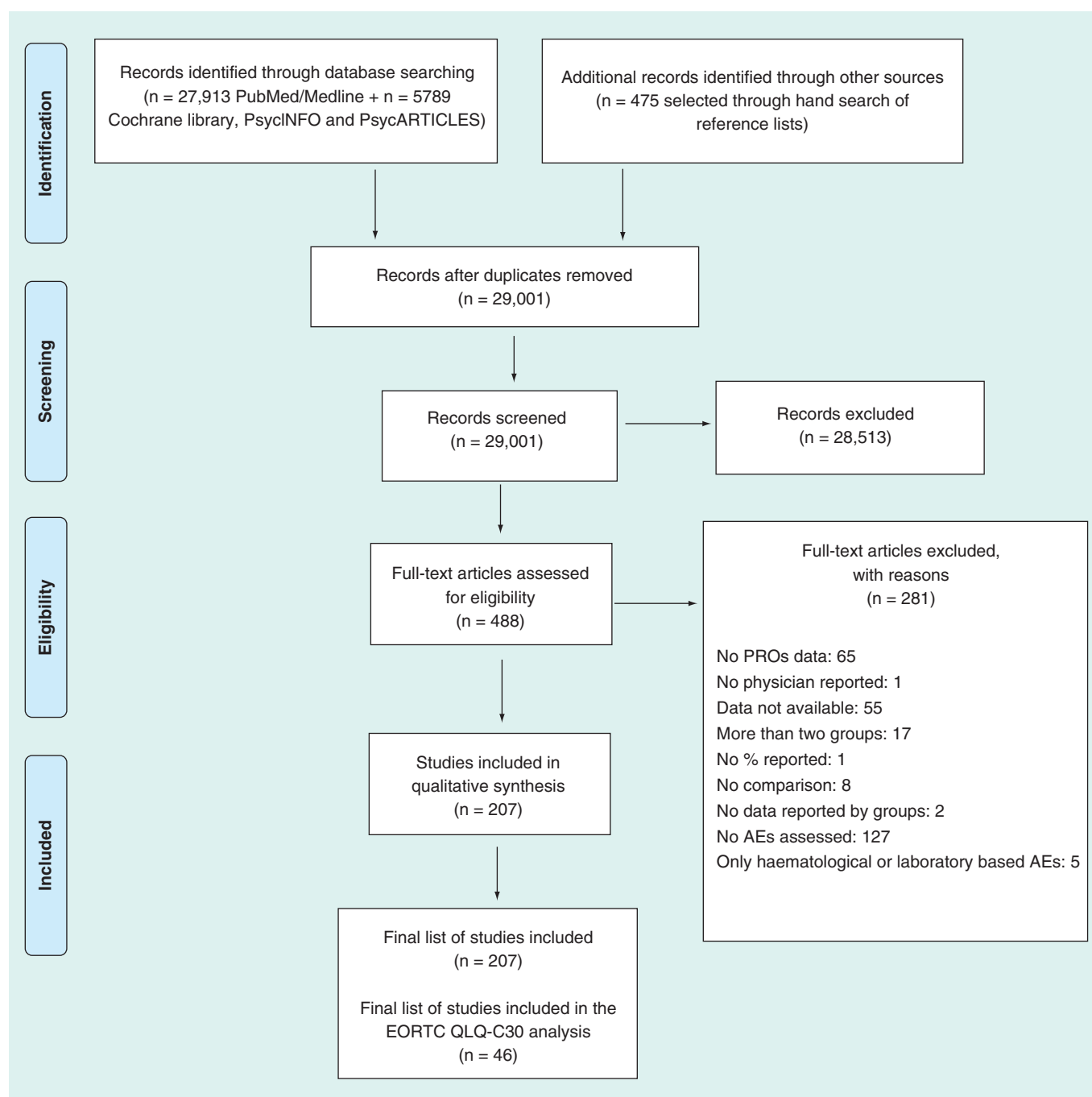


Figure 1. Schematic breakdown of literature search results of breast, colorectal, lung and prostate randomized controlled trials (preferred reporting items for systematic reviews and meta-analysis).

AE: Adverse event; EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; PRO: Patient-reported outcomes.

avored the experimental arm, in 13 (6.3%) both favored the control arm, and in 48 trials (23.2%) neither the AEs nor the PROs favored one treatment arm over the other (Table 2).

The majority of RCTs reviewed (n = 133, 64.2%) showed a discordance between PROs and AEs. Of the nine scenarios depicted in Table 2, the most frequent were those in which no differences were observed between treatment arms according to both the AE and the PRO reporting (n = 48, 23.2%), and in which the AEs were more prevalent

Table 1. Characteristics of the studies.

Item	Category	Total (N = 207)
Instrument used to evaluate AEs, n (%)	NCI-CTCAE	154 (74.4)
	WHO criteria	13 (6.3)
	Others	10 (4.8)
	Not reported	30 (14.5)
Type of cancer, n (%)	Lung	69 (33.3)
	Breast	61 (29.5)
	Colorectal	41 (19.8)
	Prostate	36 (17.4)
International, n (%)	Yes	99 (47.8)
	No	108 (52.2)
Blinding [†]	Blinded trial	67 (32.4)
	Open label trial	51 (24.6)
	Unclear	89 (43.0)
Secondary paper on PRO results [‡] , n (%)	Yes	66 (31.9)
	No	141 (68.1)
PRO study end point, n (%)	Primary	31 (15.0)
	Secondary	176 (85.0)
Disease stage, n (%)	Metastatic	150 (72.5)
	Non-metastatic	36 (17.4)
	Both	17 (8.2)
	Unclear	4 (1.9)
Overall study sample size, n (%)	≤200 patients	58 (28.0)
	>200 patients	149 (72.0)
Year of publication, n (%)	2004–2010	130 (62.8)
	2011–2017	77 (37.2)
Treatment type, n (%) [¶]	Chemotherapy	145 (70.0)
	Targeted therapy	58 (28.0)
	Hormonal therapy	29 (14.0)
	Radiotherapy	17 (8.2)
	Surgery	6 (2.9)
	Other	26 (12.6)
PRO measure used, n (%) [¶]	EORTC questionnaires	98 (47.3)
	FACT questionnaires	65 (31.4)
	VAS or LASA questionnaires	17 (8.2)
	SF-36	4 (1.9)
	Other PROs	99 (47.8)

[†] Assessed using the item 'performance bias' of the Cochrane risk of bias tool (Higgins JP, Altman DG, Gotzsche PC, *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomized trials. *BMJ*. 2011;343:d5928). The performance bias evaluates the blinding of participants and personnel. The risk of bias is assessed as 'low,' 'high' or 'unclear.' RCTs were classified in three groups: 'open-label trial' when a high risk of performance bias was found, 'blinded trial' when low risk of performance bias was found and 'unclear' when this aspect was not evaluable.

[‡] Assessed as 'yes' if at least one paper focused on PRO outcomes was published in addition to the original RCT report.

[¶] More than one option could be chosen.

AE: Adverse event; EORTC: European Organization for Research and Treatment of Cancer; FACT: Functional Assessment of Cancer Therapy; LASA: Linear Analogue Self-Assessment; NCI-CTCAE: National Cancer Institute-Common Terminology Criteria for Adverse Events; RCT: Randomized controlled trial; SF-36: Medical Outcomes Study 36-Item Short-Form Health Survey; VAS: Visual Analog Scale; PRO: Patient-Reported Outcomes.

in the experimental than the control arm, when the PRO data indicated equivalence between treatment arms (n = 42, 20.3%).

The least frequent scenario was that in which PRO data favored the control arm, when AEs favored the experimental arm (n = 2, 1.0%). In half of the trials (n = 104, 50.2%), PRO data favored the experimental arm when AEs did not, while the opposite situation was found in 29 trials (14.0%).

Table 2. Concordance between patient-reported outcomes and adverse events (N = 207).

Patient-reported outcomes	Adverse events		
	Better AEs	Equal AEs	Worse AEs
Better PRO n (%)	13 (6.3)	29 (14.0)	33 (15.9)
PRO equivalence n (%)	14 (6.7)	48 (23.2)	42 (20.3)
Worse PRO n (%)	2 (1.0)	13 (6.3)	13 (6.3)

Definitions of 'better' or 'worse' AEs and/or PROs are reported in the methods paragraph of the manuscript.
AE: Adverse event; PRO: Patient-reported outcome.

Table 3. Concordance between patient reported outcomes results and adverse events by patient reported outcomes endpoint and blinding.

Variable		Concordance		Total
		Yes	No	
PRO end point, n (%)	Primary	14 (45.2)	17 (54.8)	31
	Secondary	60 (34.1)	116 (65.9)	176
Blinding, n (%)	Blinded trial	19 (28.4)	48 (71.6)	67
	Open-label	18 (35.3)	33 (64.7)	51

Row percentages are reported for each category.
PRO: Patient-reported outcome.

Further descriptive analyses, showed that level of concordance varied according to some key characteristics of the study. Indeed, concordance between PROs and AEs was more frequent in open-label trials (35.3%) than in blinded trials (28.4%) and in RCTs with PRO as primary endpoint (45.2%) than in those with PRO as secondary end point (34.1%; [Table 3](#)).

Concordance between patient- & clinician-reported individual symptomatic AEs

46 studies using the EORTC QLQ-C30 were included in the analysis of single-symptoms concordance, and two separate analyses were performed considering either severe AEs (based on data from 45 RCTs) or any grade AEs (34 RCTs). Out of eight symptoms originally planned to be assessed, we excluded insomnia as no study reported data based on both patient and clinician report with regard to this symptom. Details are reported in [Figure 2A](#) and [B](#).

For any grade nausea/vomiting ([Figure 2A](#)), patients evaluated the symptom more favorably in the experimental arm compared with clinicians' assessments in 63% of trials, while the patients' assessments agreed with those of the clinicians in only 22.2% of the cases. Appetite loss and pain showed a similar pattern, with the PRO data indicating less symptom burden in the experimental group as compared with AE ratings in 50% of the cases. The largest agreement between patients and clinicians was observed for dyspnoea (75%) and constipation (68.8%).

Analysis of severe grade symptoms resulted in a slightly different scenario ([Figure 2B](#)). Appetite loss was the symptom where discordance between patient and clinician was highest. In 50% of the cases, appetite loss in the experimental group was evaluated by patients as being less problematic than by clinicians. The opposite pattern was found in 25% of cases; and in the remaining 25% of cases, patients and clinician's ratings were concordant. Patient–physician concordance was higher for fatigue (63%), diarrhoea (58.8%) and constipation (68.8%).

Discussion

Previous work in this area has focused on the comparison of patient versus clinicians' ratings of the severity and frequency of symptomatic AEs with results generally indicating that clinicians frequently underestimate symptom burden [7,9,19,24–28]. Our systematic review shifted the focus to the evaluation of concordance between clinician-reported AEs and PROs in supporting the same treatment arm in RCTs. Our results indicate discordance between PROs and clinician-reported AEs in almost two-thirds of the RCTs analysed. This suggests the need to pay special attention to such discordance between PROs and AE ratings when making treatment recommendations based on overall RCT results and future work should thoroughly examine if and how such information is disclosed in study summary findings and factor into overall recommendations.

In a recent review of studies in metastatic colorectal cancer [22], substantial level of discordance was reported between PROs and AEs. However, it is difficult to compare our findings with this study [22] as the authors only

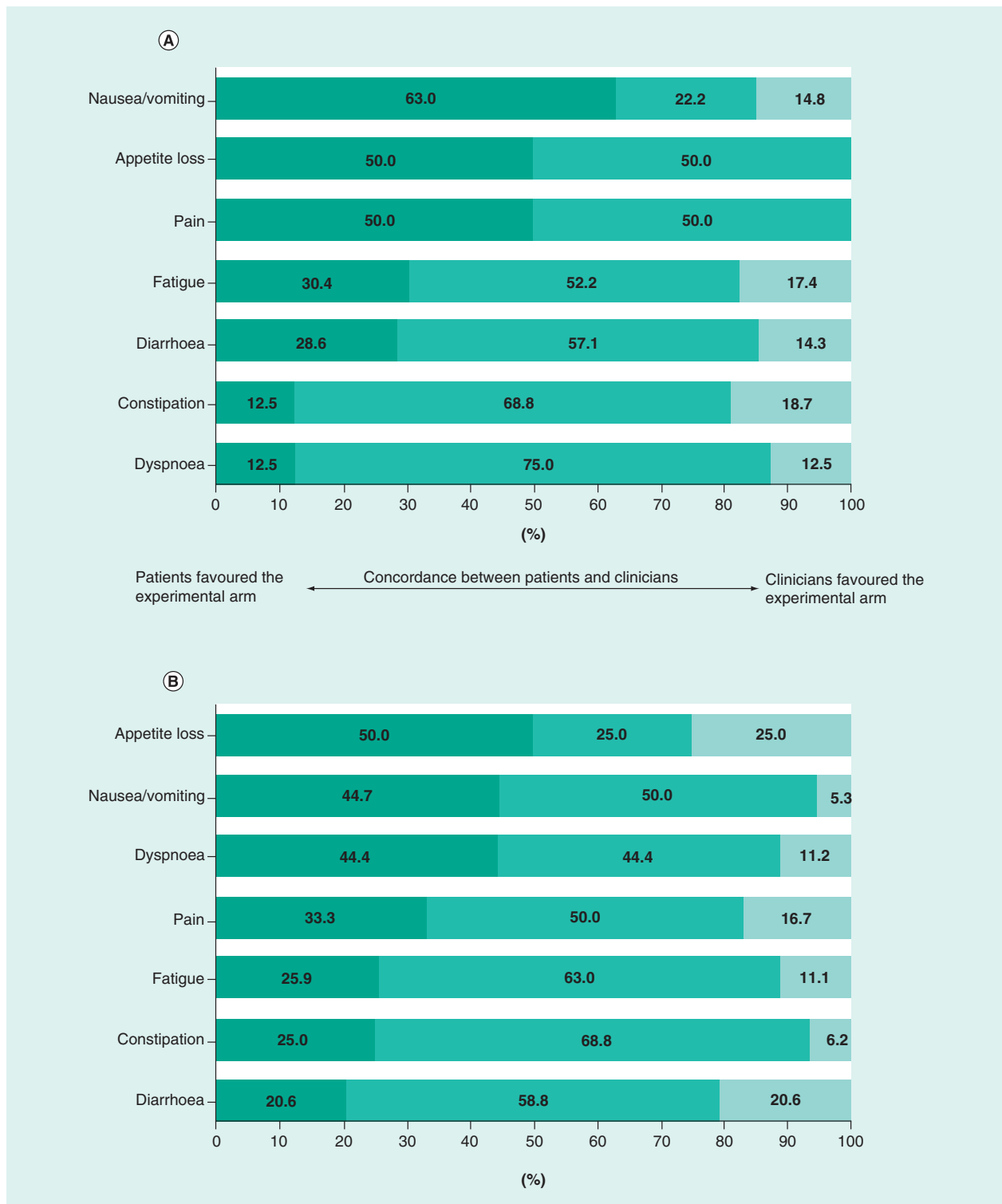


Figure 2. Concordance between patient- and clinician-reported individual symptomatic adverse events. For the purpose of this analysis, patient- reported symptoms and clinician-reported AEs were grouped into three categories. 1. Concordance between patients and clinicians: when both patients and clinicians classified the symptom as equal, or better or worse in the experimental arm; 2. Patients favored the experimental arm: situations where the symptom in the experimental arm was reported by the patients to be better than the clinicians' evaluation (i.e., equal patient reported outcome (PRO)-Worse AE, better PRO-worse AE and better PRO-equal AE); 3. Clinicians favored the experimental arm: situations where the symptom in the experimental arm was reported by clinicians to be better than the patients' evaluation (i.e., equal PRO-better AE, worse PRO-equal AE and worse PRO-better AE).

If nausea and vomiting were reported separately by clinicians, we labeled them as 'better' or 'worse' when both favored the experimental or control arm, or 'equal' when a discordance or no differences in both of them were found. **(A)** Reports the analysis of any grade AEs (n = 34). **(B)** Reports the analysis of severe AEs (n = 45).

AE: Adverse event; PRO: Patient-reported outcome.

considered the global quality of life scales/ratings of PRO instruments into consideration, while we considered all subscales of PRO questionnaires (i.e. functioning, symptoms and global quality of life scales).

Our findings indicate that, in approximately half of the RCTs reviewed, the overall perception of patients (as documented by PRO data) enrolled in the experimental arm, was more favorable than that documented by clinician-reported AEs.

A similar scenario emerged from the single-symptoms analyses. For all symptoms reviewed, regardless of grade severity, clinicians tended to report more AEs in the experimental arm as compared with patient-reported symptoms. In the majority of cases, PRO results indicated similar symptom burden between treatment arms or even less symptom burden in the experimental treatment arm (e.g., for appetite loss and nausea/vomiting). These results further underscore the importance of including a patient-reported symptoms evaluation in RCTs in order to provide important complementary information on drug effects. An implication of this finding, is that of possibly overestimating the overall symptom burden of a given experimental therapy, if PRO data are not available. Accuracy and validity of AEs in cancer clinical trials can be improved by using, for example, the PRO-CTCAE item library, as it includes a comprehensive list of issues aimed at enhancing patient-centeredness of AE reporting in cancer clinical research [3,10]. We thus encourage future clinical researches, to further improve the quality of PRO design and reporting in RCTs, as this might lead to improved validity and accuracy of AE reporting.

Our results also showed that concordance between patient- and clinician-reported AEs tend to be higher in RCTs with PRO as primary endpoint and in open-label trials. Some previous works have shown that methodological quality of PRO reporting tends to be higher in studies with PRO as primary end point, as compared with those with PRO as secondary end point [11,29]. PRO methodological quality might partly explain the different level of concordance by type of PRO end point (primary or secondary). However, this finding deserves further consideration in future studies. Also, the finding that concordance tended to be higher in open-label trials remains to be elucidated in future research.

Our study has several limitations. First, in some of the trials analysed, the authors may have reported only the most prevalent AEs, as well as only selected PRO scales. Therefore, it is possible that our synthesis is not completely exhaustive and comprehensive of all the symptoms assessed across the trials included in our review. However, selective reporting of AE and PRO data from trials is in and of itself problematic and should be discouraged. Second, while AE rating scales typically ask the clinician to report symptoms related to the treatment, most PRO measures, and particularly multidimensional questionnaires require patients to rate their symptoms as related to the disease rather than to treatment. However, the distinction between disease and treatment-related symptoms would be difficult for patients. This may, at least in part, explain discrepancies observed between clinician-reported AE and patient-based symptom burden measures. Third, with the algorithm used to determine whether the PRO data favored the experimental group, the control group or neither group, we used a simple summative approach (i.e., we determined if the majority of the PRO scales favored one treatment arm over the other). This approach does not consider the possibility that the investigators may have identified one or more primary PRO outcomes *a priori*. In fact, most current clinical trial guidelines recommend that investigators select a limited number of PROs to focus on, rather than taking a shotgun approach in which all of the scales in a PRO instrument are analyzed without a prior hypotheses about the expected effects of treatment on symptom burden [30,31]. However, this limitation might be attenuated by the fact that only 15% of our studies reported *a priori* selected PRO outcomes for analysis (data not shown). Lastly, AE and PRO are not typically assessed at the same time intervals during the study period and this may partly explain the discrepancies found in our study but, at the same time, reinforce the need to incorporate these two complementary measures in order to better establish risks and benefits of treatments.

Our study also has strengths. We included a large number of RCTs focusing on the major cancer types (breast, lung, colorectal and prostate). Also, we investigated not only PRO scales reflecting symptoms, but also those assessing functional health outcomes and overall quality of life. Finally, we did not only consider global toxicity ratings, but also examined specific symptoms individually and did so separately for severe versus all grade AEs.

Conclusion

In conclusion, our findings suggest that, often, information obtained via PRO measures and clinician-reported AEs do not favor the same treatment arm in RCT settings.

Executive summary

Background

- There is evidence that patients' perspective may provide a more reliable description of symptoms severity compared with clinicians.
- A systematic review was conducted to investigate the degree of concordance between clinician-reported symptomatic adverse events (AEs) and patient reported outcomes (PROs) in supporting the same treatment arm, in randomized controlled trials (RCTs) published between January 2004 and April 2017 in breast, colorectal, lung and prostate cancer.

Methods

- PROs and AEs were classified as better, equal or worse in the experimental arm compared with the control arm. These outcomes were compared in general and, specifically, for a core set of eight selected AEs.

Results

- A total of 207 RCTs were identified. The analysis showed that only in 35.8% (n = 74) of the studies PROs and AEs were concordant.
- In 104 (50.2%) trials, clinicians report higher levels of AEs in the experimental arms of RCTs with respect to patients. Conversely, in only 29 (14.0%) trials, the PRO results yielded poorer results for the experimental arm, while clinician-reported symptomatic AEs favored that arm.
- The same pattern of results was observed in the single-symptoms analysis.

Conclusion

- Frequently, information obtained via PRO measures and clinician-reported AEs do not favor the same treatment arm in RCT settings.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/full/10.2217/cer-2018-0092

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Stickel A, Goerling U. Quality of life in oncology. *Recent Results Cancer Res.* 210, 163–180 (2018).
2. Keefe DM, Bateman EH. Tumor control versus adverse events with targeted anticancer therapies. *Nat. Rev. Clin. Oncol.* 9(2), 98–109 (2011).
3. Basch E, Dueck AC, Rogak LJ *et al.* Feasibility assessment of patient reporting of symptomatic adverse events in multicenter cancer clinical trials. *JAMA Oncol.* 3(8), 1043–1050 (2017).
4. National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE), version 4.0. (2018) https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03/Archive/CTCAE_4.0_2009-05-29_QuickReference_8.5x11.pdf
5. Fromme EK, Eilers KM, Mori M, Hsieh YC, Beer TM. How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the Quality-of-Life Questionnaire C30. *J. Clin. Oncol.* 22(17), 3485–3490 (2004).
6. Pakhomov SV, Jacobsen SJ, Chute CG, Roger VL. Agreement between patient-reported symptoms and their documentation in the medical record. *Am. J. Manag. Care* 14(8), 530–539 (2008).
7. 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.* 7(11), 903–909 (2006).
8. Di Maio M, Basch E, Bryce J, Perrone F. Patient-reported outcomes in the evaluation of toxicity of anticancer treatments. *Nat. Rev. Clin. Oncol.* 13(5), 319–325 (2016).
- Provides description of the barriers and challenges that should be addressed when considering integration of patient-reported outcomes (PRO) toxicity monitoring in oncology clinical trials.
9. Efficace F, Rosti G, Aaronson N *et al.* Patient- versus physician-reporting of symptoms and health status in chronic myeloid leukemia. *Haematologica* 99(4), 788–793 (2014).

10. Basch E, Reeve BB, Mitchell SA *et al.* Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). *J. Natl Cancer Inst.* 106(9), doi: 10.1093/jnci/dju244 (2014).
- **Describes development of the patient reported outcomes-Common Terminology Criteria for Adverse Events (PRO-CTCAE).**
11. Efficace F, Fayers P, Pusic A *et al.* Quality of patient-reported outcome reporting across cancer randomized controlled trials according to the CONSORT patient-reported outcome extension: A pooled analysis of 557 trials. *Cancer* 121(18), 3335–3342 (2015).
- **Evaluates the completeness of PRO reporting according to the Consolidated Standards of Reporting Trials (CONSORT) PRO extension.**
12. Efficace F, Osoba D, Gotay C, Sprangers M, Coens C, Bottomley A. Has the quality of health-related quality of life reporting in cancer clinical trials improved over time? Towards bridging the gap with clinical decision making. *Ann. Oncol.* 18(4), 775–781 (2007).
13. Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res.* 74(11), 2913–2921 (2014).
14. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J. Clin.* 68(1), 7–30 (2018).
15. Efficace F, Feuerstein M, Fayers P *et al.* Patient-reported outcomes in randomised controlled trials of prostate cancer: methodological quality and impact on clinical decision making. *Eur. Urol.* 66(3), 416–427 (2014).
16. Efficace F, Jacobs M, Pusic A *et al.* Patient-reported outcomes in randomised controlled trials of gynaecological cancers: investigating methodological quality and impact on clinical decision-making. *Eur. J. Cancer* 50(11), 1925–1941 (2014).
17. NCI-PRO-CTCAE Items-English. Item Library Version 1.0. (2018)
<https://healthcaredelivery.cancer.gov/pro-ctcae/pro-ctcae.english.pdf>
18. Aaronson NK, Ahmedzai S, Bergman B *et al.* The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J. Natl Cancer Inst.* 85(5), 365–376 (1993).
19. 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.* 33(8), 910–915 (2015).
- **Compares reporting by patients and physicians of six toxicities within three randomized trials.**
20. Reilly CM, Bruner DW, Mitchell SA *et al.* A literature synthesis of symptom prevalence and severity in persons receiving active cancer treatment. *Support. Care Cancer* 21(6), 1525–1550 (2013).
21. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J. Clin. Epidemiol.* 62(10), 1006–1012 (2009).
22. Schuurhuizen C, Braamse AMJ, Konings I *et al.* Does severe toxicity affect global quality of life in patients with metastatic colorectal cancer during palliative systemic treatment? A systematic review. *Ann. Oncol.* 28(3), 478–486 (2017).
- **Reviews whether severe toxicity affects global quality of life in patients with metastatic colorectal cancer receiving palliative systemic treatment.**
23. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 47(1), 207–214 (1981).
24. Basch E, Jia X, Heller G *et al.* Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. *J. Natl Cancer Inst.* 101(23), 1624–1632 (2009).
25. Atkinson TM, Ryan SJ, Bennett AV *et al.* 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 (2016).
26. Quinten C, Maringwa J, Gotay CC *et al.* Patient self-reports of symptoms and clinician ratings as predictors of overall cancer survival. *J. Natl Cancer Inst.* 103(24), 1851–1858 (2011).
27. Xiao C, Polomano R, Bruner DW. Comparison between patient-reported and clinician-observed symptoms in oncology. *Cancer Nurs.* 36(6), E1–E16 (2013).
28. Gilbert A, Ziegler L, Martland M *et al.* Systematic review of radiation therapy toxicity reporting in randomized controlled trials of rectal cancer: a comparison of patient-reported outcomes and clinician toxicity reporting. *Int. J. Radiat. Oncol. Biol. Phys.* 92(3), 555–567 (2015).
29. Brundage M, Bass B, Davidson J *et al.* Patterns of reporting health-related quality of life outcomes in randomized clinical trials: implications for clinicians and quality of life researchers. *Qual. Life Res.* 20(5), 653–664 (2011).
30. Reeve BB, Mitchell SA, Dueck AC *et al.* Recommended patient-reported core set of symptoms to measure in adult cancer treatment trials. *J. Natl Cancer Inst.* 106(7), doi: 10.1093/jnci/dju129 (2014).
31. Kluetz PG, Slagle A, Papadopoulos EJ *et al.* Focusing on core patient-reported outcomes in cancer clinical trials: symptomatic adverse events, physical function, and disease-related symptoms. *Clin. Cancer Res.* 22(7), 1553–1558 (2016).