Published in final edited form as:

Support Care Cancer. 2016 August ; 24(8): 3669–3676. doi:10.1007/s00520-016-3297-9.

The Association between Clinician-Based Common Terminology Criteria for Adverse Events (CTCAE) and Patient-Reported Outcomes (PRO): A Systematic Review

Thomas M. Atkinson, Ph.D.[1], **Sean J. Ryan, M.A.**[1,2], **Antonia V. Bennett, Ph.D.**[3], **Angela M. Stover, Ph.D.**[3], **Rebecca M. Saracino, MA**[1], **Lauren J. Rogak, MA**[1], **Sarah T. Jewell, MLS**[5], **Konstantina Matsoukas, MLS**[1], **Yuelin Li, PhD**[1], and **Ethan Basch, MD, MPh**[1,3]

[1]Memorial Sloan Kettering Cancer Center, New York, NY, USA

[2]City University of New York, New York, NY, USA

[3]University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[5]Rutgers, The State University of New Jersey, Newark, NJ, USA

Abstract

Purpose—Symptomatic adverse events (AEs) are monitored by clinicians as part of all US-based clinical trials in cancer via the U.S. National Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE) for the purposes of ensuring patient safety. Recently there has been a charge toward capturing the patient perspective for those AEs amenable to patient self-reporting via patient-reported outcomes (PRO). The aim of this review was to summarize the empirically reported association between analogous CTCAE and PRO ratings.

Methods—A systematic literature search was conducted using PubMed, EMBASE, Web of Science, and Cochrane databases through July 2015. From a total of 5,658 articles retrieved, 28 studies met inclusion criteria.

Results—Across studies, patients were of mixed cancer types, including anal, breast, cervical, chronic myeloid leukemia, endometrial, hematological, lung, ovarian, pelvic, pharyngeal, prostate and rectal. Given this mixture, the AEs captured were variable, with many common across studies (e.g., dyspnea, fatigue, nausea, neuropathy, pain, vomiting), as well as several that were disease-specific (e.g., erectile dysfunction, hemoptysis). Overall, the quantified association between CTCAE and PRO ratings fell in the fair to moderate range and had large variation across the majority of studies ($n = 21$).

Conclusions—The range of measures used and symptoms captured varied greatly across the reviewed studies. Regardless of concordance metric employed, reported agreement between CTCAE and PRO ratings was moderate at best. To assist with reconciliation and interpretation of

CORRESPONDING AUTHOR: Thomas M. Atkinson, Ph.D., Department of Psychiatry & Behavioral Sciences, Memorial Sloan Kettering Cancer Center, 641 Lexington Ave., 7th Floor, New York, NY 10022, USA. Phone: 646-888-0089. Fax: 212-888-2584. atkinsot@mskcc.org.

All authors declare no conflicts of interest.

A portion of these results were presented at the 2013 Conference of the International Society for Quality of Life Research in Miami, FL.

these differences toward ultimately improving patient care, an important next step is to explore approaches to integrate PROs with clinician reporting of AEs.

Keywords

Drug Monitoring; Drug-Related Side Effects and Adverse Reactions; Literature Review; Neoplasms; Self Report; Toxicity

INTRODUCTION

Accurate monitoring of symptomatic adverse events (AEs) is essential in clinical trials to assess and ensure patient safety, as well as inform decisions related to treatment and/or continued trial participation [1, 2]. Standard documentation of this information in United States-based cancer clinical trials has relied solely on clinician reporting using an AE rating system known as the Common Terminology Criteria for Adverse Events (CTCAE) [3]. This library of descriptive terms was developed and is maintained by the United States National Cancer Institute (NCI) and consists of 790 items that capture information on discrete events; each is graded by a clinician on a 5-point scale. Of these, 78 items have been identified as objective, observable AEs [4].

Recently, there has been a charge toward the use of patient-reported outcomes (PROs), defined as the unfiltered direct report of a given symptom toxicity by a patient, and considered to be the “gold standard” for the capture of symptomatic AEs [5, 6]. This charge has been led by the release of the 2009 United States Food and Drug Administration (FDA) Guidance for Industry on the Use of PRO Measures in Medical Development to Support Labeling Claims [7], which subsequently led to the NCI initiative to develop a PRO version of the CTCAE (PRO-CTCAE) that will be used in future U.S.-based clinical trials in oncology [4, 8, 9].

As the integration of PRO symptomatic AE ratings into cancer clinical trials becomes commonplace, it is important to understand the degree to which this patient-driven information is complementary to clinician-based CTCAE ratings of observable AEs. The purpose of this review is to summarize the current state of the literature with respect to characterizing the association between clinician-based CTCAE and patient-based PRO ratings.

METHODS

Search Strategy

A systematic search for articles published in peer-reviewed medical journals was conducted with assistance from health sciences librarians using PubMed, EMBASE, Web of Science, and Cochrane Central Register of Controlled Trials (CENTRAL) databases. In PubMed, Medical Subject Headings (MeSH) were used. In Embase, Emtree terms were exploded. The search terms were: ("Neoplasms"[Mesh] OR cancer OR cancers OR cancerous OR tumor OR tumors OR tumour OR tumours OR neoplasm* OR neoplastic OR malignan* OR metastatic OR metastasis OR metastases) AND (physician* OR doctor OR doctors OR clinician* OR "general practitioner" OR "general practitioners" OR provider* OR

specialist* OR nurse* OR staff OR oncologist*) AND (agreement OR disagreement OR reliability OR concordance OR discrepancy OR consistency OR consensus OR "intra-class correlation coefficient" OR "intraclass correlation coefficient" OR "kappa correlation" OR "kappa coefficient" OR “self assessment” OR “physician assessment” OR “provider assessment” OR “self assessed” OR “physician assessed” OR “provider assessed” OR “patient ratings” OR “physician ratings” OR “provider ratings” OR “patients rated” OR “physicians rated” OR “providers rated” OR "Physician's Role"[Mesh]) AND ("adverse symptom event" OR "adverse symptom events" OR "adverse event" OR "adverse events" OR "adverse effect" OR "adverse effects" OR "patient reported outcome" OR "quality of life" OR "Quality of Life"[Mesh] OR ECOG OR “Eastern Cooperative Oncology Group” OR “functional problems” OR distress OR “performance status” OR depression). There were no date or language restrictions; each database was search in its entirety, through July 2015.

Selection Strategy

We deemed studies were eligible for inclusion if they (1) made an original report of a quantitative comparison between analogous CTCAE and PRO ratings, and (2) included participants aged 18 or older.

Screening Process

For the title screening, abstract review, and full-text review stages, two co-authors were randomly assigned to independently review each article for eligibility, with discrepancies arbitrated by a third co-author who was naïve to the article. At the full-text review stage, each assigned co-author completed standardized coding forms to extract the pre-determined data from the potentially eligible articles. References from the included full-text articles were searched to determine whether they should be also considered for inclusion. Study quality was assessed using a modified version of the Downs and Black Study Quality Checklist [10] (i.e., 16 of the original 27 quality indicators most relevant to the types of studies reviewed)[=]. Based on quality, an article was considered to be fit for inclusion if it met at least 33% of the modified Downs and Black Study quality indicators.

For the purposes of this review, cutoffs for all agreement metrics reported, including percentage agreement, Kendall-tau rank correlation (τ), Goodman and Kruskal’s gamma statistic (γ), Pearson correlation (r) , Cohen’s kappa (κ), weighted Cohen’s kappa (κ_w) statistics were defined as: poor (0.00–0.29), moderate (0.30–0.69), and strong (0.70–1.00). For instances where only specificity/sensitivity was reported, the following cutoffs were used: low (0.00–0.29), moderate (0.30–0.69), and high (0.70–1.00).

RESULTS

A total of 7,474 titles were identified through electronic database searching, with 75 additional records found through hand searching. After duplicates were removed, a total of 5,658 articles were retrieved. Following title screening, 908 articles were identified for

[=]Questions 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 16, 17, 18, 19, 20, and 26 from the Downs and Black Study Quality Checklist were used to assess quality.

abstract review with 93.6% agreement between the independent pairs of raters. A total of 251 full-text articles were reviewed. Reasons for article exclusion during the full-text review included: CTCAE not used (n = 161), PROs were not assessed (n = 18), the article was a review or did not include original research findings (n = 18), CTCAE and PRO ratings were not explicitly compared (n = 12), and the article described a non-cancer population (n = 5). Twenty-eight articles met eligibility criteria and were included in this review (Figure 1). Each of these articles possessed at least 47% of the relevant quality indicators from a modified version of the Downs and Black Study Quality Checklist [10], with 21 of the included articles having 75% or more of these quality indicators.

Study Characteristics

Table 1 provides a summary of clinical characteristics and findings from each of the 28 included studies. Patients were of mixed cancer types, including anal, breast, cervical, chronic myeloid leukemia, endometrial, hematological, lung, ovarian, pelvic, pharyngeal, prostate and rectal. Given this, the AEs captured were variable, with many common across studies (e.g., dyspnea, fatigue, nausea, neuropathy, pain, vomiting), as well as several that were disease-specific (e.g., erectile dysfunction, hemoptysis, xerostomia).

Well-validated PRO measures were used to capture patient-based AE ratings in 20 of the included studies. This included disease specific modules of the European Organization for Research and Treatment of Cancer (EORTC) [12–20] and Functional Assessment of Cancer Therapy (FACT) [21–23] instruments, as well as the dermatology-specific Skindex-16, [24, 25] the EuroQol EQ-5D, [26] the Short-Form 36 Health Survey (SF-36) [27], and two recently developed bowel symptom inventories [28, 29]. Patient-adapted versions of the CTCAE were used in four studies [27, 30–32], with visual analog scales used to capture patient-rated AEs in three studies [11, 33, 34].

Association between CTCAE and PRO Ratings

A poor to moderate association between CTCAE and PRO ratings was reported in the majority of included studies (n = 21). Seven such studies reported this as a general finding, with agreement statistics ranging from $\kappa = 0.00$ to $r = 0.74$ [14, 24, 27, 32, 34, 35]. Regardless of how pain was self-reported by patients (i.e., EuroQol EQ-5D or EORTC QLQ-C30), four independent studies reported moderate agreement between PRO and CTCAE ratings (r range = 0.33–0.37; κ range = 0.16–0.29) [16, 17, 20, 26]. Similarly, independent assessments of nausea and vomiting were reported as being poor, as captured by Symptom Tracking and Reporting (STAR) ($\tau = 0.11$ and -0.02) [26] to moderate, as captured by EORTC QLQ-C30 (r range = 0.47–0.48, κ range = 0.41–0.48) [15–17].

Independent studies that made use of the Bowel Problem Scale in cohorts of patients with rectal [28] or anal [29] cancer found poor agreement with clinician ratings for proctitis (κ = 0.22 and 0.11, respectively) and moderate agreement for diarrhea (κ = 0.64 and 0.68, respectively). Additionally, independent studies of neuropathy reported a poor to moderate relationship between PRO and CTCAE reports, as measured by the FACT Gynecologic Oncology Group – Neurotoxicity Module (FACT/GOG-Ntx) (r's range 0.23–0.69) [23, 36]. A study that made use of the Patient Neurotoxicity Questionnaire found that patients more

frequently reported severe neuropathy (30%) as compared to that identified by CTCAE (10%) [37].

Several studies ($n = 4$) reported a mix of poor, moderate and strong associations between CTCAE and PRO ratings. In a study of 400 patients with lung or genitourinary cancer, Basch and colleagues [30] reported a strong percentage of agreement between CTCAE and a modified CTCAE for nausea and vomiting (96% and 90%, respectively), but relatively moderate percent agreement for fatigue (55%) and strong percent agreement for pain (70%). A study of 82 patients with lung cancer led by Brabo [13] observed a moderate correlation between CTCAE and the sore mouth ($r = 0.41$, $p < 0.01$), and alopecia items ($r = 0.52$, $p < 0.01$) from EORTC QLQ-C30. However, strong correlations were observed when comparing CTCAE and the EORTC QLQ-C30 dysphagia ($r = 0.73$, $p < 0.01$) and neuropathy ($r = 0.72$, $p < 0.01$) items. As part of a study of 100 patients with mixed cancer types, Cirillo and colleagues [31] found moderate agreement between CTCAE and a modified patient CTCAE for asthenia ($\kappa = 0.32$, 95% Confidence Interval (CI): 0.17 to 0.44), but strong agreement between clinicians and patients for nausea ($\kappa = 0.74$, 95% CI: 0.64–0.85) and diarrhea ($\kappa = 0.71$, 95% CI: 0.43–0.90). A recently completed study of 66 patients with lung cancer by Tang and colleagues [38] found poor agreement between CTCAE and the Thoracic Symptom Self-Assessment Tool (TSSAT) for nausea ($\kappa_w = 0.07$, 95% CI: −0.16–0.30) and insomnia ($\kappa_w = 0.03$, 95% CI: −0.34–0.39), but strong agreement for hemoptysis ($\kappa_w = 0.71$, 95% CI: 0.35–1.00).

Three studies reported strong agreement between CTCAE and PRO ratings. This included a study of 281 patients with chemotherapy induced peripheral neuropathy, as captured by the EORTC QLQ-CIPN20 [12]; a study of edema, neuropathy and pain in 1031 patients with lung cancer, as captured by the FACT-Taxane module [22]; as well as a study of nausea in 30 patients with mixed cancer types, as measured by a visual analogue scale ($r = 0.88$, $p < 0.001$) [11].

DISCUSSION

As the oncological clinical encounter increasingly incorporates patient-reported data, it is necessary to inform clinicians how this data relates to the standard AE ratings they have been making as part of clinical trials and their routine practice. This review demonstrated that, regardless of which self-report measure was used, the majority of studies that have directly compared CTCAE and PRO ratings report a poor to moderate association between clinician and patient-based AEs, either globally, or by individual symptom (e.g., nausea, neuropathy, pain, vomiting).

This discordance between clinicians and patients provides further evidence that PRO measures provide unique, valuable information that can be complementary to CTCAE ratings. For instance, in a systematic review completed by Gotay and colleagues, it was demonstrated that PRO measures are correlated with survival in cancer clinical trials and provide information above and beyond conventional clinical assessments [39].

Currently, all AEs in clinical trials, including adverse symptoms, are documented by providers at clinic visits to meet federal requirements and facilitate evaluation of new therapies. Patient self-reporting has a relatively minimal role. However, this current practice is problematic, as our previous work has demonstrated that AEs frequently go undetected by clinicians or are reported as less severe than via patient reporting [26, 40]. Additionally, the CTCAE recall period encompasses the entire period since the last clinic visit for a given patient. In many cases, a patient follow-up visit might not occur for 2–4 weeks, which can lead to a loss of information related to AEs that are experienced early in that timeframe.

Recognizing the importance of accurately capturing the patient perspective as complementary information to clinician-based reporting, the NCI has developed a patient language version of the CTCAE (PRO-CTCAE) [4]. PRO-CTCAE is an electronic-based system for patient self-reporting of AEs from the CTCAE that were found to be amenable to patient reporting. This library of 78 symptoms was identified and developed via a process of direct cognitive interviews with patients and extensive quantitative validation in a large multicenter study [8, 9]. PRO-CTCAE utilizes a 7-day recall period and can be completed routinely by patients, and thus may be implemented to systematically capture AEs that may not apparent at the time of clinician-based CTCAE grading.

The present review is limited by a number of factors. The timing of the respective CTCAE and PRO ratings was not explicitly stated in the majority of included articles, though it was implied that these AEs were proximally assessed. Future studies should clearly indicate the amount of time that occurs between clinician- and patient-based AE reporting to eliminate the possibility that the level of agreement between these rating sources is being influenced by passage of time (e.g., separate visits, between visits versus next or previous clinic visit). Additionally, a meta-analysis would have been optimal to best characterize the relationship between CTCAE and PRO ratings. Unfortunately, given the variety of PRO measures used and numerous instances where study authors did not provide detailed statistical coefficient for each AE captured, a meta-analysis was infeasible.

To better assist clinicians with understanding the relationship between their CTCAE ratings and PRO assessments, future work should seek to systematically equate CTCAE and PRO ratings. Future approaches that aim to integrate PROs with clinician reporting of AEs, especially those of the CTCAE and PRO-CTCAE, would improve our understanding of patient and clinician ratings and assist clinicians and policy makers with the interpretation of clinical trial results.

Acknowledgments

This project was supported by a National Institutes of Health Research Training Grant (T32 CA009461-25); as well as a National Institutes of Health Support Grant (P30 CA08748-48), which provides partial support for the Behavioral Research Methods Core Facility used in conducting this investigation. The authors wish to thank Jan-Samuel Wagner for his early contributions to the project, as well as Dr. Jamie Ostroff for her support of this work.

REFERENCES

1. Basch E. The missing voice of patients in drug-safety reporting. N Engl J Med. 2010 Mar 11.362:865–869. [PubMed: 20220181]

2. Basch E. New frontiers in patient-reported outcomes: adverse event reporting, comparative effectiveness, and quality assessment. Annual review of medicine. 2014; 65:307–317.

3. Vyzula RRS. Lung cancer in patients with HIV-infection. Lung Cancer. 1996; 15:325–339. [PubMed: 8959678]

4. 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. 2014 Sep.106

5. Basch E. Beyond the FDA PRO Guidance: Steps toward Integrating Meaningful Patient-Reported Outcomes into Regulatory Trials and US Drug Labels. Value in Health. 2012 May.15:401–403. [PubMed: 22583448]

6. Trotti A, Colevas AD, Setser A, Basch E. Patient-reported outcomes and the evolution of adverse event reporting in oncology. J Clin Oncol. 2007; 25:5121–5127. [PubMed: 17991931]

7. US Department of Health and Human Services. [Accessed January 5, 2016] Guidance for Industry. Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009 Dec. Available from: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf

8. Hay JL, Atkinson TM, Reeve BB, et al. Cognitive interviewing of the US National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). Qual Life Res. 2014 Feb.23:257–269. [PubMed: 23868457]

9. Dueck AC, Mendoza TR, Mitchell SA, et al. Validity and Reliability of the U.S. National Cancer Institute's Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). JAMA Oncol. 2015

10. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. Journal of epidemiology and community health. 1998 Jun.52:377–384. [PubMed: 9764259]

11. Mandala M, Cremonesi M, Rocca A, et al. Midazolam for acute emesis refractory to dexamethasone and granisetron after highly emetogenic chemotherapy: a phase II study. Support Care Cancer. 2005 Jun.13:375–380. [PubMed: 15668754]

12. Alberti P, Rossi E, Cornblath DR, et al. Physician-assessed and patient-reported outcome measures in chemotherapy-induced sensory peripheral neurotoxicity: two sides of the same coin. Ann Oncol. 2014 Jan.25:257–264. [PubMed: 24256846]

13. Brabo EP, Paschoal ME, Biasoli I, et al. Brazilian version of the QLQ-LC13 lung cancer module of the European Organization for Research and Treatment of Cancer: preliminary reliability and validity report. Qual Life Res. 2006 Nov.15:1519–1524. [PubMed: 16960750]

14. Christodoulou M, McCloskey P, Stones N, et al. Investigation of a Patient Reported Outcome tool to assess radiotherapy-related toxicity prospectively in patients with lung cancer. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology. 2014 Aug. 112:244–249. [PubMed: 25107555]

15. Di Maio M, Gallo C, Leighl NB, et al. Symptomatic toxicities experienced during anticancer treatment: agreement between patient and physician reporting in three randomized trials. J Clin Oncol. 2015 Mar 10.33:910–915. [PubMed: 25624439]

16. 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. 2004 Sep 1.22:3485–3490. [PubMed: 15337796]

17. Greimel ER, Bjelic-Radisic V, Pfisterer J, et al. Toxicity and quality of life outcomes in ovarian cancer patients participating in randomized controlled trials. Support Care Cancer. 2011; 19:1421–1427. [PubMed: 20694564]

18. Jensen K, Lambertsen K, Torkov P, et al. Patient assessed symptoms are poor predictors of objective findings. Results from a cross sectional study in patients treated with radiotherapy for pharyngeal cancer. Acta oncologica. 2007; 46:1159–1168. [PubMed: 17851855]

19. Kirchheiner K, Nout R, Lindegaard J, et al. Do clinicians and patients agree regarding symptoms? A comparison after definitive radiochemotherapy in 223 uterine cervical cancer patients.

Strahlentherapie und Onkologie : Organ der Deutschen Rontgengesellschaft … [et al]. 2012 Oct. 188:933–939.

20. 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. 2011 Dec.103:1851–1858. [PubMed: 22157640]
21. Cella D, Huang H, Homesley HD, et al. Patient-reported peripheral neuropathy of doxorubicin and cisplatin with and without paclitaxel in the treatment of advanced endometrial cancer: Results from GOG 184. Gynecol Oncol. 2010 Dec.119:538–542. [PubMed: 20863554]
22. Hirsh V, Okamoto I, Hon JK, et al. Patient-reported neuropathy and taxane-associated symptoms in a phase 3 trial of nab-paclitaxel plus carboplatin versus solvent-based paclitaxel plus carboplatin for advanced non-small-cell lung cancer. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2014 Jan.9:83–90.
23. Shimozuma K, Ohashi Y, Takeuchi A, et al. Feasibility and validity of the Patient Neurotoxicity Questionnaire during taxane chemotherapy in a phase III randomized trial in patients with breast cancer: N-SAS BC 02. Support Care Cancer. 2009 Dec.17:1483–1491. [PubMed: 19330359]
24. Atherton PJ, Burger KN, Loprinzi CL, et al. Using the Skindex-16 and Common Terminology Criteria for Adverse Events to assess rash symptoms: results of a pooled-analysis (N0993). Support Care Cancer. 2012 Aug.20:1729–1735. [PubMed: 21922203]
25. Neben-Wittich MA, Atherton PJ, Schwartz DJ, et al. Comparison of provider-assessed and patient-reported outcome measures of acute skin toxicity during a Phase III trial of mometasone cream versus placebo during breast radiotherapy: the North Central Cancer Treatment Group (N06C4). International journal of radiation oncology, biology, physics. 2011 Oct 1.81:397–402.
26. 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 Dec 2.101:1624–1632. [PubMed: 19920223]
27. Efficace F, Rosti G, Aaronson N, et al. Patient- versus physician-reporting of symptoms and health status in chronic myeloid leukemia. Haematologica. 2014 Apr.99:788–793. [PubMed: 24241488]
28. Flores LT, Bennett AV, Law EB, et al. Patient-Reported Outcomes vs. Clinician Symptom Reporting During Chemoradiation for Rectal Cancer. Gastrointestinal cancer research : GCR. 2012 Jul.5:119–124. [PubMed: 23077685]
29. Tom A, Bennett A, Rothenstein D, et al. Prevalence of Patient-Reported Gastrointestinal Symptoms and Concordance With Clinician Toxicity Assessments in Radiation Therapy for Anal Cancer. International Journal of Radiation Oncology*Biology*Physics. 2013; 87:S570–S571.
30. 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; 7:903–909. [PubMed: 17081915]
31. Cirillo M, Venturini M, Ciccarelli L, et al. 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 Dec.20:1929–1935. [PubMed: 19622510]
32. Gravis G, Marino P, Joly F, et al. Patients' self-assessment versus investigators' evaluation in a phase III trial in non-castrate metastatic prostate cancer (GETUG-AFU 15). Eur J Cancer. 2014 Mar.50:953–962. [PubMed: 24424105]
33. Cella D, Pulliam J, Fuchs H, et al. Evaluation of pain associated with oral mucositis during the acute period after administration of high-dose chemotherapy. Cancer. 2003 Jul 15.98:406–412. [PubMed: 12872363]
34. Steinsvik EA, Fossa SD, Axcrona K, et al. Do perceptions of adverse events differ between patients and physicians? Findings from a randomized, controlled trial of radical treatment for prostate cancer. The Journal of urology. 2010 Aug.184:525–531. [PubMed: 20620412]
35. Franklin HR, Simonetti GP, Dubbelman AC, et al. Toxicity grading systems. A comparison between the WHO scoring system and the Common Toxicity Criteria when used for nausea and vomiting. Ann Oncol. 1994 Feb.5:113–117. [PubMed: 8186153]

36. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. J Clin Epidemiol. 2010 Nov.63:1179–1194. [PubMed: 20685078]

37. Bennett BK, Park SB, Lin CS, et al. Impact of oxaliplatin-induced neuropathy: a patient perspective. Support Care Cancer. 2012 Nov.20:2959–2967. [PubMed: 22426503]

38. Tang B, Giuliani M, Le LW, et al. Capturing acute toxicity data during lung radiotherapy by using a patient-reported assessment tool. Clin Lung Cancer. 2013 Mar.14:108–112. [PubMed: 22885348]

39. Gotay CC, Kawamoto CT, Bottomley A, Efficace F. The prognostic significance of patient-reported outcomes in cancer clinical trials. J Clin Oncol. 2008 Mar 10.26:1355–1363. [PubMed: 18227528]

40. Atkinson TM, Li Y, Coffey CW, et al. Reliability of adverse symptom event reporting by clinicians. Qual Life Res. 2012 Sep.21:1159–1164. [PubMed: 21984468]

41. Park JH, Kim YS, Park J, et al. Incidence and dose-volume analysis of acute bladder toxicity following pelvic radiotherapy. Tumori. 2014 Mar-Apr;100:195–200. [PubMed: 24852865]

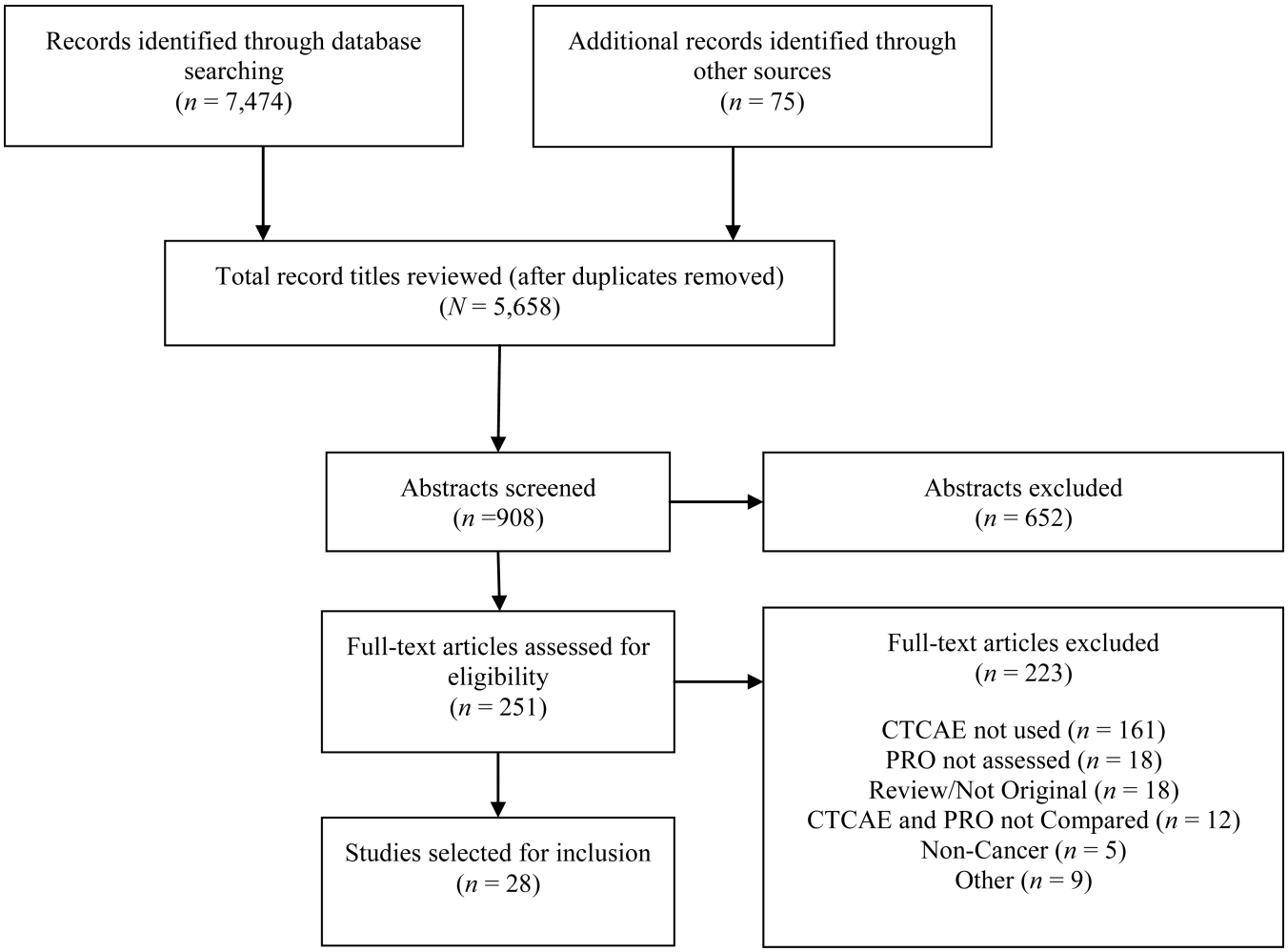

Figure 1.
PRISMA Flow Chart

Table 1

Clinical Characteristics and Summarized Findings of Included Studies by Primary Patient-Reported Outcomes Measure ($N = 28$)

Study	*n*	Disease	Adverse Events	Patient measure	Finding
European Organization for Research and Treatment of Cancer Quality of Life (EORTC)					
Alberti, 2014 [12]	281	Mixed	Chemotherapy-induced peripheral neuropathy	EORTC QLQ-CIPN20	Strong agreement between CTCAE and EORTC QLQ-CIPN20 scores, except for patients with normal results.
Brabo, 2006 [13]	82	Lung	Alopecia, Dysphagia, Neuropathy, Sore Mouth	EORTC QLQ-LC13	Strong correlation for dysphagia and neuropathy (r = 0.73 and 0.72, $p < 0.01$). Moderate correlation for sore mouth and alopecia (r = 0.41 and 0.52, $p < 0.01$).
Di Maio, 2015 [15]	1090	Breast & Lung	Anorexia, Diarrhea, Vomiting	EORTC QLQ-C30, QLQ-LC13, & QLQ-BR23	Moderate agreement for vomiting and diarrhea (κ = 0.41, 95% CI: 0.34–0.47 and 0.45, 95% CI: 0.39–0.50) but poor agreement on anorexia (κ = 0.15, 95% CI: 0.12–0.19).
Fromme, 2004 [16]	37	Prostate	Fatigue, Nausea, Pain, Vomiting	EORTC QLQ-C30	Moderate agreement for nausea/vomiting (κ= 0.48, 95% CI: 0.13–0.83). Poor agreement for pain (κ = 0.16, 95% CI: −0.17–0.48). Poor negative agreement on fatigue (κ = −0.29, 95% CI: −0.68–0.09).
Greimel, 2011 [17]	2110	Ovarian	Constipation, Myalgia, Nausea, Pain, Vomiting	EORTC QLQ-C30	Nausea, vomiting, and constipation moderately correlated ($r = 0.47$, 0.49, and 0.48$p < 0.01$). Pain and myalgia had poor to moderate correlations ($r = 0.33$ and 0.22, $p < 0.01$).
Jensen, 2007 [18]	35	Pharyngeal	Xerostomia	EORTC QLQ-C30, EORTC H&N35	Patient reported dry mouth had moderate sensitivity (0.68, 95% CI: 0.43–0.87) and a high specificity (0.81, 95% CI:0.54–0.96) with the CTCAE representing the objective standard.
Kirchheiner, 2012 [19]	223	Cervical	Edema, Hot Flashes, Insomnia	EORTC QLQ-C30, EORTC CX25	CTCAE had moderate sensitivity (0.49 for insomnia to 0.73 for hot flashes) and high specificity (0.82 insomnia, 0.93 for edema) when compared to EORTC QLQ-C30 as the gold standard.
Quinten, 2011 [20]	2270	Mixed	Constipation, Diarrhea, Emesis, Fatigue, Nausea, Pain	EORTC QLQ-C30	Poor agreement was found for pain (κ = 0.29, 95% CI: 0.26–0.33), as well as for the remaining symptoms (κ= 0.07–0.22).
Patient Adapted CTCAE					
Basch, 2006 [30]	400	Lung & Genitourinary	Fatigue, Nausea, Pain, Vomiting	Modified CTCAE	Strong percent agreement for nausea and vomiting (96% and 90%). Moderate percent agreement for fatigue/pain (55% and 70%) with patients reporting symptoms as more severe.
Basch, 2009 [26]	163	Lung	Fatigue, Nausea, Pain, Vomiting	STAR EuroQol EQ-5D	Poor to moderate concordance for pain and fatigue (τ = 0.37 and 0.29). Poor concordance

Study	*n*	Disease	Adverse Events	Patient measure	Finding
					for nausea and vomiting (τ = 0.11 and −0.02).
Bennett, 2012 [37]	22	Colorectal	Neuropathy	Patient Neurotoxicity Questionnaire	30% of patient self-reported severe neuropathy compared to 10% identified by clinicians.
Christodoulou, 2014 [14]	70	Lung	Anorexia, Anxiety, Chest Pain, Cough, Depression, Dyspnea, Dysphagia, Fatigue, Hemoptysis	Patient adapted CTCAE, EORTC QLQ-C30, QLQ-LC13, HADS	Agreement between patients' and clinicians' reported toxicity was poor to moderate for most items at end of radiotherapy (κ_w range 0.21–0.61) and at follow-up (κ_w range 0.12–0.61). The majority of the discrepancies were within one grade of toxicity.
Cirillo, 2009 [31]	100	Mixed	Asthenia, Diarrhea, Nausea	Modified CTCAE	Strong agreement for nausea and diarrhea (κ = 0.74, 95% CI: 0.64–0.85 and 0.71, 95% CI: 0.43–0.90). Moderate for asthenia (κ = 0.32, 95% CI: 0.17–0.44).
Efficace, 2014 [27]	422	Chronic Myeloid Leukemia	Abdominal Discomfort, Diarrhea, Edema, Fatigue, Headache, Muscle cramps, Musculoskeletal Pain, Nausea, Skin Problems	Patient adapted CTCAE, SF-36	Agreement between clinicians and patients ranged from 34% (muscle cramps) to 66% (nausea). Patients reported higher severity more often than their physicians for all symptoms.
Gravis et al., 2014 [32]	385	Prostate	Fatigue, Hot Flashes, Sexual Dysfunction, Weight Gain/Loss	Patient adapted CTCAE	Poor concordance between patients and physicians across all symptoms at both 3 months (κ range 0.03–0.21) and 6 months (κ range 0.00–0.21).
Functional Assessment of Cancer Therapy (FACT)					
Cella, 2010 [21]	467	Endometrial	Neuropathy	FACT/GOG-Ntx	Moderate correlation for neuropathy (r = 0.69).
Hirsh, 2014 [22]	1031	Lung	Edema, Neuropathy, Pain	FACT-Taxane	Strong agreement between FACT-Taxane and CTCAE.
Shimozuma, 2009 [23]	300	Breast	Neuropathy	FACT/GOG-Ntx	Poor to moderate correlation for sensory and motor neuropathy scores (r = 0.45 and 0.23).
Visual Analog Scale (VAS)					
Cella, 2003 [33]	323	Hematological	Dysphagia, Mouth Pain	VAS	Moderate correlation for and peak dysphagia and peak mouth pain (r = 0.67 and 0.69).
Mandala, 2005 [11]	30	Mixed	Nausea	VAS	Strong correlation for nausea (r = 0.88, p < 0.000001).
Steinsvik, 2010 [34]	333	Prostate	Bowel Problems, Erectile Dysfunction, Incontinence, Nausea	VAS	Observed correlations were poor to moderate for all AEs (r = 0.19–0.74).
Skindex-16					
Atherton, 2012 [24]	412	Mixed	Dermatology	Skindex-16	Moderate agreement for dermatological AE (r =

Support Care Cancer. Author manuscript; available in PMC 2017 August 01.

Study	*n*	Disease	Adverse Events	Patient measure	Finding
					0.55–0.62).
Neben-Wittich, 2011 [25]	166	Breast	Dermatology (Pruritus, Dermatitis, and Striae)	Skindex-16, STAT	Strong to moderate correlation for pruritus and dermatitis (r = 0.74 and 0.63), but poor correlation for striae (r = 0.15).
AE-Specific Questionnaire					
Flores, 2012 [28]	199	Rectal	Diarrhea, Proctitis	Bowel Problem Scale	Moderate agreement for diarrhea (κ = 0.64) but poor agreement for proctitis (κ = 0.22).
Franklin, 1994 [35]	74	Mixed	Nausea, Vomiting	Anti-Emetic Survey	Physician assessment had moderate agreement for nausea and vomiting (γ = 0.32 and 0.67).
Park, 2014 [41]	92	Pelvic	Incontinence, Nocturia, Pain, Urgency	IPSS, OABSS	For patients with CTCAE grade 2 toxicity, 70% were categorized with moderate to severe symptoms based on IPSS, compared with 41% based on OABSS (p = 0.03).
Tang, 2013 [38]	66	Lung	Hemoptysis, Insomnia, Nausea	TSSAT	Strong agreement for hemoptysis (κ_w = 0.71, 95% CI: 0.35–1.00), but poor agreement for nausea and insomnia (κ_w = 0.07, 95% CI: −0.16–0.30 and κ_w = 0.03, 95% CI: −0.34–0.39)
Tom, 2013 [29]	78	Anal	Diarrhea, Proctitis	Bowel Problem Scale	Poor agreement for proctitis (κ = 0.11), but moderate agreement for diarrhea (κ = 0.68).

Note: AE indicates Adverse event; CI, Confidence Interval; κ, Cohen's kappa; CTCAE, κ_W, weighted Cohen's kappa; τ, Kendall-tau rank correlation; γ, Goodman and Kruskal's gamma statistic; Common Terminology Criteria for Adverse Events; EORTC QLQ-C30, European Organization for Research and Treatment of Cancer Quality of Life Questionnaire – Core 30; EORTC QLQ-BR23, Quality of Life Questionnaire – Breast Cancer Module; EORTC QLQ-CIPN20, Quality of Life Questionnaire – Chemotherapy-Induced Peripheral Neuropathy Module; EORTC CX25, Quality of Life Questionnaire – Cervical Cancer Module; EORTC H&N35, Head and Neck Module; EORTC QLQ-LC13, Quality of Life Questionnaire – Lung Module; FACT, Functional Assessment of Cancer Therapy; FACT/GOG-Ntx, FACT Gynecologic Oncology Group – Neurotoxicity Module; FACT-Taxane, FACT Taxane Module; HADS, Hospital Anxiety and Depression Scale; IPSS, International Prostate Symptom Score; OABSS, Overactive Bladder Symptom Score; SF-36, Short-Form 36 Health Survey; STAT, Skin Toxicity Assessment Tool; STAR, Symptom Tracking and Reporting; TSSAT, Thoracic Symptom Self-Assessment Tool; VAS, Visual Analog Scale.