Statistical methods and data visualisation of patient-reported outcomes in early phase dose-finding oncology trials: a methodological review

Emily Alger,[a] Anna Minchom,[b] Olalekan Lee Aiyegbusi,[c,d] Matthew Schipper,[e] and Christina Yap[a,*]

[a]Clinical Trial and Statistics Unit, Institute of Cancer Research, London, UK
[b]Drug Development Unit, Royal Marsden/Institute of Cancer Research, London, UK
[c]Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, College of Medical and Dental Sciences, University of Birmingham, UK
[d]National Institute for Health and Care Research (NIHR) Birmingham Biomedical Research Centre, University of Birmingham, Birmingham, UK
[e]Departments of Radiation Oncology and Biostatistics, University of Michigan, Ann Arbor, MI, USA

eClinicalMedicine 2023;64: 102228

Published Online xxx https://doi.org/10.1016/j.eclinm.2023.102228

Summary

Background Traditionally, within dose-finding clinical trials, treatment toxicity and tolerability are assessed by clinicians. Research has shown that clinician reporting may have inadequate inter-rater reliability, poor correlation with patient reported outcomes, and under capture the true toxicity burden. The introduction of patient-reported outcomes (PROs), where the patient can assess their own symptomatic adverse events or quality of life, has potential to complement current practice to aid dose optimisation. There are no international recommendations offering guidance for the inclusion of PROs in dose-finding trial design and analysis. Our review aimed to identify and describe current statistical methods and data visualisation techniques employed to analyse and visualise PRO data in published early phase dose-finding oncology trials (DFOTs).

Methods DFOTs published from June 2016–December 2022, which presented PRO analysis methods, were included in this methodological review. We extracted 35 eligible papers indexed in PubMed. Study characteristics extracted included: PRO objectives, PRO measures, statistical analysis and visualisation techniques, and whether the PRO was involved in interim and final dose selection decisions.

Findings Most papers (30, 85.7%) did not include clear PRO objectives. 20 (57.1%) papers used inferential statistical techniques to analyse PROs, including survival analysis and mixed-effect models. One trial used PROs to classify a clinicians' assessed dose-limiting toxicities (DLTs). Three (8.6%) trials used PROs to confirm the tolerability of the recommended dose. 25 trial reports visually presented PRO data within a figure or table within their publication, of which 12 papers presented PRO score longitudinally.

Interpretation This review highlighted that the statistical methods and reporting of PRO analysis in DFOTs are often poorly described and inconsistent. Many trials had PRO objectives which were not clearly described, making it challenging to evaluate the appropriateness of the statistical techniques used. Drawing conclusions based on DFOTs which are not powered for PROs may be misleading. With no guidance and standardisation of analysis methods for PROs in early phase DFOTs, it is challenging to compare study findings across trials. Therefore, there is a crucial need to establish international guidance to enhance statistical methods and graphical presentation for PRO analysis in the dose-finding setting.

Funding EA has been supported to undertake this work as part of a PhD studentship from the Institute of Cancer Research within the MRC/NIHR Trials Methodology Research Partnership. AM is supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at the Royal Marsden NHS Foundation Trust, the Institute of Cancer Research and Imperial College.

*Corresponding author. Clinical Trials and Statistics Unit, The Institute of Cancer Research, London, UK.
E-mail address: christina.yap@icr.ac.uk (C. Yap).

Keywords: Patient-reported outcomes; Dose-finding; Methodological review; Oncology; Statistical methods

Research in context

Evidence before this study

Previous reviews have revealed that though patient-reported outcome (PRO) endpoints are only present within a limited number of dose-finding oncology trials (DFOTs), its use has increased over time. There are no guidelines on how PROs should be incorporated in DFOTs, analysed and interpreted, and the quality of PRO analysis in this setting is unknown. To evaluate PRO analysis techniques within DFOTs, we searched the bibliographic database MEDLINE (via PubMed) for eligible studies. Oncology trial reports or trial protocols which were published between June 2016 and December 2022 and included PRO analysis methods within the dose-finding component were included.

Added value of this study

To the best of our knowledge, this is the first study providing a contemporary review of statistical and graphical PRO analysis techniques in DFOTs. In line with the introduction of CONSORT-PRO, SPIRIT-PRO, and the founding of the SISAQOL consortium, it is becoming increasingly important to encourage quality PRO reporting within trials. We highlight the benefits of incorporating PROs within dose selection decisions as a long-term approach to assess treatment tolerability and identify methodological weaknesses and recommendations with regards to current analysis practice.

Implications of all the available evidence

This review highlights the lack of standardisation and consistency of PRO analysis in DFOTs. This work strengthens the call for new PRO analysis recommendations within the early phase dose-finding setting. Looking forward, the development of guidance to analyse and report PROs will facilitate the interpretation of PRO findings at different dose levels and will aid dose adaptation decisions and final dose selection. Development of guidance will not only generate analysis methods which can contribute to more complete treatment tolerability profiles but may also improve the accuracy of dose determination methods and enable a more accurate synthesis of PRO data in order to achieve patient-centred clinical development.

Introduction

Within early phase dose-finding oncology trials (DFOTs), the safety and tolerability of a new agent is assessed. Often the recommended phase II dose (RP2D) has generally been set at or close to the maximum tolerated dose (MTD). The MTD is often determined by observing dose-limiting toxicities (DLTs) within patients enrolled on a trial. Clinicians usually grade toxicities using the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE), with grade three or above toxicities considered a DLT.[1]

This approach to MTD or RP2D estimation relies solely on a clinician assessment of tolerability and is not informed by a patient's own evaluation of their quality of life whilst receiving treatment.[2] Whilst clinicians can assess adverse events (AEs) such as fever and blood profiles, other AEs such as fatigue or nausea can be subjective toxicities which may be difficult for them grade consistently and which they may undercapture.[3–6]

Furthermore, the traditional DLT assessment window (usually one or two cycles of treatment) may not be efficient for therapies which are administered for extended periods of time. Oncology treatments are often administered until disease progression (treatment resistance) is observed[7] and therefore, the desired assessment window to assess treatment tolerability may elongate.[1] Whilst radiotherapy treatments often have longer DLT assessment periods,[8,9] the short DLT assessment window typically used to assess the tolerability of cytotoxic agents may not capture the toxicities beyond the first cycle of treatment. A retrospective study by Postel-Vinay et al. found that 57% of grade three or four toxicities experienced by patients treated with a molecularly targeted agent (MTA) during a phase I trial occurred beyond the first cycle of treatment.[7] Additionally, the current approach to tolerability assessment does not typically capture the toxicity associated with prolonged Grade 2 toxicities. These types of toxicities may become an impediment to treatment tolerability over longer term dosing schedules. For example, Durvalumab was recently approved to be given for 12 months following chemo-radiation in for locally advanced non-small cell lung cancer. However, many patients discontinue early due to treatment toxicity.[10]

There is growing interest in the introduction of Patient-Reported Outcomes (PROs) within early phase dose-finding trials to inform tolerability of treatment. The US Department of Health defines a PRO as "any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else".[11] Incorporating PROs within DFOTs may enhance researchers' understanding of toxicity profiles and improve the accuracy of RP2D determination.[12]

The U.S. Food and Drug Administration's (FDA) Project Optimus initiative encourages the leveraging of clinical and non-clinical data to aid dose optimization within pre-marketing drug development.[13] Against this backdrop, Friends of Cancer Research have encouraged the introduction of PROs to guide drug optimization.[14]

A systematic review by Lai-Kwon et al. found that only 5.3% (548/10,372) of trials presented on ClinicalTrials.gov in 2007–2020 contained a PRO endpoint, though its use has increased significantly over time.[15] Currently, as PRO endpoints are only present in a small number of DFOTs, there is limited literature encouraging routine practice in their analysis across the whole trial process, not least dose-finding trials.[16] Inclusion of PROs in DFOTs has the potential to more accurately characterize toxicity and tolerability as they vary with treatment dose.

New guidance for the inclusion of PROs within interventional clinical trials has been developed to ensure comprehensive publishing of PRO data. CONSORT-PRO[17] (Consolidated Standards of Reporting Trials–PRO) and SPIRIT-PRO[18] (Standard Protocol Items: Recommendations for Interventional Trials-PRO) were developed to ensure the rigorous reporting of PROs within trial reports and trial protocols. Additional work has been undertaken to introduce guidance for statistical analysis plans (SAPs) in early phase clinical trials,[19] alongside ongoing development of the SPIRIT and CONSORT extensions for early phase dose-finding trials.[20,21] The newly founded SISAQOL Consortium (Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data) have generated recommendations to adapt statistical methods, missing data and statistical terminology for the incorporation of PROs within cancer randomized controlled trials (RCTs).[22] However, there is currently no international guidance for PRO statistical analysis in early phase DFOTs,[12,23] which are typically non-randomized.[24] It is unclear which PRO statistical analysis strategies are currently being employed within DFOTs.

To assess the current PRO analysis methods and data visualization techniques utilised within dose-finding oncology trials we conducted a methodological review via PubMed.

Our objectives were to: describe the study characteristics of published trials investigating PROs within early phase DFOTs, identify and evaluate current techniques to analyse PRO measures within early phase DFOTs, and to explore the data visualisation techniques used to display PRO measures graphically.

Methods

Study strategy and selection criteria

Papers eligible for this methodological review were dose-finding oncology trials or trial protocols archived on PubMed between 01/06/2016 and 31/12/2022 which described how PRO data was to be analysed. For a trial protocol, the paper should include a presentation of the statistical techniques which the researchers planned to use to analyse PRO data. For a trial report, the paper should include a presentation of the statistical techniques which were used to analyse PRO data collected during the study. Clinical trials were extracted by EA in XML format on 21/04/2023.

Eligible papers were extracted using the following search strategy: ("dose-find*" OR "dose escalat*" OR "dose find*" OR "dose expan*" OR "single ascending dose" OR "multiple ascending dose" OR "first in man" OR "first in human" OR "early phase" OR "phase 1a" OR "phase 1b" OR "phase ia" OR "phase ib" OR "RP2D") AND ("quality of life" OR "patient reported outcome*" OR "patient-reported outcome*") AND ("2016/06/01" [Date–Completion]: "2022/12/31" [Date–Completion]).

Each entry was reviewed for eligibility by one reviewer (EA). Papers were eligible if they:

(1) Reported a dose-finding component within an early phase trial with a cancer population,
(2) The intervention of interest was either a drug or radiotherapy,
(3) PRO analysis was presented within the dose-escalation component.

For data verification of eligibility of papers, 9.6% (54/562) of randomly selected papers were double-reviewed (EA and AM) to ensure eligible papers were captured. For data extraction, 14.3% (5/35) of randomly selected eligible papers were assessed by an additional reviewer (AM, OLA, CY) to ensure all relevant features were correctly extracted. All queries that arose during data extraction were discussed and any differences of opinions between reviewers were resolved through discussion.

Data analysis

For eligible papers, the following characteristics were extracted: year of publication, study population, cancer type, trial phase, anti-cancer agent, funder type, number of centers, trial design, PRO instrument, frequency of PRO assessment, minimal clinically important difference (MCID), primary endpoint. We also extracted additional PRO statistical features including: PRO objectives, PRO analysis method, PRO visualisation technique, discussion of missing PRO data reported, number of patients with missing PRO data, reasons described for missing PRO data, and methodology to manage missing PRO data described.

To evaluate the use of PROs across a diverse range of trial demographics, we firstly extracted basic trial characteristics. Extracted trial characteristics were a subset of those in Yap et al.'s[24] review on the quality of early phase DFOTs. To evaluate the statistical rigor of PRO methods and analysis, we extracted features which provided a general overview of statistical methodology. These included items extracted by Lai-Kwon et al.[15] in their ClinicalTrials.gov review of PROs in DFOTs. In addition, we collected PRO objectives, how PRO analysis was attempted, how it was presented to the reader (data

visualisation) and how missing data was recorded and mitigated. The option categories for each feature are presented in Tables 1, 2, 3, and 4.

Statistics

For each extracted feature, the percentage of papers which reported each possible outcome was presented. For features with a numerical domain, the corresponding 95% confidence interval (CI) was also calculated. No data imputation was performed.

A linear regression model with year as the independent variable was fitted to assess the trend in the analysis of PROs within DFOTs over time, and model diagnostics were evaluated to assure the suitability of this model. R version 4.2.1 was used for the statistical analysis.

Role of funding source

EA has been supported to undertake this work as part of a PhD studentship from the Institute of Cancer Research within the MRC/NIHR Trials Methodology Research Partnership.

AM is supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at the Royal Marsden NHS Foundation Trust, the Institute of Cancer Research and Imperial College. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

No funders had a role in the data collection, data analyses, interpretation, or writing of this report.

	Overall (N = 35)
Trial design	
Algorithmic	22 (62.9%)
3 + 3	16 (45.7%)
Rolling six	4 (11.4%)
Other	2 (5.7%)
Model based	7 (20.0%)
Continual reassessment method (CRM)	5 (14.3%)
Time-to-event continual reassessment method (TiTE-CRM)	1 (2.9%)
Escalation with overdose control (EWOC)	1 (2.9%)
Unclear	6 (17.1%)
Intervention type	
Drug	17 (48.6%)
Drug + radiotherapy	5 (14.3%)
Radiotherapy	13 (37.1%)
Number of PRO measures	
1	15 (42.9%)
2	11 (31.4%)
3	8 (22.9%)
5	1 (2.9%)
Number of PRO assessments	
Mean (SD)	6.31 (4.75)
Median [min, max]	5.00 [2.00, 24.0]
Type of PRO analysis	
Descriptive	15 (42.9%)
Descriptive & inferential	11 (31.4%)
Inferential	9 (25.7%)
PRO endpoint	
Exploratory	1 (2.9%)
Secondary	23 (65.7%)
Tertiary	1 (2.9%)
Unclear	10 (28.6%)

A full table presenting a summary of all extracted features is presented in Supplementary Table S2.

Table 1: **Characteristics of eligible early phase dose-finding oncology trials.**

Results

562 papers were assessed for eligibility and 35 papers were eligible for the review. A study flow diagram for this study is presented in Fig. 1. There was 94.4% agreement between reviewers (EA and AM) assessing eligible trials. For the three papers where EA and AM disagreed on paper eligibility, disagreements were resolved by discussion with additional arbitrators (OLA and CY).

30 of the eligible manuscripts detailed completed trials. Five manuscripts were published trial protocols.[8,38,39,47,55] Supplementary Table S1 summarises the eligible papers analysed within this methodological review.

An overview of papers included in this review is presented in Table 1. When PROs were labelled as an endpoint (n = 25), it was most often identified as a secondary endpoint (23, 92.0%). Two papers marked the PRO outcome as an exploratory or tertiary endpoint. All papers which presented a PRO endpoint also presented the PRO analysis within the paper.

34 distinct questionnaires were used to evaluate PROs and are presented in Table 2. Often papers utilized more than one patient-reported outcome measure (PROM). Each paper considered a median of 2 questionnaires (Range: 1–4).

Notably, no eligible trials considered the NCI-PRO-CTCAE questionnaire. This criterion was developed in 2014 by the National Cancer Institute to complement the NCI-CTCAE criteria used by clinicians to assess treatment toxicities.[61]

The majority of eligible papers which stated their trial design (29, 82.9%) used a 3 + 3 dose escalation design (16, 45.7%), however Rolling six, (TiTE)-CRM, EWOC, and other specific algorithmic designs were also present within the review. Nearly all (33, 94.3%) trials considered an adult population, however two considered pediatric populations.[53,56] In these cases, pediatric specific quality of life questionnaires were considered (PedsQL and Impact of Pediatric Illness (IPI) Parent Report Form).

Patient-reported outcomes were only considered as part of the dose-finding decisions in four (11.4%) trials.[30,32,33,47] In three cases, the maximum tolerated

Questionnaire	Number of times used as instrument	Associated trials
EORTC QLQ-C30 (Quality of life of cancer patients)	14	8,25–37
EQ-5D (Generic quality of life)	5	28,38–41
EPIC	4	42–45
IPSS	3	42,44,46
AUA	3	43,44,47
M.D. Anderson Dysphagia Inventory	2	8,28
CLAS1	2	48,49
CLAS2	2	48,49
CLAS3	2	48,49
FACT-G (General)	2	50,51
EORTC QLQ-H&N35 (Head & neck)	2	8,28
FACT-P (Prostate)	1	39
FACT-ES (Endocrine symptoms)	1	52
Impact of Pediatric Illness (IPI) Parent Report Form	1	53
Norfolk QOL-NET (Neuroendocrine tumour)	1	54
FACT-Ga (Gastric)	1	55
EORTC QLQ-LC13 (Lung)	1	25
VASB	1	27
WOMAC	1	40
PedsQL	1	56
SHIM	1	46
EORTC QLQ—STO22 (Gastric)	1	55
FACT-KSI (Kidney)	1	50
EORTC QLQ-BR23 (Breast)	1	31
DLIQ	1	25
Rectal function study questionnaire[57]	1	46
VHI	1	28
EORTC QLQ-PAN26 (Pancreatic cancer)	1	58
O'Leary Interstitial Cystitis Symptom Index	1	47
EORTC QLQ-BN20 (Brain)	1	34
EORTC QLQ-PR25 (Prostate)	1	35
FAACT	1	58
FACT-BP (Bone pain)	1	51
MDASI-BT (Brain tumour)	1	59

A table of acronyms used in this table are explained in Supplementary Table 3.

Table 2: **Type and number of each PROM questionnaire captured within this review, along with the associated trials which utilised each questionnaire.**

dosages were confirmed using the usual dose escalation design (CRM (2) and 3 + 3 (1)) and the PROs were used to confirm the tolerability of the recommended phase 2 dose (RP2D). PROs were used to determine whether the MTD was tolerable from a quality of life perspective after the MTD had been determined. In one case, a specific rise in PRO score (signaling a deterioration in health related quality of life) was defined as a DLT and used to guide 3 + 3 dose escalation.[47]

There was no significant time trend in the number of published trials which reported PRO analysis (0.25, 95% CI: −1.46 to 1.96). The mean number of time points PROs were assessed for a drug intervention was 4.88 (SD: 3.12). The mean number of time points PROs were assessed for a radiotherapy, or drug and radiotherapy combination trial was 7.67 (SD: 5.66).

Analysis strategies

18 (51.4%) extracted manuscripts provided information on the planned analysis of PROs in the methods section of the paper. In the majority of papers (85.7%), the PRO objective was not explicitly stated or it was vague, for example "assessing quality of life". Five papers defined a minimally important clinical difference (MCID) for patients' quality of life deterioration. 15 (42.9%) papers considered only exploratory statistical analysis, these methods included: plotting quality of life scores over time for each patient, and considering average, median and IQR scores for each question. The majority of papers (n = 20) considered inferential, or explanatory and inferential statistics in their analysis. 14 papers (40.0%) used hypothesis tests to analyse PROs, no papers mentioned whether this test was powered. In general, these hypothesis tests were used to assess whether there was a statistically significant difference in PRO scores across time points or across dose cohorts. Only one paper checked model assumptions before model fitting. Anota and colleagues[33] tested the proportional hazard assumption using Schoenfeld residuals before utilising a cox proportional hazard model to predict time to quality of life deterioration. No papers completed any form of model validation. Of the 20 papers which utilised some inferential statistics, 13 papers reported the statistical software used. Software included SAS (n = 5), R (n = 4), SPSS (n = 2), Stata (n = 2).

Of the 30 trial reports eligible for this review, only six papers commented on the number of patients with missing data at PRO assessments. The median number of patients who had at least one missing PRO assessment was 1 (range: 0–2). No paper presented the reason for the missing PRO data.

An overview of the statistical techniques extracted within this review is presented in Table 3.

Presentation of PRO results using figures or tables

80% of trial reports (n = 24/30) visually presented PRO data within a figure or table within their publication. Nine trial papers (30.0%) presented PRO data within a table and 17 trial papers (56.7%) presented PRO data within a figure. 11 papers (36.7%) presented PRO results over each dosage within either a figure or table.

Within the 17 papers which included figures presenting PRO data, 18 distinct plots were identified. The number of times each figure type was used to visualise PRO data is presented in Table 4.

PRO data visualisation methods extracted during this review are presented in Figs. 2 and 3.

Statistical inferential technique	Description	Drug trial (n = 8)	Radiotherapy trial (n = 12)
Wilcoxon signed-rank test	Change in PRO between baseline and another timepoint.	30	28,50,56
Wilcoxon rank-sum test	Association between PRO and dosimetric parameters at timepoints.		28
Mixed-effect model	Fit model predicting PRO score across study from baseline accounting for intra-patient correlation.	40	8,51
	Compare change in PRO score across each dose cohort when time is fitted as a random effect		43
Linear Regression	Fit model predicting average PRO score for specific symptoms across study from baseline.	54	
t-tests	Change in PRO score between baseline and another timepoint or difference in PRO score between dose cohorts.	25,34,35,58	29,38,41,46,48
Fisher's exact test	Association between MCID and each dose cohort		45
Survival Analysis	Fit model predicting time from inclusion in study until deterioration or grade 3/4 toxicity using a Cox proportional hazards model.	33	

PRO: patient-reported outcome; MCID: minimal clinically important difference; HRQoL: health related quality of life.

Table 3: **Description of PRO statistical inferential techniques captured, with associated trials separated by the treatment under investigation.**

Out of the 18 figures, 7 (38.9%) presented PRO data by dose levels, 6 of these figures were longitudinal plots and 1 was a Kaplan–Meier plot. 12 figures plotted PRO scores longitudinally across PRO assessment timepoints. Of these 12 plots, five (41.7%) used a larger score to signify an improved quality of life whereas seven (58.3%) used a larger score to signal a worse quality of life. One paper plotted each individual's overall score.[52] Nine figures presented either the mean or median PRO score for each dose cohort. Two papers plotted some PRO statistic longitudinally but did not label whether the mean or median score was used. Four papers used barplots, boxplots, or histograms to visualise PRO data. One paper presented the proportion of patients which experienced an improvement or deterioration in quality of life between their baseline and last PRO assessment,[26] and one paper used Kaplan–Meier graphs to show the deterioration of quality of life across dose, this plot is presented in Fig. 2.[33] One paper presented box plots of PRO score for each cycle of treatment (see Fig. 2) and a barplot to show the proportion of patients who experienced an improvement or worsening of symptoms at each cycle of treatment. In this paper, asterisks were used above the boxplots to signify statistically significant changes in quality of life from baseline.[40] This was mirrored by another paper which presented a barplot of mean scores for different symptoms for the determined MTD and other dosages (see Fig. 2).[36] One paper used a histogram to show the distribution of PRO scores.[41]

Figure type	Overall (N = 18) (%)	Associated trials
Longitudinal plot	12 (66.7)	30,34,37,42–46,51,52,54,59
Barplot	3 (16.7)	26,36,40
Box plot	1 (5.6)	40
Histogram	1 (5.6)	41
Kaplan–Meier graph	1 (5.6)	33

Table 4: **The number and proportion of times each figure type was used to visualise PRO data, with associated trials.**

Of the nine tables, eight presented descriptive summaries of the PRO scores and one table presented fitted model results. Two papers presented PRO scores as means and standard deviations: one determined the average score before and after treatment[25] and one determined the average difference in score compared to baseline.[32] Four papers presented median PRO scores alongside inter-quartile ranges or ranges. Of these papers, one summarised the PROs by each item and three summarised PROs by each questionnaire used to assess quality of life. One other table highlighted the proportion of each cohort which experienced every symptom before and after treatment. Of the nine tables, four (44.4%) presented PRO data at every time point and not at each dose level. The remaining five tables (55.5%) recorded PROs for each dosage and timepoint separately. One paper presented the median predicted time to deterioration of at least one PRO score and associated hazard score in a table, this is presented in Fig. 3.[33] Three papers presented p-values associated with unpowered hypothesis tests which hypothesized that quality of life altered from baseline.

Discussion

Following the recent MDICT 2022 report,[1] it is increasingly important for trials to identify optimal treatment dosages compared to MTDs in DFOTs. Establishing patient's views on drug tolerability using Patient-reported outcomes (PROs) may help to assess the effect of dosages on treatment tolerability and quality of life in DFOTs. Only one paper used PRO data to guide dose escalation/de-escalation and three (8.6%) papers made some attempt to use PRO analysis to inform their understanding of drug tolerability. Whilst there remains no specific guidance on the reporting of PROs within early phase dose-finding trials, the SISAQOL guidance and CONSORT-PRO extension should provide some direction for authors publishing PRO analysis within early phase dose-finding trials.[17] Though CONSORT-PRO offers guidance within the randomised

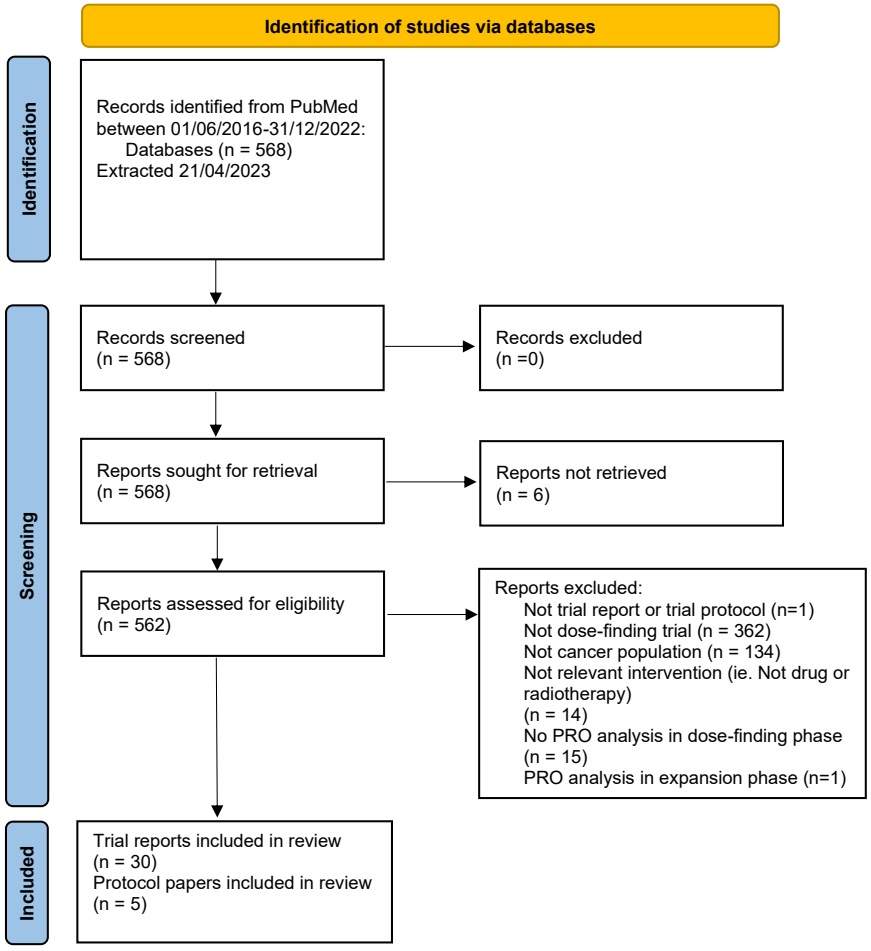


Fig. 1: PRISMA study flow diagram illustrating selection of eligible studies.[60]

trial setting, many guidelines are conveyable to the dose-finding trial and reflect current deficiencies. For example, Item 22 of CONSORT-PRO recommends that "Patient reported outcome data should be interpreted in relation to clinical outcomes including survival data, where relevant".[17]

As there is an absence of established standards for PROs within DFOTs, the evaluation criteria used to evaluate papers within this methodological review has been broadly determined by previous reviews. Thus, the features we have extracted build upon the selection of statistical issues which have previously been identified as critical for the analysis of PROs within DFOTs. These criteria remain broadly in line with established guidelines for PRO analysis in RCTs[17,18,22] and early phase SAPs[19] and ongoing efforts on the development of the SPIRIT and CONSORT extensions for DFOTs.[20,21]

Exploration of PRO analysis techniques has previously been researched in the Phase 2 setting. Among others, analysis techniques such as generalised estimating equations and ordinal log-linear models were suggested, both of which were not utilised within this review.[62]

34 distinct PROMs were used to evaluate quality of life within this review. It is significant to note that some PROMs identified within this review appear to be non-validated PRO instruments. A radiotherapy trial conducted by Sampath and colleagues[46] used a rectal function scale[57] which combined two quality of life assessments–the Prostate Brachytherapy Research Group Protocol PBRG-1 and Radiation Therapy Oncology Group Protocol. Unvalidated PRO measures may not accurately indicate significant changes to quality of life within a trial if items detailing potential patient experiences on treatment are not identified during a rigorous developmental process.

16 papers utilised the EORTC (European Organisation For Research And Treatment Of Cancer) item library as a PROM within this review. Item banks such as this can be used alongside algorithms such as computerized adaptive testing (CAT) to generate short, concise measures capable of evaluating patient symptoms and

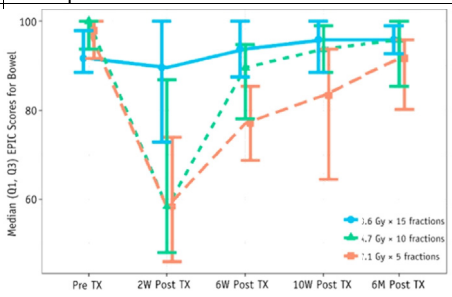
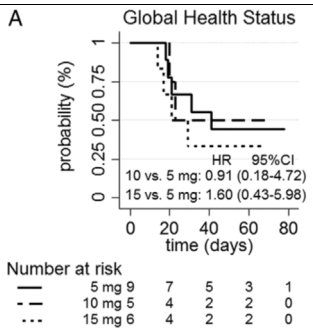
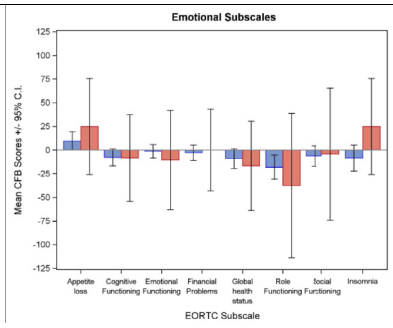
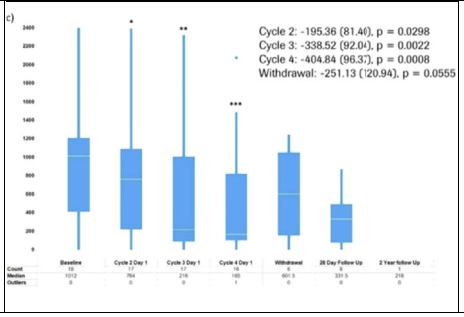
Figure type	Example	Description
Longitudinal plot by dose levels.[42]		Longitudinal plot of median PRO score for each dosage across assessment time points. It is easy to identify that the median and interquartile range is used to summarise the PRO data.
Kaplan-Meier plot by dose levels.[33]		Non-parametric Kaplan-Meier plot presenting time to significant deterioration in quality of life score (with a five point minimal clinically important difference) by dose level.
Barplot[36]		Barplot illustrating mean change from baseline score with 95% confidence intervals across multiple items.
Boxplot[40]		Boxplot summarising PRO data for each assessment, with asterisks marking statistically significant deviations in quality of life score compared to baseline. The number of patients assessed at each time point is highlighted in the table below the figure.

Fig. 2: Exemplar plots visualising PRO data identified within this review.

overall quality of life.[63] However, utilising such a bank within early phase trials (where relevant items may not be known prior to the commencement of the trial) may mean that relevant adverse events are not reported by a PROM. Conversely, whilst issuing a complete item library within a trial may record an exhaustive list of symptoms experienced by a patient, this questionnaire may be time consuming, burdensome for patients, and infeasible to administer frequently within a trial. Crucial research is required to establish core set of items from PRO item libraries (which could be from EORTC, FACIT (Functional Assessment of Chronic Illness Therapy), or PRO-CTCAE) that adequately capture common and clinically significant treatment-related symptoms for employment in the early phase dose-finding setting. The PRO core set of items must be

<table>
<tr><th>Table type</th><th>Example</th><th>Description</th></tr>
<tr><td>Table of model output examining relationship between HRQoL and dose.[33]</td><td>
<table>
<tr><th></th><th colspan="3">Deterioration in at least one HRQoL score</th><th colspan="3">Deterioration in at least one HRQoL score or grade 3/4 toxicity</th></tr>
<tr><th></th><th>N (events)</th><th>Median days (95% CI)</th><th>HR (95% CI)</th><th>N (events)</th><th>Median days (95% CI)</th><th>HR (95% CI)</th></tr>
<tr><td>TTD</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>All</td><td>20 (18)</td><td>21 (19 to 42)</td><td></td><td>20 (18)</td><td>17 (8 to 42)</td><td></td></tr>
<tr><td>5 mg</td><td>9 (9)</td><td>31 (19 to NA)</td><td>1</td><td>9 (9)</td><td>31 (19 to NA)</td><td>1</td></tr>
<tr><td>10 mg</td><td>5 (3)</td><td>29 (20 to NA)</td><td>0.83 (0.22 to 3.17)</td><td>5 (3)</td><td>8 (5 to NA)</td><td>0.96 (0.52 to 3.66)</td></tr>
<tr><td>15 mg</td><td>6 (6)</td><td>17 (14 to NA)</td><td>4.47 (1.33 to 15.09)</td><td>6 (6)</td><td>6 (1 to NA)</td><td>3.31 (1.04 to 10.46)</td></tr>
<tr><td>TTD or no-follow-up</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>All</td><td>20 (19)</td><td>21 (19 to 41)</td><td></td><td>20 (19)</td><td>14 (6 to 41)</td><td></td></tr>
<tr><td>5 mg</td><td>9 (9)</td><td>31 (19 to NA)</td><td>1</td><td>9 (9)</td><td>31 (19 to NA)</td><td>1</td></tr>
<tr><td>10 mg</td><td>6 (5)</td><td>23 (20 to NA)</td><td>1.10 (0.33 to 3.68)</td><td>5 (4)</td><td>8 (5 to NA)</td><td>1.30 (0.39 to 4.34)</td></tr>
<tr><td>15 mg</td><td>6 (6)</td><td>17 (14 to NA)</td><td>3.95 (1.20 to 13.01)</td><td>6 (6)</td><td>6 (1 to NA)</td><td>3.05 (0.98 to 9.46)</td></tr>
</table>
HRQoL, health-related quality of life; TTD, time to HRQoL score deterioration.
</td><td>Table presenting HRQoL deterioration fitted using a Cox-exponential hazard model with predicted time to deterioration (TTD) and associated hazard ratio partitioned by dosage.</td></tr>
<tr><td>Table summarising PRO scores by dose levels.[42]</td><td>
<table>
<tr><th rowspan="2">Assessment tool</th><th rowspan="2">Dose level</th><th colspan="2">2-wk posttreatment difference from baseline</th></tr>
<tr><th>No.</th><th>Median (Min, Q1, Q3, Max)</th></tr>
<tr><td>EPIC-26: Bowel*</td><td>3.6 Gy × 15 fractions</td><td>6</td><td>−4.2 (−50.0, −20.8, 0, 8.3)</td></tr>
<tr><td></td><td>4.7 Gy × 10 fractions</td><td>6</td><td>−33.3 (−87.5, −54.2, 0, 3.3)</td></tr>
<tr><td></td><td>7.1 Gy × 5 fractions</td><td>12</td><td>−31.3 (−75.0, −54.2, −14.6, 8.3)</td></tr>
<tr><td>EPIC-26: Urinary irritative*</td><td>3.6 Gy × 15 fractions</td><td>6</td><td>−6.3 (−25, −18.8, 0, 6.3)</td></tr>
<tr><td></td><td>4.7 Gy × 10 fractions</td><td>6</td><td>0 (−6.3, −6.3, 0, 12.5)</td></tr>
<tr><td></td><td>7.1 Gy × 5 fractions</td><td>12</td><td>−6.3 (−56.3, −18.8, −3.1, 12.5)</td></tr>
<tr><td>EPIC-26: Urinary incontinence*</td><td>3.6 Gy × 15 fractions</td><td>6</td><td>4.1 (−12.5, −8.3, 8.3, 8.3)</td></tr>
<tr><td></td><td>4.7 Gy × 10 fractions</td><td>6</td><td>3.1 (0, 0, 6.3, 14.5)</td></tr>
<tr><td></td><td>7.1 Gy × 5 fractions</td><td>12</td><td>3.1 (−31.3, −11.5, 6.3, 62.8)</td></tr>
<tr><td>IPSS total score†</td><td>3.6 Gy × 15 fractions</td><td>6</td><td>2 (−3, −3, 7, 9)</td></tr>
<tr><td></td><td>4.7 Gy × 10 fractions</td><td>6</td><td>1 (−1, −1, 3, 4)</td></tr>
<tr><td></td><td>7.1 Gy × 5 fractions</td><td>12</td><td>2 (−3, 0, 3, 14)</td></tr>
</table>
*EPIC-26 (range 0–100). The higher the score, the better the patient is doing.
†IPSS (range 0–35). Scores 1 to 7 = mild, 8 to 19 = moderate, 20 to 35 = severe prostate symptoms.
</td><td>Table presenting average summaries of PRO scores for each dosage and assessment time point. Asterisks indicate that a greater EPIC-26 score is related to better quality of life and higher IPSS score is related to severer prostate symptoms.</td></tr>
</table>

Fig. 3: Exemplar tables visualising PRO data identified within this review.

appropriate to patient groups and trial design/drug types.

For investigators who wish to evaluate an overall burden of side effects within an early phase dose-finding trial, many item libraries contain single questions which evaluate the overall impact of symptomatic toxicities on a patient. This includes the FACT-G item GP5 "I am bothered by side effects of treatment", part of the FACIT item library. The US Food and Drug Administration (FDA) and Critical Path Institute's PRO Consortium explored the strengths of such items at a public workshop in 2017.[64] Single items such as FACT-GP5 were recognised for their simplicity and for supporting patients to weight their side effects. Such measures were encouraged to enrich a therapy's side effect profile, particularly the consequences of adverse events on a patient's quality of life.[64]

The diversity of questionnaires used to evaluate PROs may provide some explanation as to why PRO analysis techniques were varied within this review. When PRO objectives were not stated explicitly, it was unfeasible to evaluate whether the statistical analysis approach utilised was appropriate.[65] All 14 papers which analysed PROs using a hypothesis test did not mention if the test was powered. There is a high risk that these tests were underpowered to undertake formal hypothesis testing due to small sample sizes in typical DFOTs. Just over half of papers which did present PRO analysis considered inferential statistical analysis. Of the three methods which considered some form of model fitting, only one paper presented this model fitting visually.[54] All inferential statistical methods which were presented in this review were evaluated by SISAQOL. These methods were assessed for the essential/highly desirable statistical attributes agreed by the SISAQOL consortium.[22] Anota et al.'s trial report is an exemplar paper which presents most features which were desired within this review.[33] Of note, this paper presents a specific PRO objective, clear PRO endpoint, and utilises an appropriate analysis measure (fitting HRQoL deterioration to a cox proportional hazard model by dose levels) which is reviewed favorably by SISAQOL. The cox proportional hazard model can handle censored data and has within-group statistical relevance when it comes to estimating time to quality of life score deterioration.

Within DFOTs, patients can be withdrawn from a study due to toxicities, progression, or death. It is very likely that studies which collect PROs may have missing data and therefore it is important that missing data is reported and methods which manage missing PRO data are described. However, within this review, only six papers comment on the number of patients with missing data.

A review of the data visualisation techniques currently utilised within the early phase setting suggests that when PRO data was published as a figure or table, a

longitudinal plot was most popular. However, barplots were also used to show changes in PRO scores over time. There was heterogeneity in the way PRO scores were presented, with only 36.7% of all trial reports presenting PRO data at each dose level using either a figure or table which limit their usage to inform patients' perspective of treatment tolerability across doses. The majority of papers related a larger PRO score to a worse quality of life. Within SISAQOL's preliminary findings, which aim to standardise the graphical visualization of PRO data, the consortium recommends a larger score to indicate better quality of life. They also recommend that intervals are displayed to indicate thresholds to define "improvement" and "deterioration" in quality of life.[22] The heterogeneity in PRO data visualisation was also discussed by a Consensus Panel of Oncologists, PRO researchers and patients organised by Snyder et al.[66] This panel recommended greater PRO scores to signify better quality of life, however it was noted that PROM measures should not be changed to conform to this recommendation. Instead, descriptive labels could be used to confirm the interpretation of larger scores with words such as "None", "Mild"-"Severe". This panel also recommended the use of line plots to summarise PRO scores to ensure consistent comparison between trial publications. Other publications have reported that clinicians and PRO researchers may misinterpret PRO results due to the variety of PRO measures and how each measure quantified good or poor quality of life.[67] Research presented by Brundage et al. recommended that different data visualisation techniques be used to present PRO data to patients and clinicians.[68] This paper suggested that simple linear plots be used to explain PRO score to patients, and recommended the use of normed scores and p-values to tailor PRO data visualisation techniques to a clinician audience.

This methodological review provides the most contemporary picture of PRO usage within early phase DFOTs to date; however it has some limitations. Whilst the search strategy for this review was rigorous, it may be possible that some papers were missed as we did not include specific PROMs within the search criteria. We instead searched for more general terms such as "quality of life" and "patient-reported outcomes". Nevertheless, this evaluation encompassed a total of 34 unique PRO questionnaires, making it a comprehensive assessment. It may also be the case that more published papers may have been captured if we expanded the intervention beyond a drug or radiotherapy. We acknowledge that conducting an independent validation on 14.3% of the extracted data may not completely eliminate subjectivity. However, we are content that continuous discussion among the authors minimized its impact. Lai-Kwon and colleagues' review[15] has previously highlighted that the number of publications which consider a PRO endpoint is increasing over time. The long-term nature of a clinical trial, from trial registration to publication, could mean that DFOTs which are currently analysing PROs may not yet be published and available for this review.

Within this review, no papers consider the PRO-CTCAE questionnaire to evaluate quality of life. This finding is consistent with a systematic review conducted by Fiteni et al., which found that none of the 15 published phase I trials with PRO endpoints from January 2012 to May 2016 utilised PRO-CTCAE.[2] Interestingly, 2.7% (10/119) of eligible DFOTs used the PRO-CTCAE questionnaire within Lai-Kwon et al.'s[15] ClinicalTrials.gov review from January 2007 to January 2020. Reasons why we have not found the PRO-CTCAE questionnaire being utilised in this review could include: the long lag time between trial completion and publication of trial results, the likelihood that many early phase dose-finding trials might have remained unpublished,[69,70] or that PRO analytical approaches or results within the dose-escalation component of the trials were not reported and hence ineligible. Other reasons could include the associated publication bias for positive trials. Even if PRO-CTCAE data was collected in a trial, a lack of promising results following PRO analysis, or incomplete PRO data might have discouraged authors from including such PRO data in their manuscript. What's more, clinicians who do not distinguish between symptomatic and quality of life based PROs may be ill-prepared to deploy the NCI-PRO-CTCAE measure, specifically designed for patients to solely evaluate their symptoms. This may also explain why the PRO-CTCAE PROM was not evaluated in this methodological review.

Within this review we made no distinction between PROs assessing health related quality of life and symptomatic adverse events. Due to conceptual differences in assessing symptom severity/toxicities compared to health-related quality of life deterioration, the statistical analysis methods and data visualisation techniques presented in Table 3, Fig. 2, and Fig. 3 may have different rationales and methodological focus depending on the type of PRO being analysed. It may be the case that, for some eligible papers in this methodological review, investigators have used symptomatic and quality of life based PROs synonymously. Studies included within this review have considered quality of life PROMS exclusively, symptomatic PROMs exclusively, or a mixture of quality of life and symptomatic measures. For example, the EORTC questionnaires extracted in this review primarily focuses on health related quality of life, whilst other measures such as the AUA and WOMAC assess treatment symptoms.

In conclusion, currently, a minority of trials analyse PROs within dose-finding oncology trials. There is vast heterogeneity in the way PROs are analysed and subsequently presented within publications, this prevents comparison across study findings. Urgent improvement is needed. Increasing the inclusion of PROs in dose

finding trials in a statistically rigorous and consistent manner has the potential to provide meaningful and reliable conclusions of treatment tolerability and improve selection of dose levels that will be evaluated in subsequent trials. This methodological review encourages the introduction of PRO analysis guidelines for dose-finding clinical trials. We recommend further stakeholder engagement is undertaken to ensure consensus driven recommendations for PRO analysis and visualisation. This research can build on the work of the SISAQOL consortium and Snyder et al.'s panel.[66] Future work needs to ensure that rigorous methods are in place to integrate patients' experience and perspectives into trial design and guide optimal analysis of PROs. This will help inform treatment tolerability profiles and dose-selection decisions more efficiently and with less arbitrariness, ultimately leading to patient-centered clinical development.

Contributors

Emily Alger: Conceptualization, Investigation, Formal Analysis, Methodology, Data Curation, Writing–Original Draft, Writing–Review & Editing, Visualization, Project administration. Anna Minchom: Methodology, Validation, Writing–Review & Editing, Supervision. Olalekan Lee Aiyegbusi: Methodology, Validation, Writing–Review & Editing, Supervision. Matthew Schipper: Writing–Review & Editing, Christina Yap: Conceptualization, Investigation, Methodology, Validation, Writing–Review & Editing, Visualisation, Supervision, Project administration, Funding acquisition.

All authors have read and approved the final version of this manuscript. Authors EA, AM, OLA, and CY have verified the underlying data.

Data sharing statement

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declaration of interests

OLA receives funding from the NIHR Birmingham Biomedical Research Centre (BRC), NIHR Applied Research Collaboration (ARC), West Midlands, NIHR Blood and Transplant Research Unit (BTRU) in Precision Transplant and Cellular Therapeutics at the University of Birmingham and University Hospitals Birmingham NHS Foundation, Innovate UK (part of UK Research and Innovation), Gilead Sciences Ltd., Merck, Anthony Nolan, and Sarcoma UK. He declares personal fees from Gilead Sciences, Merck and GlaxoSmithKline outside the submitted work.

AM is funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at the Royal Marsden NHS Foundation Trust. She has served on advisory boards for Janssen Pharmaceuticals, Merck Pharmaceuticals, Takeda Pharmaceuticals and Genmab Pharmaceuticals. Has received honoraria from Chugai Pharmaceuticals, Novartis Oncology, Faron Pharmaceuticals, Bayer Pharmaceuticals, Merck Pharmaceuticals, GSK and Janssen Pharmaceuticals. Has received expenses from Amgen Pharmaceuticals and LOXO Oncology. Has received research funding from Merck Pharmaceuticals and MSD.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.eclinm.2023.102228.

References

1 Araujo D, Greystoke A, Bates S, et al. Oncology phase I trial design and conduct: time for a change - MDICT guidelines 2022. *Ann Oncol*. 2023;34(1):48–60. https://doi.org/10.1016/j.annonc.2022.09.158.

2 Fiteni F, Ray IL, Ousmen A, Isambert N, Anota A, Bonnetain F. Health-related quality of life as an endpoint in oncology phase I trials: a systematic review. *BMC Cancer*. 2019;19(1):361. https://doi.org/10.1186/s12885-019-5579-3.

3 Basch E, Iasonos A, Mcdonough T, et al. Patient versus clinician symptom reporting using the National Cancer Institute common terminology criteria for adverse events: results of a questionnaire-based study. *Lancet Oncol*. 2006;7(11):903–909. https://doi.org/10.1016/S1470-2045(06)70910-X.

4 Wilkie JR, Hochstedler KA, Schipper MJ, et al. Association between physician- and patient-reported symptoms in patients treated with definitive radiation therapy for locally advanced lung cancer in a statewide consortium. *Int J Radiat Oncol Biol Phys*. 2022;112(4):942–950. https://doi.org/10.1016/j.ijrobp.2021.11.024.

5 Jagsi R, Griffith KA, Vicini F, et al. Identifying patients whose symptoms are underrecognized during treatment with breast radiotherapy. *JAMA Oncol*. 2022;8(6):887–894. https://doi.org/10.1001/jamaoncol.2022.0114.

6 Veitch ZW, Shepshelovich D, Gallagher C, et al. Underreporting of symptomatic adverse events in phase I clinical trials. *J Natl Cancer Inst*. 2021;113(8):980–988. https://doi.org/10.1093/jnci/djab015.

7 Postel-Vinay S, Gomez-Roca C, Molife LR, et al. Phase I trials of molecularly targeted agents: should we pay more attention to late toxicities? *J Clin Oncol*. 2011;29(13):1728–1735. https://doi.org/10.1200/jco.2010.31.9236.

8 Kong A, Good J, Kirkham A, et al. Phase I trial of wee1 inhibition with chemotherapy and radiotherapy as adjuvant treatment, and a window of opportunity trial with cisplatin in patients with head and neck cancer: the WISTERIA trial protocol. *BMJ Open*. 2020;10(3):e033009. https://doi.org/10.1136/bmjopen-2019-033009.

9 Walls GM, Oughton JB, Chalmers AJ, et al. Concorde: a phase I platform study of novel agents in combination with conventional radiotherapy in non-small-cell lung cancer. *Clin Transl Radiat Oncol*. 2020;25:61–66. https://doi.org/10.1016/j.ctro.2020.09.006.

10 Bryant AK, Sankar K, Zhao L, et al. De-escalating adjuvant durvalumab treatment duration in stage iii non-small cell lung cancer. *Eur J Cancer*. 2022;171:55–63. https://doi.org/10.1016/j.ejca.2022.04.033.

11 U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcome*. 2006;4(1):79. https://doi.org/10.1186/1477-7525-4-79.

12 Basch E, Yap C. Patient-reported outcomes for tolerability assessment in phase I cancer clinical trials. *J Natl Cancer Inst*. 2021;113(8):943–944. https://doi.org/10.1093/jnci/djab017.

13 Fourie Zirkelbach J, Shah M, Vallejo J, et al. Improving dose-optimization processes used in oncology drug development to minimize toxicity and maximize benefit to patients. *J Clin Oncol*. 2022;40(30):3489–3500. https://doi.org/10.1200/jco.22.00371.

14 Shah MRA, Theoret MR, Pazdur R. *Optimizing dosing in oncology drug development Q&A. Presented at: Friends of Cancer Research Annual Meeting 2021*. 2022.

15 Lai-Kwon J, Yin Z, Minchom A, Yap C. Trends in patient-reported outcome use in early phase dose-finding oncology trials - an analysis of clinicaltrials.Gov. *Cancer Med*. 2021;10(22):7943–7957. https://doi.org/10.1002/cam4.4307.

16 Hamel JF, Saulnier P, Pe M, et al. A systematic review of the quality of statistical methods employed for analysing quality of life data in cancer randomised controlled trials. *Eur J Cancer*. 2017;83:166–176. https://doi.org/10.1016/j.ejca.2017.06.025.

17 Calvert M, Blazeby J, Altman DG, et al. Reporting of patient-reported outcomes in randomized trials: the consort pro extension. *JAMA*. 2013;309(8):814–822. https://doi.org/10.1001/jama.2013.879.

18 Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the spirit-pro extension. *JAMA*. 2018;319(5):483–494. https://doi.org/10.1001/jama.2017.21903.

19 Homer V, Yap C, Bond S, et al. Early phase clinical trials extension to guidelines for the content of statistical analysis plans. *BMJ*. 2022;376:e068177. https://doi.org/10.1136/bmj-2021-068177.

20 Yap C, Bedding A, De Bono J, et al. The need for reporting guidelines for early phase dose-finding trials: dose-finding consort extension. *Nat Med*. 2022;28(1):6–7. https://doi.org/10.1038/s41591-021-01594-1.
21 Solovyeva O, Dimairo M, Weir CJ, et al. Development of consensus-driven SPIRIT and CONSORT extensions for early phase dose-finding trials: the DEFINE study. *BMC Med*. 2023;21:246. https://doi.org/10.1186/s12916-023-02937-0.
22 Coens C, Pe M, Dueck AC, 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–e96. https://doi.org/10.1016/S1470-2045(19)30790-9.
23 Lai-Kwon J, Vanderbeek AM, Minchom A, et al. Using patient-reported outcomes in dose-finding oncology trials: surveys of key stakeholders and the National Cancer Research Institute consumer forum. *Oncol*. 2022;27(9):768–777. https://doi.org/10.1093/oncolo/oyac117.
24 Yap C, Solovyeva O, Yin Z, et al. 53p assessing the reporting quality of early phase dose-finding trials. *Ann Oncol*. 2022;33:S24. https://doi.org/10.1016/j.annonc.2022.01.018.
25 Stega J, Noel MS, Vandell AG, Stega D, Del Priore G, Hoffman S. A first-in-human study of the novel metabolism-based anti-cancer agent sm-88 in subjects with advanced metastatic cancer. *Invest New Drugs*. 2020;38(2):392–401. https://doi.org/10.1007/s10637-019-00758-8.
26 Subbiah V, Grilley-Olson JE, Combest AJ, et al. Phase Ib/II trial of NC-6004 (nanoparticle cisplatin) plus gemcitabine in patients with advanced solid tumors. *Clin Cancer Res*. 2018;24(1):43–51. https://doi.org/10.1158/1078-0432.Ccr-17-1114.
27 Hocking AJ, Farrall AL, Newhouse S, et al. Study protocol of a phase 1 clinical trial establishing the safety of intrapleural administration of liposomal curcumin: curcumin as a palliative treatment for malignant pleural effusion (IPAL-MPE). *BMJ Open*. 2021;11(3):e047075. https://doi.org/10.1136/bmjopen-2020-047075.
28 Sher DJ, Timmerman RD, Nedzi L, et al. Phase 1 fractional dose-escalation study of equipotent stereotactic radiation therapy regimens for early-stage glottic larynx cancer. *Int J Radiat Oncol Biol Phys*. 2019;105(1):110–118. https://doi.org/10.1016/j.ijrobp.2019.03.010.
29 Reiss KA, Herman JM, Armstrong D, et al. A final report of a phase i study of veliparib (ABT-888) in combination with low-dose fractionated whole abdominal radiation therapy (LDFWAR) in patients with advanced solid malignancies and peritoneal carcinomatosis with a dose escalation in ovarian and fallopian tube cancers. *Gynecol Oncol*. 2017;144(3):486–490. https://doi.org/10.1016/j.ygyno.2017.01.016.
30 Goody RB, Mackay H, Pitcher B, et al. Phase 1/2 study of the addition of cisplatin to adjuvant chemotherapy with image guided high-precision radiation therapy for completely resected gastric cancer. *Int J Radiat Oncol Biol Phys*. 2016;96(5):994–1002. https://doi.org/10.1016/j.ijrobp.2016.08.034.
31 Ippolito E, Rinaldi CG, Silipigni S, et al. Hypofractionated radiotherapy with concomitant boost for breast cancer: a dose escalation study. *Br J Radiol*. 2019;92(1095):20180169. https://doi.org/10.1259/bjr.20180169.
32 Guiu B, Jouve J-L, Schmitt A, et al. Intra-arterial idarubicin_lipiodol without embolisation in hepatocellular carcinoma: the LIDA-B phase I trial. *J Hepatol*. 2018;68(6):1163–1171. https://doi.org/10.1016/j.jhep.2018.01.022.
33 Anota A, Boulin M, Dabakuyo-Yonli S, et al. An explorative study to assess the association between health-related quality of life and the recommended phase ii dose in a phase i trial: idarubicin-loaded beads for chemoembolisation of hepatocellular carcinoma. *BMJ Open*. 2016;6(6):e010696. https://doi.org/10.1136/bmjopen-2015-010696.
34 O'rawe M, Wickremesekera AC, Pandey R, et al. Treatment of glioblastoma with re-purposed renin-angiotensin system modulators: results of a phase I clinical trial. *J Clin Neurosci*. 2022;95:48–54. https://doi.org/10.1016/j.jocn.2021.11.023.
35 Smith K, Galazi M, Openshaw MR, et al. The use of transdermal estrogen in castrate-resistant, steroid-refractory prostate cancer. *Clin Genitourin Cancer*. 2020;18(3):e217–e223. https://doi.org/10.1016/j.clgc.2019.09.019.
36 Chawla SP, Goel S, Chow W, et al. A phase 1b dose escalation trial of NC-6300 (nanoparticle epirubicin) in patients with advanced solid tumors or advanced, metastatic, or unresectable soft-tissue sarcoma. *Clin Cancer Res*. 2020;26(16):4225–4232. https://doi.org/10.1158/1078-0432.Ccr-20-0591.
37 Mercier C, Claessens M, Buys MA, et al. Stereotactic ablative radiation therapy to all lesions in patients with oligometastatic cancers: a phase 1 dose-escalation trial. *Int J Radiat Oncol Biol Phys*. 2021;109(5):1195–1205. https://doi.org/10.1016/j.ijrobp.2020.11.066.
38 Diehl CD, Shiban E, Straube C, et al. Neoadjuvant stereotactic radiosurgery for intracerebral metastases of solid tumors (NepoMUC): a phase I dose escalation trial. *Cancer Commun*. 2019;39(1):73. https://doi.org/10.1186/s40880-019-0416-2.
39 Xiao YT, Zhao X, Chang Y, et al. Assessing the safety and feasibility of neoadjuvant hormone and radiation therapy followed by robot-assisted radical prostatectomy for treating locally advanced prostate cancer: protocol for an open-label, dose-escalation, single-centre, phase I clinical trial. *BMJ Open*. 2020;10(11):e038678. https://doi.org/10.1136/bmjopen-2020-038678.
40 Cassier PA, Italiano A, Gomez-Roca C, et al. Long-term clinical activity, safety and patient-reported quality of life for emactuzumab-treated patients with diffuse-type tenosynovial giant-cell tumour. *Eur J Cancer*. 2020;141:162–170. https://doi.org/10.1016/j.ejca.2020.09.038.
41 Rahman MA, Brekke J, Arnesen V, et al. Sequential bortezomib and temozolomide treatment promotes immunological responses in glioblastoma patients with positive clinical outcomes: a phase 1b study. *Immun Inflamm Dis*. 2020;8(3):342–359. https://doi.org/10.1002/iid3.315.
42 Ballas LK, Luo C, Chung E, et al. Phase 1 trial of SBRT to the prostate fossa after prostatectomy. *Int J Radiat Oncol Biol Phys*. 2019;104(1):50–60. https://doi.org/10.1016/j.ijrobp.2018.12.047.
43 Potters L, Rana Z, Lee L, Cox BW. Outcomes of a dose-escalated stereotactic body radiation phase 1 trial for patients with low- and intermediate-risk prostate cancer. *Int J Radiat Oncol Biol Phys*. 2019;104(2):334–342. https://doi.org/10.1016/j.ijrobp.2019.01.092.
44 Den RB, Greenspan J, Doyle LA, et al. A phase IB clinical trial of 15 Gy HDR brachytherapy followed by hypofractionated/SBRT in the management of intermediate-risk prostate cancer. *Brachytherapy*. 2020;19(3):282–289. https://doi.org/10.1016/j.brachy.2020.02.008.
45 Alayed Y, Loblaw A, Chu W, et al. Stereotactic body radiation therapy boost for intermediate-risk prostate cancer: a phase 1 dose-escalation study. *Int J Radiat Oncol Biol Phys*. 2019;104(5):1066–1073. https://doi.org/10.1016/j.ijrobp.2019.04.006.
46 Sampath S, Frankel P, Vecchio BD, et al. Stereotactic body radiation therapy to the prostate bed: results of a phase 1 dose-escalation trial. *Int J Radiat Oncol Biol Phys*. 2020;106(3):537–545. https://doi.org/10.1016/j.ijrobp.2019.11.005.
47 Moe A, Liow E, Redfern A, et al. A phase I open label dose-escalation study to evaluate the tolerability, safety and immunological efficacy of sub-urothelial durvalumab injection in adults with muscle-invasive or high-risk non-muscle-invasive bladder cancer (SUBDUE-1, SUB-urothelial DUrvalumab injection-1 study): clinical trial protocol. *BJU Int*. 2021;128(S1):9–17. https://doi.org/10.1111/bju.15365.
48 Zamagni A, Buwenge M, Macchia G, et al. Accelerated middle half body radiotherapy in bone metastases from prostate cancer: a phase I study (sharon project). *Anticancer Res*. 2019;39(9):5065–5069. https://doi.org/10.21873/anticanres.13699.
49 Farina E, Capuccini J, Macchia G, et al. Phase I-II study of short-course accelerated radiotherapy (sharon) for palliation in head and neck cancer. *Anticancer Res*. 2018;38(4):2409–2414.
50 Correa RJM, Ahmad B, Warner A, et al. A prospective phase I dose-escalation trial of stereotactic ablative radiotherapy (SABR) as an alternative to cytoreductive nephrectomy for inoperable patients with metastatic renal cell carcinoma. *Radiat Oncol*. 2018;13(1):47. https://doi.org/10.1186/s13014-018-0992-3.
51 Muller DA, Wages NA, Wilson DD, et al. Stat rad: prospective dose escalation clinical trial of single fraction Scan-Plan-QA-Treat stereotactic body radiation therapy for painful osseous metastases. *Pract Radiat Oncol*. 2020;10(6):e444–e451. https://doi.org/10.1016/j.prro.2020.03.008.
52 Schmidt M, Lenhard H, Hoenig A, et al. Tumor suppression, dose-limiting toxicity and wellbeing with the fetal estrogen estetrol in patients with advanced breast cancer. *J Cancer Res Clin Oncol*. 2021;147(6):1833–1842. https://doi.org/10.1007/s00432-020-03472-8.
53 Heiss JD, Jamshidi A, Shah S, et al. Phase I trial of convection-enhanced delivery of IL13-pseudomonas toxin in children with

diffuse intrinsic pontine glioma. *J Neurosurg Pediatr*. 2018;23(3):333–342. https://doi.org/10.3171/2018.9.Peds17225.
54 Kim HS, Shaib WL, Zhang C, et al. Phase 1b study of pasireotide, everolimus, and selective internal radioembolization therapy for unresectable neuroendocrine tumors with hepatic metastases. *Cancer*. 2018;124(9):1992–2000. https://doi.org/10.1002/cncr.31192.
55 Vatandoust S, Bright T, Roy AC, et al. Phase I open-label trial of intraperitoneal paclitaxel in combination with intravenous cisplatin and oral capecitabine in patients with advanced gastric cancer and peritoneal metastases (IPGP study): study protocol. *BMJ Open*. 2019;9(5):e026732. https://doi.org/10.1136/bmjopen-2018-026732.
56 Veldhuijzen Van Zanten SEM, El-Khouly FE, Jansen MHA, et al. A phase I/II study of gemcitabine during radiotherapy in children with newly diagnosed diffuse intrinsic pontine glioma. *J Neuro Oncol*. 2017;135(2):307–315. https://doi.org/10.1007/s11060-017-2575-9.
57 Merrick GS, Butler WM, Dorsey AT, Galbreath RW, Blatt H, Lief JH. Rectal function following prostate brachytherapy. *Int J Radiat Oncol Biol Phys*. 2000;48(3):667–674. https://doi.org/10.1016/S0360-3016(00)00698-2.
58 Gong J, Thomassian S, Kim S, et al. Phase I trial of Bermekimab with nanoliposomal irinotecan and 5-fluorouracil/folinic acid in advanced pancreatic ductal adenocarcinoma. *Sci Rep*. 2022;12(1):15013. https://doi.org/10.1038/s41598-022-19401-3.
59 Cohen AL, Anker CJ, Johnson B, et al. Repeat radiation with bevacizumab and minocycline in bevacizumab-refractory high grade gliomas: a prospective phase 1 trial. *J Neuro Oncol*. 2020;148(3):577–585. https://doi.org/10.1007/s11060-020-03551-3.
60 Page MJ, Mckenzie JE, Bossuyt PM, et al. The prisma 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. https://doi.org/10.1136/bmj.n71.
61 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;106(9):dju244. https://doi.org/10.1093/jnci/dju244.
62 Regnault A, Loubert A, Gorsh B, et al. A toolbox of different approaches to analyze and present PRO-CTCAE data in oncology studies. *J Natl Cancer Inst*. 2023;115(5):586–596. https://doi.org/10.1093/jnci/djad018.
63 Flynn KE, Dombeck CB, Dewitt EM, Schulman KA, Weinfurt KP. Using item banks to construct measures of patient reported outcomes in clinical trials: investigator perceptions. *Clin Trials*. 2008;5(6):575–586. https://doi.org/10.1177/1740774508098414.
64 Kluetz PG, Kanapuru B, Lemery S, et al. Informing the tolerability of cancer treatments using patient-reported outcome measures: summary of an FDA and critical path institute workshop. *Value Health*. 2018;21(6):742–747. https://doi.org/10.1016/j.jval.2017.09.009.
65 Pe M, Dorme L, Coens C, 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–e469. https://doi.org/10.1016/S1470-2045(18)30418-2.
66 Snyder C, Smith K, Holzner B, Rivera YM, Bantug E, Brundage M. Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data. *Qual Life Res*. 2019;28(2):345–356. https://doi.org/10.1007/s11136-018-2020-3.
67 Brundage M, Blackford A, Tolbert E, Smith K, Bantug E, Snyder C. Presenting comparative study pro results to clinicians and researchers: beyond the eye of the beholder. *Qual Life Res*. 2018;27(1):75–90. https://doi.org/10.1007/s11136-017-1710-6.
68 Brundage MD, Smith KC, Little EA, Bantug ET, Snyder CF. Communicating patient-reported outcome scores using graphic formats: results from a mixed-methods evaluation. *Qual Life Res*. 2015;24(10):2457–2472. https://doi.org/10.1007/s11136-015-0974-y.
69 Devito NJ, Goldacre B. Evaluation of compliance with legal requirements under the FDA amendments act of 2007 for timely registration of clinical trials, data verification, delayed reporting, and trial document submission. *JAMA Intern Med*. 2021;181(8):1128–1130. https://doi.org/10.1001/jamainternmed.2021.2036.
70 Shepshelovich D, Goldvaser H, Wang L, Abdul Razak AR. Comparison of published and unpublished phase I clinical cancer trials: an analysis of the clinicltrials.Gov database. *Invest New Drugs*. 2018;36(5):933–938. https://doi.org/10.1007/s10637-017-0549-6.