Single-arm studies involving patient-reported outcome data in oncology

SISAQOL-IMI Work Package 3; Calvert, Melanie; Cruz Rivera, Samantha; Aiyegbusi, Olalekan Lee

DOI:
10.1016/S1470-2045(23)00110-9

TITLE: Single-arm studies involving patient-reported outcome data in oncology: a literature review on current practice

AUTHORS: Limin Liu[1], Jungyeon Choi[2], Jammbe Z. Musoro[3], Willi Sauerbrei[4], Cecilie Delphin Amdal[5,6], Ahu Alanya[3], Yolanda Barbachano[7], Joseph C. Cappelleri[8], Ragnhild Sørum Falk[5], Mallorie H. Fiero[9], Antoine Regnault[10], Jaap C. Reijneveld[11], Rickard Sandin[12], Doranne Thomassen[13], Satrajit Roychoudhury[8], Els Goetghebeur[1], Saskia le Cessie[1,2,13]; on behalf of SISAQOL-IMI Work Package 3

AFFILIATIONS:

[1]Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Belgium. (L Liu PhD, Prof E Goetghebeur PhD, Prof S le Cessie PhD)

[2]Department of Clinical Epidemiology, Leiden University Medical Center, the Netherlands. (J Choi MSc, Prof S le Cessie PhD)

[3]European Organisation for Research and Treatment of Cancer (EORTC) Headquarters, Brussels, Belgium (J Z Musoro PhD, A Alanya PhD)

[4]Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center - University of Freiburg, Germany (Prof W Sauerbrei PhD)

[5]Research Support Services, Oslo University Hospital, Norway (C D Amdal MD PhD, R S Falk PhD)

[6]Department of Oncology, Oslo University Hospital, Norway (C D Amdal MD PhD)

[7]Medicines and Healthcare products Regulatory Agency, UK (Y Barbachano PhD*)

[8]Pfizer Inc, US (J C Cappelleri PhD, S Roychoudhury PhD)

[9]US Food and Drug Administration, US (M H Fiero PhD)

[10]Modus Outcomes, France (A Regnault PhD)

[11]VU University Medical Center, Department of Neurology & Brain Tumor Center, Amsterdam, Netherlands (Prof J C Reijneveld PhD)

[12]Pfizer Ltd, UK (R Sandin PhD)

[13]Department of Biomedical Data Sciences, Leiden University Medical Center, the Netherlands. (D Thomassen MSc, Prof S le Cessie PhD)

*Moved to another institution

Corresponding Author
Dr Limin Liu, Department of Applied Mathematics, Computer Science and Statistics, Campus Sterre, S9, Krijgslaan 281, 9000 Gent, Belgium; Tel: +32 (0) 9 264 96 87; limin.liu@ugent.be

Total number of words: 3694

Summary

Patient-reported outcomes (PROs) are increasingly used in single-arm cancer studies. We reviewed 60 recent publications of single-arm studies of cancer treatment involving PRO data for current practice on design, analysis, reporting, and interpretation. We further examined their handling of potential bias and how they informed decision-making. Most studies (97%) analyzed PROs without stating a predefined research hypothesis. Thirteen studies (22%) used a PRO as a (co)primary endpoint. Definitions of PRO objectives, study population, endpoints, and strategies of handling missing data varied widely. Twenty-three studies (38%) compared the PRO data to external information, most often by using a clinically important difference value; one study used a historical control group. Appropriateness of methods to handle missingness and intercurrent events including death were seldom discussed. Most studies (85%) concluded that PRO results supported treatment. Conducting and reporting of PROs in cancer single-arm studies lacks standards, and a critical discussion of statistical methods and possible biases. These findings will guide the Setting International Standards in Analysing Patient-Reported Outcomes and Quality of Life Data in Cancer Clinical Trials-Innovative Medicines Initiative (SISAQOL-IMI) in developing recommendations for the use of PRO-measures in single arm studies.

Funding

The SISAQOL-IMI project has received funding from the Innovative Medicines Initiative (IMI) 2 Joint Undertaking under grant agreement No. 945052. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation program and the European Federation of Pharmaceutical Industries and Associations (EFPIA).

1. Introduction

The patient's voice is increasingly heard in the evaluation of risks and benefits of cancer therapies in clinical trials.[1] Legislative authorities, therefore, advocate incorporating patient-reported outcome (PRO) data in cancer research.[2–4] In recent years, the majority (>70%) of oncology indications approved by the US Food & Drug Administration (FDA) or the European Medicines Agency (EMA) included PRO data in the regulatory submission.[5,6]

While randomized controlled trials (RCTs) have been considered the gold standard to demonstrate treatment efficacy since the 1960s, they are not always feasible or ethical to conduct.[7] Also in pragmatic research settings, single-arm studies may be preferred over RCTs because they can better reflect real-life clinical practice. Patients may not wish to be randomized in a RCT or fail the stricter inclusion criteria, leading to more selective patient groups in the tightly controlled setting. This limits the external validity of RCTs and implementation of findings, especially for studies of supportive and palliative settings, with patient-centered treatment goals, and patient-tailored interventions. Therefore, regulatory authorities also accept results from single-arm studies under certain circumstances, e.g. in rare diseases, specific disease subtypes with small patient numbers, where no effective standard treatment exists, or to support expedited development using accelerated approval, conditional approval or other regulatory approval pathways.[8–13] In recent decades, oncology treatment approval and reimbursement submissions based on evidence from single-arm trials have increased .[14,15] Between 2014 and 2019, 187 trials led to 176 approvals for novel anticancer drugs by the FDA, of which 34% were single-arm clinical trials.[16] A primary concern of single-arm studies with PRO endpoints is their inability to control for bias. It is therefore difficult to assess whether any change in PRO is due to treatment, or to confounding factors such as the natural disease course, the absence of blinding of treatment, regression to the mean, or to the fact that patients' responses may shift over time.[5,17,18] Furthermore, missing responses which may be informative of the patient health status and intercurrent events such as treatment discontinuation or the start of concomitant medication must be considered.

Much attention has been given recently to translating a research question into an estimand; a well-defined target of estimation. The International Council for Harmonization (ICH) has published an addendum on estimands to their guideline E9 on Statistical Considerations for Clinical Trials.[19] The estimand requires explicit specification of the study population, treatment (and comparator if applicable), outcomes, the handling of intercurrent events and population level summary.[20] It emphasizes that intercurrent events (events that happen after

randomization and affect the existence or interpretation of the measurements) should be addressed in the research question. In particular, death requires attention, because PRO data after death do not exist and cannot be defined, since death is a terminal event.

Recommendations for including PROs in clinical trials have not addressed the challenges with single-arm studies.[3,4,21,22] Therefore, to investigate the feasibility of, and develop recommendations for, the use of PRO measures for non-RCTs in cancer studies with a specific focus on single-arm studies are some of the essential aims of the Setting International Standards in Analysing Patient-Reported Outcomes and Quality of Life Endpoints in Cancer Clinical Trials - Innovative Medicines Initiative (SISAQOL-IMI) Consortium. This international multi-stakeholder consortium brings together PRO experts, statisticians, and representatives from academic institutions, pharmaceutical companies, cancer institutes, regulators, and patient organizations. The consortium was formed in 2021 to develop consensus-based recommendations for analyzing and interpreting data on health-related quality of life and other PROs in cancer clinical trials. It builds on the work of the previous SISAQOL consortium, which published recommendations for PRO analysis for cancer randomized controlled trials (RCTs) in 2020.[23,24]

This literature review aimed to describe the current state of published single-arm oncology studies involving PROs, serving as a basis for the development of guidance for the design, analysis, and reporting of single-arm cancer studies with PROs. To our knowledge, no such literature review nor similar work on single-arm studies of PROs exists. In this review, we evaluated the design, analysis, reporting, and interpretation of these studies, and considered how authors dealt with potential biases. Moreover, we describe whether the conclusion on the treatment effect on PROs found evidence in favor or against treatment.

2. Methods

Search strategy and selection criterion

We followed the PRISMA guidelines for Systematic Reviews and Meta-Analyses.[25] First, we searched PubMed, Web of Science, Cochrane, and EMBASE for single-arm studies in oncology that used PROs. Search queries were built with the help of a medical librarian (Appendix: page 2). The search was conducted on February 19th, 2021.

As we intended to describe current practice, we selected the most recent 60 eligible papers for full paper review. This number was deemed sufficient to gain a general appreciation of recent practice in published single-arm oncology studies, and was feasible to review within the allocated time within the SISAQOL-IMI project: the SISAQOL-IMI consortium needed its

results as input for the formulation of subsequent recommendations. Starting with the most recent search results, the references were traced back in time until 60 publications were identified that met the following inclusion criteria: i) single-arm studies, i.e., uncontrolled, non-randomized studies; ii) anti-cancer treatment (e.g., marketable drugs or devices, radiotherapy, or surgery); iii) cancer patients; iv) PROs as endpoints; v) PRO sample size larger or equal to 10; and vi) written in English.

Systematic reviews, meta-analyses, study protocols, purely methodological publications, response letters/brief reports/ethical guidelines, animal studies, qualitative studies, multiple single-arm studies under one protocol (such as basket and umbrella trials), interim analyses, and studies that evaluated non-medical treatment (such as psychological, supportive, and nutritional interventions) were excluded. When there were multiple papers on one study, we selected the most PRO-focused paper. Two reviewers independently performed the abstract screening (LL and JC). Disagreements were initially handled by discussion between these two reviewers. Remaining disagreements were resolved during consensus meetings involving a third reviewer (SlC or EG).

Data collection

In the full-text review, we collected the following information from each cancer study: i) general information, such as the type and stage of disease, and type of treatment; ii) study objectives; iii) information on the PRO measure, study and PRO population, collection approach (e.g., paper, digital, interview), assessment time points, how PROs were summarized (e.g., mean magnitude of change, worst score over a certain period), and intercurrent events; iv) design and PRO analysis considerations including the use of clinically important differences and comparison data, and statistical methods used; v) reporting of PROs and PRO analysis; vi) addressing potential bias in the PRO study results; and vii) interpretation of the PRO results. The complete code book and a glossary of definitions used for coding can be found in Appendix (page 4).

Full-text review of selected articles were independently conducted by at least two reviewers. All papers were reviewed by the same statistician (LL) and additional reviewers who were statisticians, epidemiologists, regulators, clinicians, and PRO experts. Disagreements between reviewers were solved following the same procedures used for screening the abstracts.

Descriptive analyses were presented as frequencies and percentages for categorical data and median and range for continuous data.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

3. Results

Abstracts were screened in reverse chronological order per year and in 2018 the required 60 papers were obtained. Out of the 364 screened abstracts published since 2018, 279 were excluded (see Figure 1 for details). In the full text review of the 60 selected studies, two articles had PRO sample size below 10, and were replaced by two other studies from 2018 with cancer types (skin and liver) that were not yet covered in the review. Details of the individual papers are given in Appendix (page 13).

General characteristics of the reviewed articles

Breast cancer was most prevalent (n=11), followed by lung cancer (n=7) and studies which considered more than one cancer type (n=7), for instance, cancer patients who were referred for radiotherapy for bone metastases (Table 1).

The majority of studies considered advanced-stage or metastatic cancer (n=36) (Table 1). The primary treatment aim was more often to prolong survival (n=44) than to reduce disease symptoms (n=7) or side effects of curative treatment (n=9). More than half of the studies were investigating oncology drugs (n=39, including five trials studying drugs in combination with radiotherapy), while fewer studies examined other treatment types such as medical device (n=1) or surgical procedures (n=5). The median sample size for the PRO analysis was 43 (range 10 to 991), including 12 studies with PRO-sample size larger than 100.

PRO study objectives and analysis

Table 2 gives an overview of the PRO study objectives and their associated analytic issues. Among the 13 studies that used PROs as a primary outcome, PROs were the sole primary outcome in 11 studies. In these 11 studies, pain score (n=6) and health-related quality of life (n=3) were most commonly measured, and the aim of the treatment was most often to reduce disease symptoms (n=5) or to reduce side effects of treatment (n=4). Five studies explicitly mentioned that (some of) their PROs were exploratory outcomes, studied in parallel to the defined primary and secondary outcomes. For six studies, it was unclear whether PROs were primary or secondary outcomes.

Only two studies stated a clear predefined PRO research hypothesis which they aimed to examine in the study. In most cases (n=58), PRO study aims were described without prior

assumptions or hypotheses stated. PROs were used to assess either tolerability (n=20) or benefit (n=23) of treatment, or both (n=7).

In most studies, the target population used for the primary analysis and PRO analysis coincided, but in four studies these populations differed. In one study, for example, the main target population was leukemia patients between 3 and 21 years old while the PRO analysis considered patients older than 8 years. Nearly half of the studies limited the PRO analysis to patients with a baseline and at least one PRO assessment (n=27). Seventeen studies included patients with at least a baseline PRO assessment and 15 studies used patients with at least one PRO assessment during follow-up.

The frequency of PRO assessment ranged from 2 to 26 (median 4, Q1=3, Q3=6). In most cases, PROs were collected at baseline, during, and after treatment (n=26), followed by being measured at baseline and while on treatment (n=19). The rationale behind measuring PROs at certain times or intervals was discussed in one study only.

A common way to report outcomes was by using mean magnitude of change from baseline for each time point (n=42), followed by the mean or median PRO scores per time point (n=34). The number and percentages of responders and non-responders (e.g., improved/not improved, worsened/not worsened relative to baseline per time point) (n=15), and other categorizations at each time point (e.g., patients in good, moderate, and bad condition) (n=14) were also regularly reported.

Sixteen studies reported that no deaths occurred during treatment or follow-up. In the studies describing death (n=35), 24 studies reported survival probabilities or number of deaths. Sixteen studies reported (some in addition to reporting survival probabilities) a median overall survival time, ranging from 3·6 to 51·3 months (median 14·8 months). In nine studies, no information on survival was provided.

The 60 reviewed studies reported a total of 142 intercurrent events other than death (on average, 2.4 types of intercurrent events per study). Common events were treatment discontinuation due to progressive disease (n=31 studies) or due to adverse events (n=25 studies). In eleven studies, no information regarding intercurrent events was reported.

PRO data collection discontinued after intercurrent events in nearly all cases. In three cases, however, it was explicitly mentioned that the collection of PROs was continued. In general, researchers were more likely to continue collecting clinical outcome(s) than PROs during follow-up after intercurrent events.

More than half of the studies reported p-values or confidence intervals (n=36), while the remainder used descriptive statistics. Paired t-tests (n=5) or non-parametric tests (n=12) such as the Wilcoxon signed-rank test were the most popular statistics used to compare PROs to baseline values. Linear mixed modelling (n=10) was most often applied, to look for trends and changes in PROs over time, followed by repeated measures ANOVA (n=5), and a non-parametric repeated measures analysis (n=1). Three studies assessed the relationships between baseline characteristics and PROs using Pearson correlation coefficients, Spearman correlations or logistic regression. Time to response/non-response was studied in four studies, either by Kaplan-Meier curves (n=3) or the Cox proportional hazards model (n=1).

Twenty-three studies (nearly 40%), compared the PRO data to external information. Most of these studies used a clinically important difference value to define responders and non-responders, based on the change in PRO values from baseline. Often (19/23) a reference or rationale for the chosen clinically important difference value was given. One study compared the results to summary values from a healthy population. Historical control data were used in only one phase II study. This was a study on breast cancer patients where self-reported symptoms were compared to a historical control group from a previously conducted RCT in a similar population of women using mixed-effects linear regression models.[26]

Addressing limitations of the PRO analysis

Table 3 shows whether missing data and other potential sources of bias in analyzing PRO data were addressed. Twenty-six studies did not use any specific method for handling missing data and worked with the available data, implicitly assuming missing completely at random. Methods mentioned for dealing with missing data were linear mixed modelling (n=6), single imputation methods such as last observation carried forward or mean imputation (n=4), and multiple imputation (n=1). Note that these studies considered some form of missingness at random, while there typically were no data post specific intercurrent events, a priori not after death. It was not clear in nearly one third of the studies (n=19) how missing data were handled in the PRO analysis, and this happened surprisingly more often in studies that used PRO as primary outcome (8/13) compared to secondary (11/42) outcomes.

Various sources of potential bias and limitations of the PRO analysis were discussed (Table 3). Limitations due to the absence of a randomized control group were most often mentioned (n=36). Fifteen single-arm studies discussed that their choice of the study population may be restrictive and that the generalizability of the trial results to a larger population may be limited. Ten studies mentioned other potential explanations for the observed treatment effects, for

example, confounding by other (unmeasured) clinical characteristics; one study mentioned confounding by response shift. Seven studies noted that informative missing values might have influenced the results. Other potential limitations discussed were potential inadequate choice of PRO instrument (n=5), subjective reporting of PROs (n=1), and small sample size (n=1). Confounding in between group comparisons was discussed by the study that compared PROs to an external control group, one study that compared PROs of treatment responders to non-responders, and two other studies that compared the main clinical outcome to external controls.

Conclusion on the treatment regarding PROs

Table 4 gives an overview of stated conclusions on the treatment based on the PRO analysis in the reviewed papers. In most single-arm studies (n=51), the results of the PRO analysis were seen as providing evidence in favor of the treatment by reporting the following: i) (some) PROs were improved by the treatment during the study (n=22); ii) PROs were maintained or not significantly changed under treatment (n=11); or iii) PROs did not deteriorate substantially in the long-term (n=18). Eight studies mentioned that the anticipated treatment benefit for PROs based on the primary analysis was not confirmed. One study explicitly mentioned that a negative treatment effect on the PROs was observed, with the conclusion being "not supportive".

4. Discussion

In this review which aimed to describe the current state of published single-arm oncology studies involving PROs, we found few single-arm cancer studies published between 2018-2021 stated a clear research hypothesis for the PRO analysis or used PROs as primary outcome. Besides a clear hypothesis, other essential information on the design and analysis was often missing. There was a lack of clarity on the role of specific PROs and why the PRO analysis was conducted, which is essential to enable interpretation of study results and conclusions.

Previous reviews of PRO analysis in RCTs also found that PRO analyses failed to meet current reporting standards such as the CONSORT-PRO and ISOQOL reporting standards.[21,27–30] One potential reason for this could be the finding that PROs are often secondary endpoints that may be perceived as less relevant for decision making and therefore less well thought out in the study protocol.

Most studies did not collect PROs after patients stopped treatment. There may be barriers to measure PROs after treatment discontinuation: patient burden of reporting, in particular in the palliative setting, higher risk of missing values and reporting bias (healthier patients are more likely to report) and costs of monitoring/reporting. However, especially in the absence of

documented reasons for treatment termination or study withdrawal, it becomes impossible to evaluate the situation after treatment termination without making strong assumptions. Treatment termination is a non-randomized event that is highly dependent on the patient's health status and thus probably also on the response to the treatment itself.

No studies specifically discussed intercurrent events, such as disease progression or the occurrence of side effects that may influence treatment continuation and PRO values. Many studies used mixed models or other (implicit) imputation methods, without any motivation. Considering how intercurrent events are handled is important for the analysis and interpretation of the results, because the way in which intercurrent events are taken into account in the analysis may affect the interpretation of study results. For instance, when stopping drug treatment because of side effects, for a quality of life PRO measure, patients and health care providers could be interested in values after treatment has been stopped, while for side effects, the main interest may be limited to PROs while on treatment. We stress the need to consider these issues beforehand, i.e., to carefully formulate a research question which incorporates the handling of intercurrent events, and to choose appropriate statistical methods that match the research question and corresponding estimand. Death as intercurrent event in particular requires an adapted approach since PROs after death are not defined. Several studies imputed values after death or assumed an implied trajectory values after death based on what happened before death, for example by using mixed models. This may yield results without clinical interpretation. An overview on different ways to handle death with corresponding statistical approaches is given in Kurland et al.[31] Most of the reviewed studies were published before the ICH-E9 estimand framework was introduced. Hopefully, in the future, the estimand framework will lead to better formulated research questions with more explicit specification of the handling of intercurrent events.

Information about missing data and how missing values were dealt with, with a justification for the approach, was generally lacking or incomplete. Many studies just summarized observed PRO values ignoring missing values. For many statistical methods, ignoring missing values may come with assumptions like missing completely at random (MCAR) that are likely not reasonable in cancer trials and therefore may lead to bias. More attention to the underlying assumptions of handling missing data is needed and sensitivity analysis should be considered to evaluate the impact of deviation from the assumptions used in the statistical analysis (such as missing at random). When studies performed comparisons, often changes from baseline within the single-arm study were considered. Also regularly, changes from baseline were

compared to a pre-specified clinically important difference value. Unfortunately, change from baseline may be caused by other factors than the treatment. For instance, response shift mostly (not always) leads to better scores.[18] This means that health status did not improve but patients adjusted their answers to their previous experiences and might judge a symptom less severe than earlier. And worse scores might not be related to the treatment itself but to the course of disease in terms of health status, e.g., worsening due to (early) disease progression. Only one of the 60 studies used historical control data.[38] Also here, confounding continues to be an important issue, albeit more informed, as an adequate randomized control group is still lacking. These and further limitations of single-arm studies were not always considered during study design nor mentioned in the discussion.

Despite the profound methodological concerns, the majority of the reviewed studies concluded that PRO results were supportive of the use of treatment, either because positive effects on PROs were observed, or PROs did not deteriorate substantially. Only one study found the treatment not recommendable based on the PRO results. This may be encouraged in part by the lack of well-defined research hypotheses in most studies, which may have resulted in more positive findings being reported.

Our review has several limitations. While the considered literature reflects recent practice, the limitation to the 60 most recently published papers hinders disclosing information about trends over time. We focused on single-arm studies, excluding multiple single-arm trials under one master protocol (i.e., basket design or umbrella design).[39,40] A dedicated review of these increasingly popular designs would be interesting future work. We used broad inclusion criteria in our review to obtain a general overview of PRO use in single arm studies. This broad scope of our review can also be seen as a limitation. For example, narrowing down the review to studies in a palliative and supportive care setting where the primary aim is to improve QOL, would provide more valuable insights on specific patient-centered studies. Conducting such a review would also be an important next step of research. Finally, our findings that most publications report PRO results favorable to the treatment may be related to publication bias toward positive studies which occurs regardless of endpoints being assessed.

To conclude, our comprehensive and detailed review on current practice of handling PROs in single arm oncology studies – the first of its kind (to our knowledge) – uncovered many specific issues that should be more carefully considered to improve the design, analysis, reporting, and interpretation of PRO data in single-arm studies. These findings will steer the forthcoming

SISAQOL-IMI recommendations and guidelines in single-arm PRO studies toward improving patient care and research.

Contributors

SC, EG, LL, JC, SR, JZM, WS and CDA conceived, designed, and planned the study. LL and JC analysed the data. LL, JC, SC and EG wrote the first draft of the manuscript. LL, JC, JZM, WS, CDA, AA, YB, JCC, RSF, MHF, AR, JCR, RS, SR, EG, and SC reviewed papers. All authors contributed to the development of the study, discussion of results, reviewed and revised the manuscript critically, and approved the final version for submission. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Declaration of interests

This publication reflects the views of the individual authors and should not be construed to represent official views or policies of the European Medicines Agency (EMA), the US Food and Drug Administration (FDA), US National Cancer Institute (NCI), Medicines and Healthcare products Regulatory Agency (MHRA), Institute for Quality and Efficiency in Health Care (IQWiG), Health Canada, the Norwegian Medicines Agency (NOMA), the American Society of Clinical Oncology (ASCO) or the European Society for Medical Oncology (ESMO) or any other institution, organization, or entity. This publication reflects the authors' view and that neither IMI nor the EU, EFPIA are responsible for any use that may be made of the information contained therein.

JCR reports payments for clinical lessons for neurologists by IMeedu Het Centrale Brein, support for attending the European Epilepsy Congress, Geneva, July 2022 by Arvelle Therapeutics, and is chair (unpaid) of the EORTC Quality of Life Group, and Member Scientific Advisory Board (unpaid) of the Dutch Epilepsy Foundation. AR is currently an employee of Modus Outcomes. SR is a current employee of Pfizer Inc and ex-employee of Novartis Pharma; RS is a current employee of Pfizer AB. All other authors declare no competing interests. This study received no funding from the US National Institutes of Health (NIH). None of the authors were fully or partly NIH funded, employed by NIH, or are in receipt of an NIH grant for this work.

References

1. Bottomley A, Pe M, Sloan J, Basch E, Bonnetain F, Calvert M, et al. Moving forward toward standardizing analysis of quality of life data in randomized cancer clinical trials. *Clin Trials Lond Engl* 2018; 15(6): 624–30.

2. Basch E, Abernethy AP, Mullins CD, Reeve BB, Smith ML, Coons SJ, et al. Recommendations for incorporating patient-reported outcomes into clinical comparative effectiveness research in adult oncology. *J Clin Oncol* 2012; 30(34): 4249–55.

3. US Food and Drug Administration. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes* 2006; 4:79.

4. European Medicines Agency-Committee for Medicinal Products for Human Use. Guideline on the evaluation of anticancer medicinal products in man – Revision 5. 2017. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-evaluation-anticancer-medicinal-products-man-revision-5_en.pdf (assessed Sep 22, 2017)

5. Gnanasakthy A, Barrett A, Evans E, D'Alessio D, Romano CD. A review of patient-reported outcomes labeling for oncology drugs approved by the FDA and the EMA (2012-2016). *Value Health J Int Soc Pharmacoeconomics Outcomes Res* 2019; 22(2): 203–9.

6. Teixeira MM, Borges FC, Ferreira PS, Rocha J, Sepodes B, Torre C. A review of patient-reported outcomes used for regulatory approval of oncology medicinal products in the European Union between 2017 and 2020. *Front Med* 2022; 9: 968272.

7. Andre EB, Reynolds R, Caubel P, Azoulay L, Dreyer NA. Trial designs using real-world data: The changing landscape of the regulatory approval process. *Pharmacoepidemiol Drug Saf* 2020; 29(10): 1201–12.

8. Davi R, Mahendraratnam N, Chatterjee A, Dawson CJ, Sherman R. Informing single-arm clinical trials with external controls. *Nat Rev Drug Discov* 2020; 19(12): 821–2.

9. Ghadessi M, Tang R, Zhou J, Liu R, Wang C, Toyoizumi K, et al. A roadmap to using historical controls in clinical trials – by Drug Information Association Adaptive Design Scientific Working Group (DIA-ADSWG). *Orphanet J Rare Dis* 2020; 15(1): 69.

10. Gray CM, Grimson F, Layton D, Pocock S, Kim J. A framework for methodological choice and evidence assessment for studies using external comparators from real-world data. *Drug Saf* 2020; 43(7): 623–33.

11. Khozin S, Blumenthal GM, Pazdur R. Real-world data for clinical evidence generation in oncology. *Jnci-J Natl Cancer Inst* 2017; 109(11).

12. Simon R, Blumenthal GM, Rothenberg ML, Sommer J, Roberts SA, Armstrong DK, et al. The role of nonrandomized trials in the evaluation of oncology drugs. *Clin Pharmacol Ther* 2015; 97(5): 502–7.

13. US Food and Drug Administration. Rare diseases: natural history studies for drug development - guidance for Industry. 2019. https://www.fda.gov/media/122425/download (assessed Mar 1, 2019)

14. Hatswell AJ, Baio G, Berlin JA, Irs A, Freemantle N. Regulatory approval of pharmaceuticals without a randomised controlled study: analysis of EMA and FDA approvals 1999-2014. *Bmj Open* 2016; 6(6).

15. Tibau A, Ocana A, Anguera G, Seruga B, Templeton AJ, Barnadas A, et al. Oncologic drugs advisory committee recommendations and approval of cancer drugs by the US Food and Drug Administration. *Jama Oncol* 2016; 2(6): 744–50.

16. Hilal T, Gonzalez-Velez M, Prasad V. Limitations in clinical trials leading to anticancer drug approvals by the US Food and Drug Administration. *JAMA Intern Med* 2020; 180(8): 1108–15.

17. Woolacott N, Corbett M, Jones-Diette J, Hodgson R. Methodological challenges for the evaluation of clinical effectiveness in the context of accelerated regulatory approval: an overview. *J Clin Epidemiol* 2017; 90: 108–18.

18. Vanier A, Oort FJ, McClimans L, Ow N, Gulek BG, Bohnke JR, et al. Response shift in patient-reported outcomes: definition, theory, and a revised model. *Qual Life Res Int J Qual Life Asp Treat Care Rehabil* 2021; 30(12): 3309–22.

19. The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use, Committee for Medicinal Products for Human Use-EMA. ICH-E9 (R1) Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. 2020. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf (assessed Jan 30, 2020).

20. Fiero MH, Pe M, Weinstock C, King-Kallimanis BL, Komo S, Klepin HD, et al. Demystifying the estimand framework: a case study using patient-reported outcomes in oncology. *Lancet Oncol* 2020; 21(10): e488–94.

21. Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD, et al. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *Jama-J Am Med Assoc* 2013; 309(8): 814–22.

22. Calvert M, Kyte D, Mercieca-Bebber R, Slade A, Chan AW, King MT, et al. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the SPIRIT-PRO extension. *JAMA* 2018; 319(5): 483.

23. Coens C, Pe M, Dueck AC, Sloan J, Basch E, Calvert M, et al. International standards for the analysis of quality-of-life and patient-reported outcome endpoints in cancer randomised controlled trials: recommendations of the SISAQOL Consortium. *Lancet Oncol* 2020; 21(2): e83–96.

24. Bottomley A, Pe M, Sloan J, Basch E, Bonnetain F, Calvert M, et al. Analysing data from patient-reported outcome and quality of life endpoints for cancer clinical trials: a start in setting international standards. *Lancet Oncol* 2016; 17(11): e510–4.

25. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021; 372: n160.

26. Santa-Maria CA, Bardia A, Blackford AL, Snyder C, Connolly RM, Fetting JH, et al. A phase II study evaluating the efficacy of zoledronic acid in prevention of aromatase inhibitor-associated musculoskeletal symptoms: the ZAP trial. *Breast Cancer Res Treat* 2018; 171(1): 121–9.

27. Pe M, Dorme L, Coens C, Basch E, Calvert M, Campbell A, et al. Statistical analysis of patient-reported outcome data in randomised controlled trials of locally advanced and metastatic breast cancer: a systematic review. *Lancet Oncol* 2018; 19(9): e459–69.

28. Fielding S, Ogbuagu A, Sivasubramaniam S, MacLennan G, Ramsay CR. Reporting and dealing with missing quality of life data in RCTs: has the picture changed in the last decade? *Qual Life Res Int J Qual Life Asp Treat Care Rehabil* 2016; 25(12): 2977–83.

29. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ* 2011; 342: d40.

30. Brundage M, Blazeby J, Revicki D, Bass B, de Vet H, Duffy H, et al. Patient-reported outcomes in randomized clinical trials: development of ISOQOL reporting standards. *Qual Life Res* 2013; 22(6): 1161–75.

31. Kurland BF, Johnson LL, Egleston BL, Diehr PH. Longitudinal data with follow-up truncated by death: match the analysis method to research aims. Stat Sci 2009; 24(2): 211.

32. Bachet JB, Wyrwicz L, Price T, Cremolini C, Phelip JM, Portales F, et al. Safety, efficacy and patient-reported outcomes with trifluridine/tipiracil in pretreated metastatic colorectal cancer: results of the PRECONNECT study. *ESMO Open* 2020; 5(3): e000698.

33. Gewandter JS, Chaudari J, Ibegbu C, Kitt R, Serventi J, Burke J, et al. Wireless transcutaneous electrical nerve stimulation device for chemotherapy-induced peripheral neuropathy: an open-label feasibility study. *Support Care Cancer* 2019; 27(5): 1765–74.

34. Konidaris G, Paul E, Kuznik A, Keeping S, Chen CI, Sasane M, et al. Assessing the value of cemiplimab for adults with advanced cutaneous squamous cell carcinoma: a cost-effectiveness analysis. *Value Health* 2021; 24(3): 377–87.

35. Laetsch TW, Myers GD, Baruchel A, Dietz AC, Pulsipher MA, Bittencourt H, et al. Patient-reported quality of life after tisagenlecleucel infusion in children and young adults with relapsed or refractory B-cell acute lymphoblastic leukaemia: a global, single-arm, phase 2 trial. *Lancet Oncol* 2019; 20(12): 1710–8.

36. Quick AM, Zvinovski F, Hudson C, Hundley A, Evans C, Stephens JA, et al. Patient-reported sexual function of breast cancer survivors with genitourinary syndrome of menopause after fractional CO2 laser therapy. *Menopause* 2021; 28(6): 642–9.

37. Takahashi M, Masuda N, Nishimura R, Inoue K, Ohno S, Iwata H, et al. Palbociclib-letrozole as first-line treatment for advanced breast cancer: Updated results from a Japanese phase 2 study. *Cancer Med* 2020; 9(14): 4929–40.

38. Santa-Maria CA, Bardia A, Blackford AL, Snyder C, Connolly RM, Fetting JH, et al. A phase II study evaluating the efficacy of zoledronic acid in prevention of aromatase inhibitor-associated musculoskeletal symptoms: the ZAP trial. *Breast Cancer Res Treat* 2018; 171(1): 121–9.

39. Janiaud P, Serghiou S, Ioannidis JPA. New clinical trial designs in the era of precision medicine: an overview of definitions, strengths, weaknesses, and current use in oncology. *Cancer Treat Rev* 2019; 73: 20–30.

40. Park JJH, Hsu G, Siden EG, Thorlund K, Mills EJ. An overview of precision oncology basket and umbrella trials for clinicians. *CA Cancer J Clin* 2020; 70(2): 125–37.