

BMJ Open SPIRIT-PRO Extension explanation and elaboration: guidelines for inclusion of patient-reported outcomes in protocols of clinical trials

Melanie Calvert ^{1,2,3,4,5}, Madeleine King,⁶ Rebecca Mercieca-Bebber ⁷, Olalekan Aiyegbusi,^{1,3} Derek Kyte,¹ Anita Slade,^{1,4} An-Wen Chan,⁸ E Basch,⁹ Jill Bell,¹⁰ Antonia Bennett,¹¹ Vishal Bhatnagar,¹² Jane Blazeby ^{13,14}, Andrew Bottomley,¹⁵ Julia Brown,¹⁶ Michael Brundage,¹⁷ Lisa Campbell,¹⁸ Joseph C Cappelleri,¹⁹ Heather Draper,²⁰ Amylou C Dueck,²¹ Carolyn Ells,²² Lori Frank,²³ Robert M Golub,²⁴ Ingolf Griebisch,²⁵ Kirstie Haywood,²⁶ Amanda Hunn ²⁷, Bellinda King-Kallimanis,¹² Laura Martin,²⁸ Sandra Mitchell,²⁹ Thomas Morel,³⁰ Linda Nelson,³¹ Josephine Norquist,³² Daniel O'Connor,¹⁸ Michael Palmer,³³ Donald Patrick,³⁴ Gary Price,¹ Antoine Regnault,³⁵ Ameeta Retzer,¹ Dennis Revicki,³⁶ Jane Scott,³⁷ Richard Stephens,³⁸ Grace Turner ^{1,4}, Antonia Valakas,³⁹ Galina Velikova,⁴⁰ Maria von Hildebrand,¹ Anita Walker,¹ Lari Wenzel⁴¹

To cite: Calvert M, King M, Mercieca-Bebber R, *et al*. SPIRIT-PRO Extension explanation and elaboration: guidelines for inclusion of patient-reported outcomes in protocols of clinical trials. *BMJ Open* 2021;**11**:e045105. doi:10.1136/bmjopen-2020-045105

► Prepublication history and supplemental material for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-045105>).

Received 24 September 2020
Revised 21 December 2020
Accepted 08 January 2021



► <http://dx.doi.org/10.1136/bmjopen-2020-046450>



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Professor Melanie Calvert;
M.Calvert@bham.ac.uk

ABSTRACT

Patient-reported outcomes (PROs) are used in clinical trials to provide valuable evidence on the impact of disease and treatment on patients' symptoms, function and quality of life. High-quality PRO data from trials can inform shared decision-making, regulatory and economic analyses and health policy. Recent evidence suggests the PRO content of past trial protocols was often incomplete or unclear, leading to research waste. To address this issue, international, consensus-based, PRO-specific guidelines were developed: the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT)-PRO Extension. The SPIRIT-PRO Extension is a 16-item checklist which aims to improve the content and quality of aspects of clinical trial protocols relating to PRO data collection to minimise research waste, and ultimately better inform patient-centred care. This SPIRIT-PRO explanation and elaboration (E&E) paper provides information to promote understanding and facilitate uptake of the recommended checklist items, including a comprehensive protocol template. For each SPIRIT-PRO item, we provide a detailed description, one or more examples from existing trial protocols and supporting empirical evidence of the item's importance. We recommend this paper and protocol template be used alongside the SPIRIT 2013 and SPIRIT-PRO Extension paper to optimise the transparent development and review of trial protocols with PROs.

BACKGROUND

Clinical trial protocols are essential documents intended to include the study rationale, intervention, trial design methods, study processes, outcomes, sample size, data

Strengths and limitations of this study

- The Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT)-patient-reported outcome (PRO) Extension aims to improve the completeness and transparency of trial protocols where PROs are a primary or key secondary outcomes and was developed following Enhancing Quality and Transparency of Health Research Network Guidance.
- This explanation and elaboration paper provides information to promote understanding and facilitate uptake of the recommended PRO protocol SPIRIT-PRO checklist items for clinical trials.
- A comprehensive protocol template and selected examples from existing trial protocols are provided to facilitate implementation.
- The protocol template and explanation and elaboration paper were developed with multistakeholder international input including: trialists, PRO methodologists, psychometricians, patient partners, industry representatives, journal editors, regulators and ethicists.
- Although the guidance is limited in focus to clinical trials, many of the SPIRIT-PRO items may also provide useful prompts about PRO content for cohort studies and other non-randomised designs.

collection procedures, proposed analyses and ethical considerations. Provision of sufficient detail is necessary to enable the research team to conduct a high-quality, reproducible study. It also facilitates external appraisal of



the scientific, methodological and ethical rigour of the trial by relevant stakeholders.^{1,2} Although trial protocols serve as the foundation for study planning, conduct, reporting and appraisal, they vary greatly in content and quality.^{1,2} Appraisals of the patient-reported outcome (PRO) content of over 350 past trial protocols revealed that many protocols lack specific information needed for high-quality PRO data collection and evidence generation (online supplemental 1).^{3–5} As a result, research personnel and potential research participants may not appreciate the purpose of PRO data collection,⁶ and the need for standardised PRO assessment methods. This may result in high levels of missing data and poor quality or non-reporting of PRO trial results, which may hinder the potential for PRO evidence to be used in regulatory decision-making, health policy and clinical care.^{6–8} For example, a recent review of cancer portfolio trials illustrates this point; recommended PRO protocol content was frequently not addressed and PRO data from 61 trials, including 49 568 participants, was unpublished.⁹ Another trial also cited poor PRO completion rates as the reason for not publishing PRO data—and the corresponding trial protocol included only sparse guidance related to the PRO study.⁷

In 2013, core protocol guidelines applicable to all types of trials was published based on expert consensus and research evidence, in the form of the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) 2013 statement. Its corresponding SPIRIT 2013 explanation and elaboration (E&E) paper provides important information to promote full understanding of, and assist protocol writers to implement, the 33 checklist recommendations.^{1,2} However, SPIRIT 2013 does not provide specific recommendations about PRO endpoints. PROs can provide valuable information on the risks, benefits and tolerability of an intervention. PRO data are intrinsically subjective, requiring completion by patient-participants within a specific time frame and, as a result, present a range of scientific and logistical challenges for researchers which should be addressed in the trial protocol.^{6,10–12}

To address this issue, international stakeholders worked to develop the SPIRIT-PRO Extension, with the aim of improving PRO content of trial protocols and supporting documents, for use in conjunction with SPIRIT 2013 Guidelines and E&E papers.^{1,2} The SPIRIT-PRO Extension was published in 2018 and comprises 11 extensions (new, PRO-specific items) and 5 elaborations (an elaboration of an existing SPIRIT 2013 item as applied to clinical trials assessing PROs) recommended for inclusion in clinical trial protocols that have PROs as primary or key secondary outcomes (table 1).¹⁰ The SPIRIT-PRO Extension paper reports the 16 items and describes the methods used to develop the checklist, but does not provide detailed implementation instructions or examples. This SPIRIT-PRO E&E paper aims to promote understanding of the guidelines, provide real examples of SPIRIT-PRO items being addressed from a range of

different trials and facilitate uptake of the recommended checklist items. A table of contents detailing where to find an example and explanation of each item is provided in table 2. In addition, we describe the development of a new PRO protocol template (online supplemental file 2) for use in protocol development. Additional information and resources regarding the SPIRIT Initiative are available on the SPIRIT website (www.spirit-statement.org).

The development of the SPIRIT-PRO Extension followed the Enhancing Quality and Transparency of Health Research Network's methodological framework for guideline development,¹³ and has been published elsewhere.¹⁰ Briefly, these methods included:

1. a systematic review of existing PRO-specific protocol guidelines to generate the list of potential PRO-specific protocol items;¹⁴
2. refinements to the list and removal of duplicate items by the International Society for Quality of Life Research (ISOQOL) Protocol Checklist Taskforce;
3. an international stakeholder survey of trial research personnel, PRO methodologists, health economists, psychometricians, patient advocates, funders, industry representatives, journal editors, policy makers, ethicists and researchers responsible for evidence synthesis (distributed by 38 international partner organisations);
4. an international Delphi exercise;
5. a consensus meeting in May 2017 to finalise the guidelines and implementation strategy.

International stakeholders provided feedback on the final wording of the SPIRIT-PRO Extension during a final 3-week consultation period. Following minor edits, the guidelines were finalised and agreed by the SPIRIT-PRO Group.¹⁰

DEVELOPMENT OF THE PRO PROTOCOL TEMPLATE

A PRO protocol template was developed to support implementation of the SPIRIT-PRO guidance (ethical approval ERN_19–0939). The draft template was reviewed by members of the project team and broader SPIRIT-PRO Group, including patient partners. In addition, an international advisory group (IAG), comprising global PRO leads from major pharmaceutical companies, regulators and academics, was convened to review and provide additional feedback on the template. Teleconference meetings were held with members of the SPIRIT-PRO Group and the IAG to discuss the feedback received. Based on the feedback, the template was revised and sent to all for final comments. After a final consultation period, the PRO protocol template was revised and finalised.

PATIENT AND PUBLIC INVOLVEMENT

Patient partners were involved in the design, conduct, reporting and dissemination plans of our research, including development of the SPIRIT-PRO Extension, the E&E paper, protocol template, tools to support implementation by patient partners and are included as coauthors.¹⁵

Table 1 SPIRIT 2013 and SPIRIT-PRO Extension checklist: recommended items to address in a clinical trial protocol

| SPIRIT section | SPIRIT item no. | SPIRIT item description | SPIRIT-PRO item no. | SPIRIT-PRO extension or elaboration item description | Page number of example/explanation |
|--|-----------------|--|--------------------------|--|------------------------------------|
| Administrative information | | | | | |
| Title | 1 | Descriptive title identifying the study design, population, interventions and, if applicable, trial acronym | | | |
| Trial registration | 2a | Trial identifier and registry name (if not yet registered, name of intended registry) | | | |
| Protocol version | 2b | All items from the WHO Trial Registration Data Set | | | |
| Funding | 3 | Date and version identifier | | | |
| Roles and responsibilities | 4 | Sources and types of financial, material and other support | | | |
| | 5a | Names, affiliations and roles of protocol contributors | SPIRIT5a-PRO Elaboration | Specify the individual(s) responsible for the PRO content of the trial protocol. | 11–12 |
| | 5b | Name and contact information for the trial sponsor | | | |
| | 5c | Role of study sponsor and funders, if any, in study design; collection, management, analysis and interpretation of data; writing of the report and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities | | | |
| | 5d | Composition, roles and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team and other individuals or groups overseeing the trial, if applicable (see item 21a for data monitoring committee) | | | |
| Introduction | | | | | |
| Background and rationale | 6a | Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention | SPIRIT6a-PRO Extension | Describe the PRO-specific research question and rationale for PRO assessment and summarise PRO findings in relevant studies. | 12 |
| Objectives | 6b | Explanation for choice of comparators | | | |
| | 7 | Specific objectives or hypotheses | SPIRIT7-PRO Extension | State specific PRO objectives or hypotheses (including relevant PRO concepts/domains). | 12–14 |
| Trial design | 8 | Description of trial design, including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio and framework (eg, superiority, equivalence, non-inferiority, exploratory) | | | |
| Methods: participants, interventions and outcomes | | | | | |
| Study setting | 9 | Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected; reference to where list of study sites can be obtained | | | |

Continued

Table 1 Continued

| SPiRiT section | SPiRiT item no. | SPiRiT item description | SPiRiT-PRO item no. | SPiRiT-PRO extension or elaboration item description | Page number of example/explanation |
|----------------------|-----------------|---|--------------------------|--|------------------------------------|
| Eligibility criteria | 10 | Inclusion and exclusion criteria for participants; if applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists) | SPiRiT10-PRO Extension | Specify any PRO-specific eligibility criteria (eg, language/reading requirements or preredomatisation completion of PRO). If PROs will not be collected from the entire study sample, provide a rationale and describe the method for obtaining the PRO subsample. | 14 |
| Interventions | 11a | Interventions for each group with sufficient detail to allow replication, including how and when they will be administered | | | |
| | 11b | Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request or improving/worsening disease) | | | |
| | 11c | Strategies to improve adherence to intervention protocols and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests) | | | |
| | 11d | Relevant concomitant care and interventions that are permitted or prohibited during the trial | | | |
| Outcomes | 12 | Primary, secondary and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion) and timepoint for each outcome; explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended | SPiRiT12-PRO Extension | Specify the PRO concepts/domains used to evaluate the intervention (eg, overall health-related quality of life, specific domain, specific symptom) and, for each one, the analysis metric (eg, change from baseline, final value, time to event) and the principal timepoint or period of interest. | 14–16 |
| Participant timeline | 13 | Time schedule of enrolment, interventions (including any run-ins and washouts), assessments and visits for participants; a schematic diagram is highly recommended (see figure in Chan <i>et al</i>) ^{1,2} | SPiRiT13-PRO Extension | Include a schedule of PRO assessments, providing a rationale for the timepoints and justifying if the initial assessment is not preredomatisation. Specify time-windows, whether PRO collection is prior to clinical assessments, and, if using multiple questionnaires, whether order of administration will be standardised. | 16–17 |
| Sample size | 14 | Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations | SPiRiT14-PRO Elaboration | When a PRO is the primary endpoint, state the required sample size (and how it was determined) and recruitment target (accounting for expected loss to follow-up). If sample size is not established based on the PRO endpoint, then discuss the power of the principal PRO analyses. | 17–19 |

Continued

Table 1 Continued

| SPRIT section | SPRIT item no. | SPRIT item description | SPRIT-PRO item no. | SPRIT-PRO extension or elaboration item description | Page number of example/explanation |
|--|----------------|--|--------------------|---|------------------------------------|
| Recruitment | 15 | Strategies for achieving adequate participant enrolment to reach target sample size | | | |
| Methods: assignment of interventions (for clinical trials) | | | | | |
| Allocation | | | | | |
| Sequence generation | 16a | Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions | | | |
| Allocation concealment mechanism | 16b | Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned | | | |
| Implementation | 16c | Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions | | | |
| Blinding (masking) | 17a | Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts) and how | | | |
| | 17b | If blinded, circumstances under which unblinding is permissible and procedure for revealing a participant's allocated intervention during the trial | | | |
| Methods: data collection, management and analysis | | | | | |

Continued

Table 1 Continued

| SPIRIT section | SPIRIT item no. | SPIRIT item description | SPIRIT-PRO item no. | SPIRIT-PRO extension or elaboration item description | Page number of example/explanation |
|-------------------------|-----------------|---|---------------------------------|--|------------------------------------|
| Data collection methods | 18a | Plans for assessment and collection of outcome, baseline and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known; reference to where data collection forms can be found, if not in the protocol | SPIRIT-18a (i)-PRO Extension | Justify the PRO instrument to be used and describe domains, number of items, recall period and instrument scaling and scoring (eg, range and direction of scores indicating a good or poor outcome). Evidence of PRO instrument measurement properties, interpretation guidelines and patient acceptability and burden should be provided or cited if available, ideally in the population of interest. State whether the measure will be used in accordance with any user manual and specify and justify deviations if planned. | 19–20 |
| | | | SPIRIT-18a (ii)-PRO Extension | Include a data collection plan outlining the permitted mode(s) of administration (eg, paper, telephone, electronic, other) and setting (eg, clinic, home, other). | 20–21 |
| | | | SPIRIT-18a (iii)-PRO Extension | Specify whether more than one language version will be used and state whether translated versions have been developed using currently recommended methods. | 21–22 |
| | | | SPIRIT-18a (iv)-PRO Extension | When the trial context requires someone other than a trial participant to answer on his or her behalf (a proxy-reported outcome), state and justify the use of a proxy respondent. Provide or cite evidence of the validity of proxy assessment if available. | 22–23 |
| | 18b | Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols | SPIRIT-18b (i)-PRO Extension | Specify PRO data collection and management strategies for minimising avoidable missing data. | 23–24 |
| | | | SPIRIT-18b (ii)-PRO Elaboration | Describe the process of PRO assessment for participants who discontinue or deviate from the assigned intervention protocol. | 24–25 |
| Data management | 19 | Plans for data entry, coding, security and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values); reference to where details of data management procedures can be found, if not in the protocol | | | |

Continued

Table 1 Continued

| SPIRIT section | SPIRIT item no. | SPIRIT item description | SPIRIT-PRO item no. | SPIRIT-PRO extension or elaboration item description | Page number of example/explanation |
|---------------------------------|-----------------|---|------------------------------|--|------------------------------------|
| Statistical methods | 20a | Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol | SPIRIT20a-PRO Elaboration | State PRO analysis methods, including any plans for addressing multiplicity/type I (α) error. | 25–26 |
| | 20b | Methods for any additional analyses (eg, subgroup and adjusted analyses) | | | |
| | 20c | Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis) and any statistical methods to handle missing data (eg, multiple imputation) | SPIRIT20c-PRO Elaboration | State how missing data will be described and outline the methods for handling missing items or entire assessments (eg, approach to imputation and sensitivity analyses). | 26–27 |
| Methods: monitoring | | | | | |
| Harms | 21a | Composition of data monitoring committee; summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests and reference to where further details about its charter can be found, if not in the protocol (alternatively, an explanation of why a data monitoring committee is not needed) | | | |
| | 21b | Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial | | | |
| Auditing | 22 | Plans for collecting, assessing, reporting and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct | SPIRIT22-PRO Extension | State whether or not PRO data will be monitored during the study to inform the clinical care of individual trial participants and, if so, how this will be managed in a standardised way. Describe how this process will be explained to participants; for example, in the participant information sheet and consent form. | 27–29 |
| | 23 | Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and sponsor(s) | | | |
| Ethics and dissemination | | | | | |
| Research ethics approval | 24 | Plans for seeking research ethics committee/institutional review board approval | | | |
| | 25 | Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, research ethics committees/institutional review boards, trial participants, trial registries, journals, regulators) | | | |

Continued

| Table 1 Continued | | SPIRIT item no. | SPIRIT item description | SPIRIT-PRO item no. | SPIRIT-PRO extension or elaboration item description | Page number of example/explanation |
|-------------------------------|-----|--|-------------------------|---------------------|--|------------------------------------|
| Consent or assent | 26a | Who will obtain informed consent or assent from potential trial participants or authorised surrogates and how (see item 32) | | | | |
| | 26b | Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable | | | | |
| Confidentiality | 27 | How personal information about potential and enrolled participants will be collected, shared and maintained to protect confidentiality before, during and after the trial | | | | |
| Declaration of interests | 28 | Financial and other competing interests for principal investigators for the overall trial and each study site | | | | |
| Access to data | 29 | Statement of who will have access to the final trial data set and disclosure of contractual agreements that limit such access for investigators | | | | |
| Ancillary and post-trial care | 30 | Provisions, if any, for ancillary and post-trial care and for compensation to those who are harmed by trial participation | | | | |
| Dissemination policy | 31a | Plans for investigators and sponsor(s) to communicate trial results to participants, healthcare professionals, the public and other relevant groups (eg, via publication, reporting in results databases or other data-sharing arrangements), including any publication restrictions | | | | |
| | 31b | Authorship eligibility guidelines and any intended use of professional writers | | | | |
| | 31c | Plans, if any, for granting public access to the full protocol, participant-level data set and statistical code | | | | |
| Appendixes | | | | | | |
| Informed consent materials | 32 | Model consent form and other related documentation given to participants and authorised surrogates | | | | |
| Biological specimens | 33 | Plans for collection, laboratory evaluation and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable | | | | |

PRO, patient-reported outcome; SPIRIT, Standard Protocol Items: Recommendations for Interventional Trials.

Table 2 Table of contents and resources

| SPIRIT-PRO item no. | Item title | SPIRIT-PRO extension or elaboration item description | Page number of example/explanation |
|-----------------------------|--|--|------------------------------------|
| 5a | SPIRIT5a-PRO Elaboration | Specify the individual(s) responsible for the PRO content of the trial protocol. | 11–12 |
| 6a | SPIRIT6a-PRO Extension | Describe the PRO-specific research question and rationale for PRO assessment and summarise PRO findings in relevant studies. | 12 |
| 7 | SPIRIT7-PRO Extension | State-specific PRO objectives or hypotheses (including relevant PRO concepts/domains). | 12–14 |
| 10 | SPIRIT10-PRO Extension | Specify any PRO-specific eligibility criteria (eg, language/reading requirements or prerandomisation completion of PRO). If PROs will not be collected from the entire study sample, provide a rationale and describe the method for obtaining the PRO subsample. | 14 |
| 12 | SPIRIT12-PRO Extension | Specify the PRO concepts/domains used to evaluate the intervention (eg, overall health-related quality of life, specific domain, specific symptom) and, for each one, the analysis metric (eg, change from baseline, final value, time to event) and the principal timepoint or period of interest. | 14–16 |
| 13 | SPIRIT13-PRO Extension | Include a schedule of PRO assessments, providing a rationale for the timepoints and justifying if the initial assessment is not prerandomisation. Specify time-windows, whether PRO collection is prior to clinical assessments, and, if using multiple questionnaires, whether order of administration will be standardised. | 16–17 |
| 14 | SPIRIT14-PRO Elaboration | When a PRO is the primary endpoint, state the required sample size (and how it was determined) and recruitment target (accounting for expected loss to follow-up). If sample size is not established based on the PRO endpoint, then discuss the power of the principal PRO analyses. | 17–19 |
| 18a | SPIRIT-18a (i)-PRO Extension | Justify the PRO instrument to be used and describe domains, number of items, recall period and instrument scaling and scoring (eg, range and direction of scores indicating a good or poor outcome). Evidence of PRO instrument measurement properties, interpretation guidelines and patient acceptability and burden should be provided or cited if available, ideally in the population of interest. State whether the measure will be used in accordance with any user manual and specify and justify deviations if planned. | 19–20 |
| | SPIRIT-18a (ii)-PRO Extension | Include a data collection plan outlining the permitted mode(s) of administration (eg, paper, telephone, electronic, other) and setting (eg, clinic, home, other). | 20–21 |
| | SPIRIT-18a (iii)-PRO Extension | Specify whether more than one language version will be used and state whether translated versions have been developed using currently recommended methods. | 21–22 |
| | SPIRIT-18a (iv)-PRO Extension | When the trial context requires someone other than a trial participant to answer on his or her behalf (a proxy-reported outcome), state and justify the use of a proxy respondent. Provide or cite evidence of the validity of proxy assessment if available. | 22–23 |
| 18b | SPIRIT-18b (i)-PRO Extension | Specify PRO data collection and management strategies for minimising avoidable missing data. | 23–24 |
| | SPIRIT-18b (ii)-PRO Elaboration | Describe the process of PRO assessment for participants who discontinue or deviate from the assigned intervention protocol. | 24–25 |
| 20a | SPIRIT-20a-PRO Elaboration | State PRO analysis methods, including any plans for addressing multiplicity/type I (α) error. | 25–26 |
| 20c | SPIRIT-20c-PRO Elaboration | State how missing data will be described and outline the methods for handling missing items or entire assessments (eg, approach to imputation and sensitivity analyses). | 26–27 |
| 22 | SPIRIT-22-PRO Extension | State whether or not PRO data will be monitored during the study to inform the clinical care of individual trial participants and, if so, how this will be managed in a standardised way. Describe how this process will be explained to participants; for example, in the participant information sheet and consent form. | 27–29 |
| Additional resources | | | |
| | Table 1. SPIRIT 2013 and SPIRIT-PRO Extension checklist: recommended items to address in a clinical trial protocol | | 3–8 |
| | Glossary of terms | | 10–11 |
| | Protocol template | | Online supplemental file 2 |

PRO, patient-reported outcome; SPIRIT, Standard Protocol Items: Recommendations for Interventional Trials.

GLOSSARY**Concept**

'The specific measurement goal (ie, the thing that is to be measured by a PRO instrument). In clinical trials, a PRO instrument can be used to measure the effect of a medical intervention on one or more concepts. PRO concepts represent aspects of how patients function or feel related to a health condition or its treatment'.¹⁶

Domain

'A subconcept represented by a score of an instrument that measures a larger concept comprised multiple domains. For example, psychological function is a larger concept with multiple domains (emotional and cognitive function) that are measured by relevant items'.¹⁶

Endpoint*

The variable to be analysed. It is a precisely defined variable intended to reflect an outcome of interest that is statistically analysed to address a particular research question. A precise definition of an endpoint typically specifies the type of assessments made, the timing of those assessments, the assessment tools used and possibly other details, as applicable, such as how multiple assessments within an individual are to be combined¹⁷ (eg, change from baseline at 6 weeks in mean fatigue score).¹⁸

Health-related quality of life

'A multidimensional concept that usually includes self-report of the way in which physical, emotional, social or other domains of well-being are affected by a disease or its treatment'.¹⁹

Important or key secondary PROs/endpoints

Some PRO measures (particularly health-related quality-of-life measures) are multidimensional, producing several domain-specific outcome scales; for example, pain, fatigue, physical function, psychological distress. For any particular trial, it is likely that a particular PRO or PRO domain(s) will be more relevant than others, reflecting the expected effect(s) of the trial intervention(s) in the target patient population. These relevant PRO(s) and/or domain(s) may additionally constitute the important or key secondary PROs (identified a priori and specified as such in the trial protocol and statistical analysis plan) and will be the focus of hypothesis testing. In a regulatory environment, these outcomes may support a labelling claim. Because these outcomes are linked with hypotheses (Consolidated Standards of Reporting Trials (CONSORT)-PRO Extension 2b),¹⁹ they may be subject to p value adjustment (or 'α spending'). Beyond efficacy/effectiveness, PROs may also be used to capture and provide evidence of safety and tolerability (eg, using the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events).²⁰

Instrument

'A means to capture data (eg, a questionnaire) plus all the information and documentation that supports its

use. Generally, that includes clearly defined methods and instructions for administration or responding, a standard format for data collection and well-documented methods for scoring, analysis and interpretation of results in the target patient population'.¹⁶

Intervention/Treatment

A process or action that is the focus of a clinical study. Interventions include drugs, medical devices, procedures, vaccines and other products that are either investigational or already available. Interventions can also include non-invasive approaches, such as education or modifying diet and exercise.²¹

Item

'An individual question, statement or task (and its standardised response options) that is evaluated by the patient to address a particular concept'.¹⁶

Observer-reported outcome

'A measurement based on a report of observable signs, events or behaviours related to a patient's health condition by someone other than the patient of a healthcare professional'.²²

Outcome*

The variable to be measured. It is the measurable characteristic that is influenced or affected by an individual's baseline state or an intervention as in a clinical trial or other exposure¹⁷ (eg, a fatigue score).

Patient-reported outcome

A PRO is 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 and may include patient assessments of health status, quality of life or symptoms.^{16 19} PROs are assessed by self-reported questionnaires, referred to as PRO measures or instruments.¹⁷

Primary outcome

The most important outcome in a trial, prespecified in the protocol, providing the most clinically relevant evidence directly related to the primary objective of the trial.

Proxy-reported outcome

'A measurement based on a report by someone other than the patient reporting as if he or she is the patient'.¹⁶

Secondary outcomes

Outcomes prespecified in the protocol to assess additional effects of the intervention; some PROs may be identified as important or key secondary outcomes.

SPIRIT Elaboration item

An elaboration of an existing SPIRIT item as applied to a specific context; in this instance, as applied to clinical trials assessing PROs.

SPIRIT-PRO Extension item

An additional checklist item describing PRO protocol content to address an aspect of PRO assessment that is not adequately covered by SPIRIT, as judged by available evidence and expert opinion.

Time-window

A predefined time frame before and after the protocol-specified PRO assessment timepoint, whereby the result would still be deemed to be clinically relevant.²³

*The terms outcome and endpoint are often used interchangeably, although this is not always consistent with the range of definitions available. For the definitions included in this glossary, an endpoint is defined from PRO data (ie, the outcome) by fully specifying four components: measurement variable (eg, fatigue ‘in the past week’ as measured by the QLQ-C30), analysis metric (eg, change in fatigue from baseline, final fatigue value, time to clinically important increase in fatigue (and ‘event’), method of aggregation (eg, median fatigue, proportion of patients with severe fatigue, proportion of patients with clinically important change in fatigue) and timepoint. Note that using these definitions, several endpoints can be defined from the same outcome source data, revealing the distinction and relationship between ‘outcome’ and ‘endpoint’ for PROs.

PURPOSE AND DEVELOPMENT OF THE EXPLANATION AND ELABORATION PAPER AND PRO PROTOCOL TEMPLATE

The SPIRIT-PRO Extension, this E&E and the included PRO protocol template are intended to guide the development of trial protocols for ethical review, where PROs are a primary or key secondary outcome, including single-arm and multi-arm trials. We recommend that authors also consider inclusion of checklist items when PROs are exploratory in nature, as appropriate. Protocols may be formatted in accordance with local requirements, however they need to address the SPIRIT-PRO items completely and transparently. The examples provided in this E&E document and protocol template are not intended to be prescriptive about how information is included in protocols, nor how trials be conducted. Trialists may, for example, wish to include a PRO-specific, dedicated section in the protocol with content informed by the SPIRIT-PRO checklist, while others may wish to add PRO content to existing sections of the protocol.

Modelled after other reporting guidelines,^{2 24 25} this E&E paper presents each checklist item with at least one example from a trial protocol, followed by an explanation of the rationale and main issues to address, to facilitate understanding and usage. The guidelines are intended to be used in conjunction with the SPIRIT-PRO Extension, SPIRIT 2013 Statement and E&E paper and other relevant extensions.^{1 2 10 26} Empirical data and references to support each SPIRIT-PRO item are provided. Real-world examples for each SPIRIT-PRO item, quoted verbatim, are presented to reflect how key elements could be

appropriately described in a trial protocol. These examples were obtained from E&E paper authors, public websites, journals, trial investigators and industry sponsors. Some examples illustrate a specific component of a checklist item, while others encompass all key recommendations for an item. Reference numbers cited in the original quoted text are denoted by (*Reference*) to distinguish them from references cited in this E&E paper. Health-related quality of life (HRQL) has been used consistently to replace terms for quality of life in examples.

ADMINISTRATIVE INFORMATION

SPIRIT-5a-PRO Elaboration

Specify the individual(s) responsible for the PRO content of the trial protocol.

Example

Trial name: multicentre randomized controlled trial of conventional versus laparoscopic surgery for colorectal cancer within an enhanced recovery programme (EnROL)

PRO endpoints: 1°, 2°

EnROL Trial Management Group

Chief Investigator/Clinical Coordinator

RK, [address, telephone]

Co-Investigator

DK, [address, telephone]

Deputy Clinical Coordinator

TR, [address, telephone]

Trial Management/QA

SP, [address, telephone]; *JB*, [address, telephone], *LD*, [address, telephone]; *AF*, [address, telephone]

Nurse Advisor

SB, [address, telephone]

Statistics

SD, [address, telephone]

Health Economics

PF, [address, telephone]

Quality of Life

JMB, [address, telephone]

Translational Science Advisor

PQ, [address, telephone]

Collaborating Surgeons

HW, [address, telephone]; *MG*, [address, telephone].²⁷

Explanation

For trials assessing PROs, input from a person with expertise in PRO methodology early in the development phase of the protocol will improve its completeness and quality.¹⁰ Providing names and contact details of those contributing to the PRO-specific aspects of the protocol provides recognition, accountability and transparency. It aids identification of competing interests and prevents ghost authorship. It also provides a named point of contact to resolve any PRO-specific queries from other research team members, protocol reviewers and sites

(during trial start-up and conduct). Acknowledgements of PRO protocol input from patient-partners as per guidelines for the reporting of patient and public involvement is also recommended.²⁸ Patient and public involvement in all aspects of trial design, including but not limited to: selection of outcomes and measures, timepoints, mode of assessment and reporting, can help minimise burden and ensure that data collected is patient-centred and relevant to participants and to the future patients who will benefit from the research.

Only 7 of 75 (9%) protocols that included PROs from the UK National Institute for Health Research (NIHR) Health Technology Assessment programme explicitly described who was responsible for the PRO component (online supplemental 1).³

SPRIT-6a-PRO Extension

Describe the PRO-specific research question and rationale for PRO assessment and summarise PRO findings in relevant studies.

Example

Trial name: a phase III, randomised, double-blind, placebo-controlled study evaluating the efficacy and safety of idelalisib (GS-1101) in combination with rituximab for previously treated chronic lymphocytic leukaemia (CLL)

PRO endpoints: 2°

Endpoint selection rationale

Health-related quality of life

'Direct patient reporting of outcomes using standardised methods has become an increasingly important component of therapeutic assessment. Evaluation of PROs is particularly relevant in patients who cannot be cured of disease (Reference).

PRO questionnaires have been previously used in CLL to understand how patients differ from the general population in terms of health concerns (References), to understand differences in perceptions of well-being in younger versus older patients (References), to determine how treatment affects HRQL (References), and to assess the pharmacoeconomic cost of improvements in HRQL (Reference).

Patients with CLL have overtly impaired well-being relative to comparable controls (References). Fatigue is cited as a common complaint, being present in the substantial majority of patients. Impairment of HRQL prior to any treatment is apparent in those with B symptoms or in patients with anaemia, supporting the concept of initiating treatment when patients experience symptomatic disease. Factors associated with lower overall HRQL have included older age, greater fatigue, severity of comorbid health conditions, advanced stage and ongoing treatment for CLL (Reference). Younger patients appear to have worse emotional and social well-being but older patients experience worse physical HRQL (Reference). In comparative evaluation of chemotherapy-containing regimens, differences in HRQL between therapies (eg, fludarabine vs fludarabine-cyclophosphamide

vs chlorambucil) reflected differences in toxicity while greater efficacy was associated with improved HRQL (References).

In this phase III study of GS-1101 and rituximab, it is postulated that incremental GS-1101-mediated tumor control will be correlated with greater positive changes in HRQL and that assessments of the drug's safety profile will be supported by HRQL evaluations.²⁹

Explanation

A summary of available PRO evidence and a clearly defined PRO question is required in the background section of the protocol, or a dedicated PRO section if appropriate. Researchers should demonstrate the need for the research and identify the PRO-specific research question to demonstrate the scientific approach and integrity of the PRO study. This should include a review of existing PRO evidence from relevant trials and observational studies (eg, same/similar target population or intervention). This will avoid duplication of research, establish the burden of disease from the patient perspective, identify likely effects of treatment and inform objectives, hypotheses, selection of measures, endpoint definition and analyses (covered by subsequent SPIRIT-PRO items).

Many protocols include PROs without specifying the PRO-specific research question and without a rationale or any reference to PROs in related studies.^{3 4 9}

Provision of this information can inform and motivate research personnel to take note of PRO assessment methods and adhere to standardisation of PRO assessment (eg, when, where, how and who of PRO assessment, as outlined in the protocol under subsequent SPIRIT-PRO items).^{6 11} Staff who understand the importance of PROs in a trial are able to share this understanding with participants. The combined effect of motivated and co-operative staff and participants may help reduce missing PRO data rates.³⁰ This information is also relevant to research ethics committees/institutional review boards (IRBs) and funders responsible for reviewing the scientific integrity and ethical aspects of the trial.

SPRIT-7-PRO Extension

State-specific PRO objectives or hypotheses (including relevant PRO concepts/domains).

Examples

Trial group name: Trans Tasman Radiation Oncology Group (TROG)

Trial name: a randomised phase III trial of high-dose palliative radiotherapy (HDPRT) versus concurrent chemotherapy+HDPRT (C-HDPRT) in patients with good performance status, locally advanced or metastatic non-small cell lung cancer (NSCLC) with symptoms predominantly due to intrathoracic disease who are not suitable for radical chemo-radiotherapy (TROG 11.03 P-LUNG GP)

PRO endpoints: 1°, 2°

Objectives

'The primary objective is to compare, in this group of patients, HDPRT versus C-HDPRT, with respect to

- ▶ the relief of dyspnoea, cough, haemoptysis and chest pain as assessed by change in total symptom burden from baseline to 6 weeks after the completion of treatment;
- ▶ response for each component symptom separately (dyspnoea, cough, haemoptysis, chest pain).

The secondary objectives are to compare the two regimens in terms of dysphagia during treatment, thoracic symptom response rate, duration of thoracic symptom response, HRQL, toxicity, progression-free survival (PFS) and overall survival.

The exploratory/tertiary objectives are

- ▶ to determine how much improvement in HRQL and symptom palliation would be necessary to make the inconvenience due to the longer duration of radiotherapy of C-HDPRT worthwhile, relative to HDPRT. This objective will be addressed in the patient preferences substudy;
- ▶ to analyse serum protein glycosylation changes and exosomes to identify potential biomarkers of disease response and progression. Prospectively collect and bank tumour tissue and blood samples from this cohort of patients for future evaluation of potential biological markers.³¹

Trial name: evaluating different rate control therapies in permanent atrial fibrillation: a prospective, randomised, open-label, blinded endpoint trial comparing digoxin and beta-blockers as initial RAte control Therapy Evaluation in permanent Atrial Fibrillation (RATE-AF)

PRO endpoints: 1°, 2°, exploratory

Hypothesis

'Null hypothesis for primary outcome: no difference in patient-reported quality of life (measured using the physical functioning domain of the 36-Item Short Form Survey (SF-36) questionnaire) when comparing a strategy of digoxin versus beta-blocker therapy for initial rate control in patients with permanent AF.

Alternative hypothesis: use of digoxin or beta-blocker therapy as initial rate control in patients with permanent AF is superior based on patient-reported quality of life (measured using the physical functioning domain of the SF-36 questionnaire).

Primary objective

- ▶ Patient-reported quality of life (HRQL), with a predefined focus on physical well-being using the SF-36 physical component summary at 6 months.

Secondary objectives

- ▶ Generic and AF-specific patient-reported HRQL using the SF-36 global and domain-specific scores, the Atrial Fibrillation Effect on QualiTY of Life (AFEQT) overall score and the 5-level EQ-5D version (EQ-5D-5L)

summary index and visual analogue scale at 6 and 12 months.

- ▶ Echocardiographic left ventricular ejection fraction and diastolic function (E/e' and composite of diastolic indices) at 12 months.
- ▶ Functional assessment, including 6-min walking distance achieved, change in European Heart Rhythm Association class and cognitive function at 6 and 12 months.
- ▶ Change in B-type natriuretic peptide levels as a surrogate for total cardiac strain at 6 months.
- ▶ Change in heart rate from baseline and group comparison using 24-hour ambulatory ECG.³²

Explanation

The PRO objectives should reflect the research question to be addressed in the trial (SPIRIT-6a-PRO Extension) and be described in the context of the population, intervention, comparator, outcome and timepoint and the estimand framework.³³ Study objectives may focus on measuring treatment benefit (superiority), non-inferiority, equivalence. Alternatively, or in addition to one of these objectives, the trial may focus on assessing the safety and tolerability from the patient perspective, or may be more exploratory in nature, where results are presented but no comparative conclusions can be drawn. The PRO-specific study objectives need to clearly align with the proposed analyses methods (SPIRIT-20a-PRO Elaboration). Critically, as described in work by the Setting International Standards in Analysing Patient-Reported Outcomes and Quality of Life Endpoints (SISAQOL) Consortium,³⁴ four key attributes need to be considered a priori for each PRO domain:

1. Broad PRO research objective/research question.
2. Between-group PRO objective.
3. Within-treatment group PRO assumption for the treatment or control arm.
4. Within-patient/within-treatment PRO objective (please note this component of the objective directly addresses the SPIRIT-12-PRO Extension).

More detailed information on how these can be applied are described in the SISAQOL consensus recommendations.³⁴ Although the SISAQOL recommendations were published for oncology trials, the principles apply more broadly. Prespecification of objectives and hypotheses encourages identification of key PRO domains and timepoints. This is particularly important because PRO data are multidimensional in two important ways. First, there is often more than one relevant PRO in a trial, particularly when the high-level outcome of interest is HRQL. Many HRQL questionnaires yield separate scores for distinct dimensions, such as physical, emotional and social functioning, as well as key symptoms such as fatigue and pain. Second, PRO assessments are typically scheduled at several timepoints during a trial, such as baseline, end of treatment, then a series of long-term follow-ups. Prespecification of objectives and hypotheses—focussing on the most important PRO domains and timepoints—is a good

way to reduce multiple statistical testing and avoid selective reporting of PROs based on statistically significant results. Exploratory, hypothesis generating, analyses can also be undertaken but should be specified as such in the final trial report.¹⁹ This links to the SPIRIT-20a-PRO Elaboration, which includes plans for addressing multiplicity/type 1 (α) error. The objectives are generally phrased using neutral wording (eg, ‘to compare the effect of treatment A vs treatment B on fatigue’) rather than in terms of a particular direction of effect.^{2 35} In contrast, the PRO hypothesis states the predicted effect of the interventions on the trial outcomes (eg, ‘patients allocated to treatment A will have less fatigue than those allocated to treatment B’).^{36–38}

Despite the importance of clearly defined PRO objectives and hypotheses, a review of trial protocols determined that 23% failed to include PRO-specific objectives and 81% were missing a clear PRO hypothesis.³

METHODS: PARTICIPANTS, INTERVENTIONS AND OUTCOMES

SPIRIT-10-PRO Extension

Specify any PRO-specific eligibility criteria (eg, language/reading requirements or prerandomisation completion of PRO).³⁰ If PROs will not be collected from the entire study sample, provide a rationale and describe the method for obtaining the PRO subsample.

Example

Trial group name: South West Oncology Group (SWOG)

Trial name: health status and quality of life in patients with early stage Hodgkin’s disease: a companion study to SWOG-9133 (SWOG S9208)

PRO endpoints: 1°, 2°

‘Eligibility criteria

Ages eligible for study: 18 years and older (adult, older adult)

Sexes eligible for study: All

Accepts healthy volunteers: No

Sampling method: Non-probability sample

Study population: community sample.

Criteria: disease characteristics: patients must be eligible for and registered to SWOG-9133.

Patient characteristics: patients must be able to complete the questionnaires in English. If they are not able to complete questionnaires in English, patients may be registered to SWOG-9133 without participating in SWOG-9208.’

The Symptom and Personal Information Questionnaire #1, the Cancer Rehabilitation Evaluation System Short Form and Cover Sheet must be completed prior to registration and randomisation on SWOG-9133.³⁹

Explanation

Any eligibility criteria relevant to PRO assessment should be considered during the trial design and clearly specified in the protocol for consistent use by research

personnel. In some trials, the baseline PRO assessment is required before randomisation as an eligibility criterion.³⁰ This helps to ensure there will be a valid baseline questionnaire from all patients, which is essential for calculation of change scores, or inclusion as a covariate in modelling longitudinal PRO data. For unblinded trials, this also ensures PRO data are collected before participants are aware of the randomisation which may affect some aspects of the participant’s response, for example, anxiety/emotional well-being.⁴⁰ In the absence of such an eligibility criterion, there is a risk that the baseline assessment may be conducted after randomisation but before the intervention is administered, resulting in detection bias. The maximum time between this assessment and randomisation should be defined and should not be too long.

It may not always be possible to collect PROs from all study participants, for example, due to non-availability of questionnaires in appropriate languages (see SPIRIT-18a(iii)-PRO Extension),⁹ literacy requirements or due to cognitive function (see SPIRIT-18a(iv)-PRO Extension). These PRO-relevant exclusions typically should not preclude the affected participants from enrolling in the trial, unless the PRO is the primary outcome. Evidence suggests eligibility criteria as stated in trial protocols often differ from what is finally reported in the trial publication,⁴¹ and data on use of other language or culturally appropriate PRO instruments is often missing from the protocol.^{13 42}

Where the needs of specific groups have been identified (eg, not fluent in English) but not accommodated in the study protocol (eg, non-English language versions not available, assistance with reading and writing English version not permitted), this should be stated, and the rationale for the sampling method described and justified. Trialists should aim to be as inclusive as possible. Given the significance of PROs, the research community has a moral obligation to, where possible, to address gaps in availability of culturally validated PRO instruments. In the meantime, the implications for generalisability of findings should be discussed in subsequent publications.¹⁹

SPIRIT-12-PRO Extension

Specify the PRO concepts/domains used to evaluate the intervention (eg, overall HRQL, specific domain, specific symptom) and, for each one, the analysis metric (eg, change from baseline, final value, time to event) and the principal timepoint or period of interest.

Examples

Trial name: evaluating different rate control therapies in permanent atrial fibrillation: a prospective, randomised, open-label, blinded endpoint trial comparing digoxin and beta-blockers as (RATE-AF)

PRO endpoints: 1°, 2°, exploratory

Primary outcome

‘Patient-reported quality of life (HRQL)—SF-36 physical component summary score at 6 months.

Secondary outcomes

Patient-reported HRQL

- ▶ SF-36 global and domain-specific scores at 6 and 12 months.
- ▶ EQ-5D-5L summary index and visual analogue scale at 6 and 12 months.
- ▶ AFEQT overall score at 6 and 12 months.³²

Trial name: a phase III, randomised, double-blind, placebo-controlled study evaluating the efficacy and safety of idelalisib (GS-1101) in combination with rituximab for previously treated CLL

PRO endpoints: 2°

‘Change in HRQL domain and symptom scores based on the Functional Assessment of Cancer Therapy: Leukaemia (FACT-Leu)—defined as the change from baseline and the time to definitive increments or decrements of 10%, 20% and 40% from baseline; time to definitive increment (better than baseline by the specified amount) is the interval from randomisation to the first timepoint when the HRQL measure is consistently better than at baseline (including that timepoint as well as all the subsequent timepoints) in a subject whose last HRQL score is better than at baseline; and time to definitive HRQL decrement (worse than baseline by the specified amount) is the interval from randomisation to the earliest of death or the first timepoint when the HRQL measure is consistently worse than at baseline (including that timepoint as well as all the subsequent timepoints) in a subject whose last performance status score is worse than at baseline.’²⁹

Explanation

For each outcome, including PROs, the trial protocol should define four components: the specific measurement variable, which corresponds to the data collected directly from trial participants (eg, Beck Depression Inventory score, all-cause mortality); the participant-level analysis metric, which corresponds to the format of the outcome data that will be used from each trial participant for analysis (eg, change from baseline, final value, time to event); the method of aggregation, which refers to the summary measure format for each study group (eg, mean, proportion with score >2) and the specific measurement timepoint of interest for analysis.¹ Many PRO questionnaires are multidimensional, assessing multiple facets of the impact of a disease and its treatment and usually include multiple assessments over the course of the trial. The multidimensional nature of PROs is most apparent in HRQL questionnaires, which often include various aspects of functioning and symptoms, which are often scored as distinct ‘domains’. These domains may not be affected equally by the trial interventions. The SPIRIT-7-PRO Extension encourages protocol writers to identify

the domains that are most likely to be affected in the trial objectives and hypotheses, drawing on previous evidence (SPIRIT-6a-PRO Extension). The SPIRIT-12-PRO Extension item reinforces the statement of these key domains, and also the most important timepoints (ie, where greatest impact of interventions are expected), and develops that concept further by encouraging protocol contributors to think about how these PRO domains and timepoints will be analysed, that is, the analysis metric.³⁴ To ensure transparency and credibility of the analysis, it is recommended that there is prespecification of the PRO concepts/domains, analysis metric(s) and timepoint(s) of interest, whether the PRO is a primary, secondary or exploratory outcome. These should closely align with the study hypotheses/objectives and the nature and trajectory of the disease or condition under investigation.^{34 43} The selected key domains, timepoints and analysis metric should be used to specify the PRO endpoints, integrated in the full endpoint model of the trial.

A clearly defined endpoint model, organising all trial outcomes (PRO and non-PRO), typically in primary, secondary and exploratory endpoints, allows rigorous control of the evidence demonstration, especially the control of the statistical testing. Each PRO endpoint in the model should explicitly specify a single domain and a single time horizon. The endpoint model enables procedures for type I error control (risk of false positive finding) (see SPIRIT-20a-PRO Elaboration). Broadly, the concepts and domains (sub concepts) measured by a PRO may be ‘proximal’ in nature, that is, direct impact of the disease and treatment (eg, symptoms such as pain, fatigue, nausea, rash and anxiety) or more distal, ‘knock-on’ effects, (eg, functional status and global quality of life), as illustrated for ovarian cancer in (figure 1, inspired by the Wilson and Cleary model⁴⁴). Of note, the Food and Drug Administration (FDA) are increasingly focused on the individual measurement of well-defined concepts that impact on HRQL but are more proximal to a therapy’s effect on the patient and the patient’s disease: symptomatic adverse events, physical function and, where appropriate, a measure of the key symptoms of the disease.⁴⁵

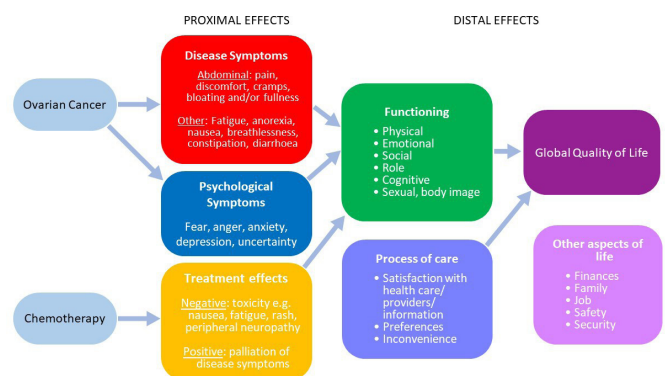


Figure 1 Proximal and distal effects of therapy on patient symptoms and quality of life. Adapted from Wilson and Cleary.⁴⁴

Common analysis metrics may include magnitude of event at time t , proportion of responders at time t , overall PRO score over time or response patterns/profiles. These should be prespecified alongside the levels of statistical and clinical significance for the study and any responder definition in use.¹⁶ Timepoints for analysis should be chosen to best address the research question, while taking into account aspects such as the natural history of the disease/condition and its treatment, the PRO measurement properties and recall period and participant completion burden.^{16 46}

The example, idelalisib and rituximab improve PFS over rituximab alone in unfit patients with relapsed CLL: a phase III study, illustrates a ‘time to event’ PRO endpoint or analysis metric, where the event is definitive improvement or definitive deterioration in a PRO. This approach allows repeated PRO measurements to be converted to a single measure: time to definitive increment or decrement. This requires quite complex and specific criteria for degree and duration of change. Also, this particular example does not specify any key domains of the FACT-Leu, but rather applies this analysis metric to all HRQL domain and symptom scores. In contrast, the RATE-AF example identifies a single score (the SF-36 physical component score), and a specific timepoint (6 months) as the primary outcome, with other SF-36 domains, questionnaires and timepoints specified as secondary outcomes.

SPiRiT-13-PRO Extension

Includes a schedule of PRO assessments, providing a rationale for the timepoints and justifying if the initial assessment is not prerandomisation. Specify time-windows, whether PRO collection is prior to clinical assessments, and, if using multiple questionnaires, whether order of administration will be standardised.

Examples

Trial group: TROG

Trial name: a randomised phase III trial of HDPRt versus concurrent C-HDPRt in patients with good performance status, locally advanced or metastatic NSCLC with symptoms predominantly due to intrathoracic disease who are not suitable for radical chemo-radiotherapy (TROG 11.03 P-LUNG GP) (figure 2).³¹

Trial group: UK Medical Research Council Scottish Cancer Trials Breast Group in association with: Breast International Group

Trial name: MRC phase III randomised trial to assess the role of adjuvant chest wall irradiation in ‘intermediate risk’ operable breast cancer following mastectomy (MRC SUPREMO TRIAL (BIG 2–04) (figures 3 and 4))

Explanation

A clear and concise schedule of PRO assessments (figures 2–4) can: assist trial staff to be organised and prepared for participant visits, inform study participants about the methods and expectations of trial participation

PRO endpoints: 1^o, 2^o

| Assessment Time Point | Beginning of Acceptable time limit | End of Acceptable time limit |
|--|---|---|
| 1. Pre-Registration | Date of signing consent | Date patient made aware of treatment arm to which they have been assigned |
| 2. Pre-Treatment | Date patient made aware of treatment arm to which they have been assigned | Date of first radiation dose |
| 3. During Treatment | Day of treatment | Before next treatment (eg if patient given QoL form on Monday then QoL should be completed before treatment on Tuesday) |
| 4. End of Treatment | Date of last treatment | Five days after last treatment |
| 5. Two weeks after end of treatment | 1 week after end of treatment | 3 weeks after end of treatment |
| 6. Six weeks after end of treatment | Five weeks after end of treatment | Eight weeks after end of treatment |
| 7. Three months after end of treatment | Nine weeks after end of treatment | Eighteen weeks after end of treatment |
| 8. Six months after end of treatment | Five months after end of treatment | Nine months after end of treatment |

Figure 2 Example schedule of patient-reported outcome (PRO) assessments in the Trans Tasman Radiation Oncology Group 11.03 P-Lung GP Trial. QoL, quality of life.

and facilitate review of participant burden by research ethics committees/IRBs.³⁰ The scheduled PRO assessments should provide the data required to address the study’s PRO objectives. When selecting appropriate timepoints for assessment, it is important to consider the natural history of disease/progression, the hypothesised impact of therapy over time and practical considerations such as alignment of assessments with clinic visits and recall period of PRO measures. PRO assessments should be described in the protocol text and in the schedule of assessment table along with the other clinical data collection activities, for ease of reference. This is recommended whether the PRO is completed by the participant during study visits or outside of the study visits (eg, at home).

The timing of the baseline PRO assessment relative to other study-related events is important and therefore should be specified in the schedule of assessments. Collecting PRO data prior to randomisation helps ensure an unbiased baseline assessment, and if specified as an eligibility criterion, can promote data completeness (SPiRiT-10-PRO Extension). Baseline PRO data are often used as a covariate in analyses and are essential to calculating change from baseline, however, collecting data from enrolled patients prior to randomisation can be logistically challenging. One approach is to have participants complete the baseline PRO assessment immediately after providing consent, while the site staff obtain the randomisation assignment from the study system. However, there may be scenarios in which prerandomisation PRO assessment is unnecessary or not possible, for example, emergency surgery trials.

Stating the time-windows for each PRO assessment clearly in the protocol text and schedule of assessments table or footnote will help staff adhere to them. Examples of time-windows for PRO assessment are similar to time-windows for other types of assessments, such as a study visit that may occur on day 10–14 postbaseline, or on day 30±3 days postsurgery. Time-windows for each scheduled

| Investigations | Patients involved | Screening | Post (+/-) chemo pre 1 RT | Post (+/-) RT | 1 yr† | 2 yr | 3 yr | 4 yr | 5 yr | 6 yr | 7 yr | 8 yr | 9 yr | 10 yr | Recurrence ² |
|---|-----------------------------------|-----------|---------------------------|---------------|---|------|------|------|----------------|------|------|------|------|-------|-------------------------|
| | | | | | Baseline 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| Informed consent | All | X | | | | | | | | | | | | | |
| Medical history & examination (b) | All | X | | X | X | X | X | X | X | X | X | X | X | X | X |
| Staging tests | All | X | | | | | | | | | | | | | |
| Contralateral mammography | All | X | | | A mammogram of the opposite breast, if appropriate, is recommended at least in alternate years for 10 years from the date of mastectomy | | | | | | | | | | |
| Blood sampling | If consented to TRANS-SUPREMO | X | | | | | | | | | | | | | X |
| Tumour paraffin block from primary tumour ³ | All | X | | | | | | | | | | | | | |
| Tumour paraffin block at recurrence if available ² | All | | | | | | | | | | | | | | X |
| Acute/ Late morbidity ³ | All | | | X | X | X | X | X | X | X | X | X | X | X | |
| Cardiac symptoms and examination | If consented to cardiac sub study | X | X ⁴ | X | X | | | | X | | | | | X | X |
| Blood sampling for BNP | If consented to cardiac sub study | X | X ⁴ | X | X | | | | X | | | | | X | X |
| Electrocardiogram | If consented to cardiac sub study | X | | | X ⁵ | | | | X ⁵ | | | | | X | X ⁵ |
| Echocardiogram (c) | If consented to cardiac sub study | X | | | X ⁵ | | | | X ⁵ | | | | | X | X ⁵ |
| QoL and EQ5D economic assessment | If consented to QoL sub study | X | | | X | X | | | X | | | | | X | |

(a) Patients in the control arm MUST follow the same follow up schedule as irradiated patients.
 (b) The only exception are patients in the cardiac sub study who receive chemotherapy. This is the only group of patients who must attend a post chemotherapy visit.
 (c) For patients receiving chemotherapy, follow up will be on completion of radiotherapy or at 3 months after chemotherapy in non-irradiated patients.
 (d) For patients not receiving chemotherapy follow up will be on completion of radiotherapy or at 3 months after surgery in non-irradiated patients.
 (e) Questioning for symptoms of recurrent breast cancer, examination of loco-regional area and other relevant clinical areas for evidence of recurrence depending on clinical features.
 (f) In centres where isotope ventriculography is the standard examination for patients requiring anthracycline containing chemotherapy, an echocardiogram will also be required at baseline. Echocardiography will be used for all subsequent time points in the study.
 (g) Baseline (pre randomisation) quality of life assessment will be conducted in the clinic. All subsequent quality of life assessment questionnaires will be mailed to the patient.

Final protocol version 29–30th August 20

Figure 3 Example schedule of patient-reported outcome assessments in the MRC SUPREMO TRIAL (BIG 2–04).¹⁴⁰ BNP, B-type natriuretic peptide; QoL, quality of life.

PRO assessment require an unambiguous reference point, to ensure that PRO data collection captures clinically relevant timepoints of interest. In deciding the size of the time-window for a PRO assessment, consider the trade-off between a smaller, more precise, time-window and a larger more feasible window. One approach is to specify a time-window that is a little larger than the ideal and not allow exceptions; this approach is more consistent than setting a smaller time-window and allowing exceptions. Often PRO assessments that occur during active treatment, for example, chemotherapy, have a smaller time-window to capture acute toxicity that arise and resolve relatively quickly, while those occurring many months or years after treatment completion can have a larger window if the participant's outcomes are expected to stabilise over time.

When the PRO assessment occurs during a research or clinic visit, it is recommended that PRO assessment is standardised to be completed prior to clinical consultation, assessments or procedures. For example, if a PRO instrument assesses participants' experiences of pain in the past 7 days, and the study visit includes a bone marrow biopsy, the schedule of assessments should indicate that the PRO assessment be completed prior to the biopsy. This will prevent the pain assessment from capturing pain associated with the biopsy, reduce risk of missing data as participants may not feel well enough to complete PROs following their procedure and offer a 'routine' for study staff responsible for data collection.

When more than one PRO questionnaire is scheduled, it is recommended that the order of questionnaires is

standardised, with those higher in the endpoint hierarchy being collected first.

These two forms of standardisation of PRO administration are examples of the more general principle in research methodology that standardisation of methods reduces unwanted sources of variation, whether random (ie, no net effect on estimates of interest, such as the impact of interventions on PROs) or systematic (ie, causing bias).

SPiRiT-14-PRO Elaboration

When a PRO is the primary endpoint, state the required sample size (and how it was determined) and recruitment target (accounting for expected loss to follow-up). If sample size is not established based on the PRO endpoint, then discuss the power of the principal PRO analyses.

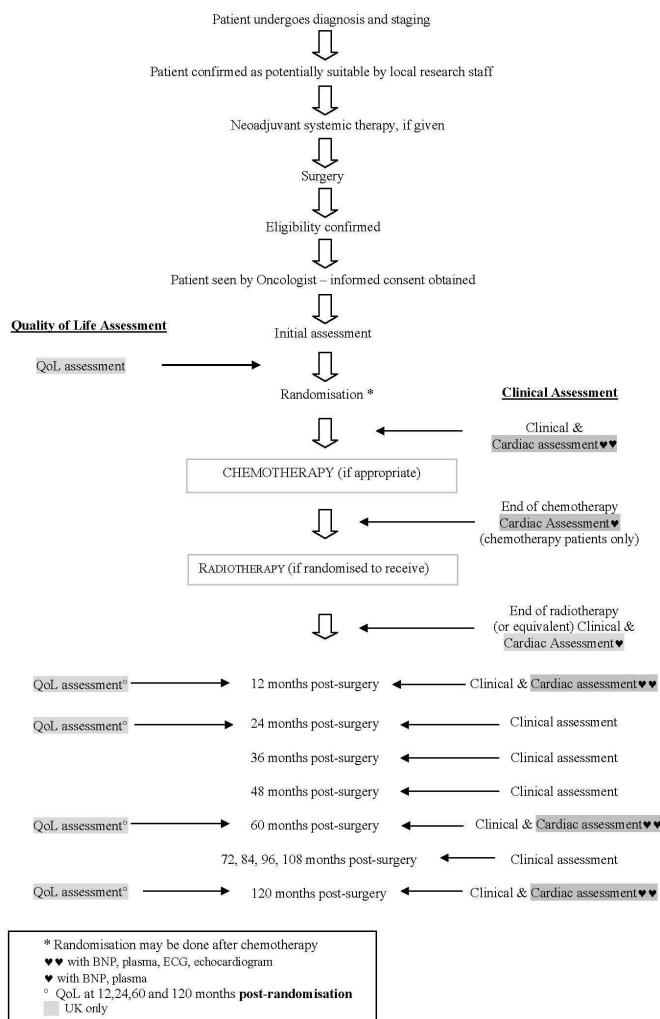
Examples

Trial name: the chronic autoimmune thyroiditis quality of life selenium trial (CATALYST)

PRO endpoints: 1°

'The primary outcome is thyroid-related quality of life during 12 months' intervention, as measured by a composite score from the ThyPRO questionnaire. Sample size estimation is based on this outcome.

The trial should be sufficiently powered to identify a difference between the intervention and the control group of 4 points on the 0–100 ThyPRO composite scale, corresponding to a small to moderate effect. In previously obtained data, the SD of ThyPRO-scores (sigma level) was 20 points. With a correlation between observations on the



Final protocol version 29–30th August 2010

Figure 4 Flow diagram schedule of patient-reported outcome assessments in the MRC SUPREMO TRIAL (BIG 2–04).¹⁴⁰ BNP, B-type natriuretic peptide; QoL, quality of life.

same participant of 0.50, and a power of 80% and a type I error probability (two-sided α level) of 0.05, a sample size of 236 experimental participants and 236 control participants is required. The sample size estimate is based on a design with five repeated measurements having a compound symmetry covariance structure'.⁴⁷

Trial name: cosmesis and body image after single-port laparoscopic or conventional laparoscopic cholecystectomy: a multicentre double-blinded randomised controlled trial (SPOCC-trial)

PRO endpoints: 1°, 2°

'The primary endpoint of the study concerns patient's satisfaction with cosmesis and body image 12 weeks after surgery. This endpoint is assessed using a validated cosmesis and body image score (CBIS) that was previously used in surgery for Crohn's disease and in donor nephrectomy. This score is calculated on an 8-item multiple choice type questionnaire ranging between 8 and 48 points.

A clinically relevant improvement of the CBIS is defined as an improvement of 20% of the cosmesis score

(8 points). Given the reported SD of the CBIS between 4 and 6 and using ($\alpha=0.05$ and $\beta=0.90$), two groups of 49 patients are needed. This is based on a two-sided significance level (α) of 0.05 and a power of 0.90. Estimating a 10% dropout rate, which is common in randomised controlled trials, 55 patients will be randomised per arm.⁴⁸

Trial group: TROG

Trial name: a randomised phase III study of radiation doses and fractionation schedules for ductal carcinoma in situ of the breast (BIG 3-07/TROG 07.01)

PRO endpoints: 2°

The sample size for this trial, based on the primary endpoint (time to local recurrence of invasive or intraductal breast cancer in the ipsilateral breast), was 1600 patients. The sample size for the PRO substudy was determined a priori, and was less than that required for the primary endpoint, as explained in the protocol excerpt below. Therefore, patients recruited after the PRO-specific target sample size was achieved did not complete PRO questionnaires, saving trial resources in data collection and management.

'Sample size determination: for the quality of life study aiming to detect a difference between the tumour bed boost and no boost groups of 0.2 SD of a continuous scale such as fatigue or physical symptoms, with 80% power at a two-sided α level of 5%, the required sample size is 790 patients. To allow for attrition at a rate of 5% per year, 1020 patients are required to participate in the quality of life study.'⁴⁹

Explanation

As with any primary endpoint, including those that focus on PRO, the criteria and methods for estimating the necessary sample size should be specified, with adjustments for expected discontinuation from the clinical study.⁴⁶ Ideally, the criteria for clinical significance (eg, minimal important difference, clinically meaningful within-patient change threshold, responder definition) should be specified when known.^{50 51} It is important to note that the FDA is more interested in what constitutes a meaningful within-patient change in score from the patient perspective.¹⁶

In cases where the PRO is specified as a key secondary endpoint, the statistical power based on the estimated sample size for the primary endpoint, should be determined. If overpowered, specifying a smaller PRO-specific sample size will save trial resources, as illustrated in the BIG 3-07/TROG 07.01 example. When sufficient power may be achieved by collecting PROs from a representative subset of participants, the sampling strategy should be clearly described.

Only 50.7% of NIHR Health Technology Assessment clinical trial protocols address sample size and statistical power for PRO specified as secondary endpoints.³ If the clinical trial is international in scope, the sampling across countries may be influenced by availability of language

translations.⁵² In addition, the variability of measurement between countries may inflate type 2 error (reduces power).

METHODS: DATA COLLECTION, MANAGEMENT AND ANALYSIS

SPiRiT-18a(i)-PRO Extension

Justify the PRO instrument to be used and describe domains, number of items, recall period, instrument scaling and scoring (eg, range and direction of scores indicating a good or poor outcome). Evidence of PRO instrument measurement properties, interpretation guidelines and patient acceptability and burden should be provided or cited if available, ideally in the population of interest. State whether the measure will be used in accordance with any user manual and specify and justify deviations if planned.

Examples

Trial name: impact of a multimodal support intervention after a 'mild' stroke (YOU CALL-WE CALL)

PRO endpoints: 1°, 2°

'A number of quality of life tools were reviewed (eg, SF-36, Stroke Impact Scale, Quality of Life Index (QLI)) and the tool chosen was a compromise between psychometric properties and adequacy of content for mild stroke. The 32-item questionnaire QLI (*Reference*) which was developed from Ferran's conceptual model of quality of life and which has been used with a stroke clientele (*Reference*) was chosen as the primary outcome. Each item of the QLI as relating to four life domains (health and functioning, socioeconomic, psychological/spiritual and family), is evaluated in terms of satisfaction and importance on a 6-point scale. Scores for each domain and a global score are expressed from 0 to 30, with a higher score indicating a better quality of life. These four life domains relate well with the main issues covered through the WE CALL intervention. It has shown to have adequate psychometric properties (concurrent validity, test-retest reliability and high internal consistency: $\alpha=0.90$) (*Reference*) and thus should be responsive to therapy-induced change (*Reference*). A 1-point difference was observed in the first 6 months poststroke descriptive follow-up (n=63) for an effect size of 0.33 (*Reference*). A 2-point difference is considered a clinically meaningful change leading to a moderate effect size of 0.66.⁵³

Trial name: a randomised phase II/III multicentre clinical trial of definitive chemoradiation, with or without cetuximab, in carcinoma of the oesophagus (SCOPE 1: Study of Chemoradiotherapy in Oesophageal Cancer Plus or Minus Erbitux)

PRO endpoints: 2°

HRQL instruments

'Generic domains of HRQL will be assessed with the EORTC core Quality of Life Questionnaire, the EORTC QLQ-C30 (*Reference*). This instrument has been well

validated in many international clinical trials in oncology including oesophageal adenocarcinoma and squamous cell cancer. Disease-specific and Chemoradiotherapy (CRT)-associated symptoms and side effects will be assessed with the oesophageal cancer-specific module, the EORTC QLQ-OES18 (*Reference*). This has been validated and tested in patients receiving definitive CRT. The module includes scales assessing dysphagia, eating restrictions, reflux, dry mouth and problems with saliva and deglutition. The Dermatology Life Quality Index (DLQI) will also be administered (*Reference*). This is a well-validated, easy-to-use index which assesses the impact of dermatological conditions on patients' HRQL (*Reference*). It has been included to accurately assess the impact of the acneiform eruption commonly seen with cetuximab.⁵⁴

Explanation

The justification for the selection of PRO instrument(s) is required in the trial protocol. This will help trial personnel and participants understand why specific measures are being used and how they directly address the trial objectives and stakeholder needs.¹¹ For example, regulatory agencies often focus on physical symptoms and functioning to inform licensing and labelling claims, whereas patients and health-policy makers may be more interested in broader aspects of HRQL, such as engaging in social activities and emotional well-being.^{55 56} For regulatory trials, it is prudent to seek regulatory advice at an early stage of trial development regarding the acceptability of the instrument and the approach to PRO assessment. Stakeholder-relevant PROs can be identified through patient involvement, qualitative research or core outcome sets,^{56 57} which alongside clinical outcomes, often include outcomes such as symptom burden, functioning, and disease control, which can be measured using PRO instruments.

Appropriately developed and evaluated PRO instruments can provide more sensitive and specific measurements of the effects of medical intervention, thereby increasing the efficiency of clinical trials that attempt to measure the meaningful treatment benefits of those therapies.⁵⁸⁻⁶⁰ Irrespective of whether the trial is conducted for regulatory purposes, FDA guidance and ISOQOL guidance provide a useful conceptual framework to assist in the selection of measures.^{16 61} Identifying and selecting valid, reliable tools that are acceptable to patients from the target population may prove challenging. The Consensus Based Standards for the Selection of Health Measurement Instruments (COMET) initiative and the Evaluating the Measurement of Patient Reported Outcomes programme provide useful guidance to support the review of measurement properties.^{56 62 63}

Ideally, the PRO instrument(s) will have been validated in the target population and this evidence cited. This will help reviewers understand if claims being supported by the PRO instrument can be substantiated by the evidence for using that instrument in that, or a related, population. Further details on the domains, number of items, recall



period, instrument scaling and scoring (eg, range and direction of scores indicating a good or poor outcome) should be provided. This will assist trial personnel in the collection and analysis of the PRO data. Questionnaires should be used in accordance with user manuals to promote good data quality and ensure standardised scoring. Deviations from user manuals or different ways of capturing PRO data may invalidate the measure; therefore, any deviations should be declared and transparently reported.¹⁹

If in the trial there are plans to use a questionnaire which has not been validated in the trial's target population, or if a new instrument is being developed alongside the trial, it is important to explain this in the protocol. Including an outline of any plans for the evaluation of its measurement properties using the trial data, if this will be undertaken, and if not why. This should be in accordance with established current guidelines for PRO validation.⁶⁴

Although all the reviewed NIHR Health Technology Assessment clinical trial protocols identified the PRO instrument to be used in the trial, few justified their use in relation to the study hypotheses, PRO instrument measurement properties or expected participant burden (41.3%, 37.3% and 14.7%, respectively).³

Patient partners involved in the design of the study can assist with the selection of PRO instruments and provide feedback on the likely acceptability of the questions, and participant burden (eg, time taken for completion, cognitive burden, emotional burden, repetition across questionnaires).⁶⁵ The number of PRO instruments/questions to be assessed in a trial requires careful justification. Minimising participant burden has been identified as a strategy to reduce risk of missing PRO data, improve recruitment and retention.³⁰

SPIRIT-18a(ii)-PRO Extension

Includes a data collection plan outlining the permitted mode(s) of administration (eg, paper, telephone, electronic, other) and setting (eg, clinic, home, other).

Examples

Trial name: early surgery versus optimal current step-up practice for chronic pancreatitis: a multicentre randomised controlled trial

PRO endpoints: 1°, 2°

'The Izbicki pain score will be assessed every 2 weeks during a follow-up period of 18 months. For this end, the Izbicki pain score will be assessed via a web questionnaire. Patients who do not have an email will be given a folder with Izbicki pain score forms and return envelopes. Patients will be contacted by telephone every 2 weeks and reminded to fill in the questionnaire and send it to the trial coordinators. The Izbicki pain score is a one page questionnaire, easily completed in <3 min. The folder with the Izbicki score forms will be re-filled at every outpatient clinic visit (scheduled every 6 months).'⁶⁶

Trial name: a randomised multistage phase II/III study of sunitinib comparing temporary cessation with allowing continuation, at the time of maximal radiological response, in the first-line treatment of locally advanced/metastatic renal cancer (STAR)

PRO endpoints: 1°, 2°

Quality of life questionnaires during the first 6 months will be administered in clinic in order to support participant use before postal questionnaires are instituted after 6 months for the EQ-5DTM/EQ-VAS (Functional Assessment of Cancer Therapy - General (FACT-G) and Functional Assessment of Cancer Therapy-Kidney Symptom Index (FSKI) will continued to be collected at clinic visits). Clinic staff should remind participants of the importance of the quality of life assessments at each clinic visit.

Due to the importance of HRQL data in this trial, measures will be taken to ensure maximum compliance of questionnaire completion. For the 2 weekly questionnaires which participants complete at home from the 24-week timepoint, where the participant consents to this, reminders for completion are sent by email or text message to the participant by the research team at the Clinical Trials Research Unit (CTRU): this is an optional part of the STAR Informed Consent Form. Where a HRQL questionnaire would be completed at a hospital clinic visit, but the local research team forget to give this to the participant, of the participant no longer attends clinic visits at hospital during their follow-up period, a questionnaire for the local research team will send this out by post to the participant's home after checking the participant's status and establishing it is appropriate to do so.^{67 68}

Explanation

Standardisation of all aspects of PRO administration is vital to PRO data quality. It is therefore critical that research personnel and trial participants understand how, when and where PRO data will be collected in the study.¹⁰ The study protocol should specify the permitted mode(s), method(s) and setting(s) of PRO data collection, including the permitted 'back-up' options and preplanned reminders. For example, when PRO assessment is conducted in clinic via a tablet computer, paper forms could be permitted (and available) as a back-up option for instances when the tablet is not available or functioning properly. Offering alternative modes of completion may help improve response rates.³⁰ Of note, the FDA has previously recommended that there is a back-up plan for electronic PRO data collection (eg, web-based, phone-based or paper-based) implemented in case of malfunctions with electronic devices.⁶⁹

Electronic PRO assessment is increasingly available in trials, but traditional paper-based methods may still be useful or required in some situations. It is therefore important to know whether there are systematic differences induced by mode of administration. A recent meta-analysis that included 31 studies that randomised

participants to different data collection modes found no evidence of bias associated with paper versus electronic administration.⁷⁰ These results support the use of multiple modes of administration within a research study, which may be a useful strategy for reducing missing PRO data. If evidence of equivalence between different modes of administration is available for the specific PRO questionnaires in a trial, it should be considered in determining the PRO administration plan. If electronic administration has not been attempted before for the trial PRO questionnaires, and only minor modifications to layout/presentation are needed with respect to the paper-based versions, it is advisable to pilot-test usability and conduct cognitive debriefing to assess equivalence.^{70 71} The International Society for Pharmacoeconomics and Outcomes Research provide useful guidance on key considerations for PRO data collection including multiple modes.^{71 72}

The setting for PRO data collection, for example, in clinic or at home (or clinic at baseline, with follow-up at home), should be described and standardised across trial intervention groups and sites. Differential use of settings and modes of administration by treatment arm should be avoided as these may lead to different response rates and potentially biased results.¹¹

The protocol should also specify the types of assistance trial staff can provide patients for completing the PRO assessment. Respondents should be encouraged to self-complete as far as possible. Some respondents may require some assistance, however, the greater the degree of assistance, the greater the potential to influence a respondent's responses. Assistance should therefore be limited, provided only by a trained member of the research team, or a trained third party, the permissible types of assistance should be clearly specified in the protocol and reviewed in staff training. Allowable assistance might include instructions on how patients can input their answer on the tablet, clarifying the response options, reading questions to the participant or recording the participants' answer on the form/tablet. This level of assistance facilitates self-administration of the PRO instrument. Completion of a PRO instrument with an interpreter, caregiver or family member should be avoided as these individuals have not been trained, and may influence the individual's responses, either directly by expressing opinions that influence the participant to alter their answers, or indirectly, for example, if the respondent seeks to avoid embarrassment or to provide a more acceptable answer (social desirability bias).

Furthermore, use of a human language interpreter should be avoided. When planning a study, common languages spoken by patients attending the recruiting centres should be considered so that validated language translations of chosen PRO instruments can be obtained (see SPIRIT-18a(iii)-PRO Extension).

Interviewer administration of PRO instruments should be avoided, but where necessary, should be clearly justified in the protocol. Interviewers should read questions verbatim, ideally using a PRO instrument that has been

validated in that mode. Similarly, proxy or observer completion requires a proxy-validated/observer-reported version of the PRO instrument (see SPIRIT-18a(iv)-PRO Extension).

SPIRIT-18a(iii)-PRO Extension

Specify whether more than one language version will be used and state whether translated versions have been developed using currently recommended methods.

Example

Trial name: A Phase III, randomised, open label trial of lenalidomide/dexamethasone with or without elotuzumab in relapsed or refractory multiple myeloma (ELOQUENT – 2)

PRO endpoints: 2°, exploratory

Outcomes research assessments

HRQL assessments

'To assess the impact of treatment, subject's quality of life will be measured using three validated HRQL instruments: the European Organization for Research and Treatment of Cancer Quality of life Questionnaire-Core (EORTC QLQ-C30), the myeloma-specific module (QLQ-MY20) and the Brief Pain Inventory-Short Form (BPI-SF).

Non-English-speaking subjects will complete the questionnaire using validated language transitions developed and recommended for each instrument. The BPI-SF has demonstrated both reliability and validity across cultures and languages, and has been used to study the effectiveness of pain treatment. (*Reference*) A score of 6 on a scale of 0–10 on any single item is generally considered to be clinically significant. (*Reference*) Pretesting was carried out in the UK, Norway, Sweden, Denmark and Germany. Field testing of the module has been conducted in a range of phase III trials. (*Reference*) The module has been validated in a large number of languages (see www.eortc.be/home/qol).⁷³

Explanation

Trials involving participants with different language requirements require measures that have been translated and culturally adapted using appropriate methodology.^{10 12 74 75} Providing culture-appropriate and language-appropriate PRO instruments for use in the trial can lead to a reduction in missing data, ability to recruit people from ethnic minority groups, lower attrition rates and improved generalisability of trial results.⁷⁶ If the countries/languages are not known at the time of protocol writing then more general protocol content may be appropriate:

multiple language validated versions are available [provide references where these can be found] and the correct language for this patient should be used.

At present, the extent to which this is happening is not clear. A review of protocols and/or subsequent publications from cancer clinical trials with a PRO endpoint, registered on the NIHR portfolio examined reporting of

ethnically diverse recruitment and the use of culturally and linguistically validated PRO instruments. The review found a lack of transparency around the use of culturally and linguistically appropriate PRO instruments. Of the 88 studies reviewed, only 14 (17%) reported any type of data on ethnic diversity. Although eight studies were multicentre, multinational cancer clinical trials, none identified if translated versions of PRO instruments were being used.⁷⁷

There are clear guidelines for translating PRO instruments,^{74 78} and plans to use translated versions, should be specified in the protocol, citing references when available.¹⁰ Specification of use of translated versions in the protocol will help reporting in accordance with CONSORT-PRO.^{19 74 75} It must not be assumed that linguistic translation equates to cross-cultural adaptation (preparing the instrument for use in another setting). A number of studies^{79 80} have recommended that cross-cultural equivalence is also an important consideration.^{74 75 81}

Using different language versions of PRO instruments to collect data in a trial ideally requires evidence to support the psychometric equivalence of data being reported, especially if data are going to be pooled for clinical trial evaluation.^{52 82 83} Where such evidence is unavailable, prespecification in the statistical analysis plan (SAP) of exploratory analyses to assess whether there are differences between PROs by language group may be appropriate. The language in which each patient complete the questionnaire should be recorded in the database to inform such analyses.⁵²

SPIRIT-18a(iv)-PRO Extension

When the trial context requires someone other than a trial participant to answer on his or her behalf (a proxy-reported outcome), state and justify the use of a proxy respondent. Provide or cite evidence of the validity of proxy assessment if available.

Examples

Trial name: cognitive rehabilitation in paediatric acquired brain injury—a randomised controlled trial (CORE-pABI)

PRO endpoints: 1°, 2°

‘[Paediatric Acquired Brain Injury] constitutes a major disruption to child development and may affect cognitive, behavioural, emotional, social as well as academic function.

The primary outcome measure is the BRIEF, parent report (*Reference*). BRIEF is an 86-item standardised questionnaire that captures parents perceptions of a child’s executive function in his or her everyday environment. Each item’s frequency of occurrence is rated on a 3-point Likert scale from 1 (never) to 3 (often). It has demonstrated good reliability, with high test–retest reliability ($r=0.88$ for teachers,⁸² for parents), internal consistency (Cronbach’s $\alpha=0.80–0.98$) and moderate correlations

have been detected between teacher and parent ratings ($r=0.32–0.34$). The questionnaire has been applied to several clinical groups in Norway.⁸⁴

Trial name: evaluating the effectiveness and cost effectiveness of dementia care mapping (DCM) to enable person-centred care for people with dementia and their carers: a cluster randomised controlled trial in care homes (DCM EPIC study 1.0)

PRO endpoints: 2°

‘Relative/friend criteria

To be eligible to provide proxy data about a resident, relatives/friends must: have visited the resident on a regular basis over the past month (ie, at least once per week). Be willing to provide data at a time convenient to them. Have sufficient proficiency in English to contribute to the data collection required for the research.⁸⁵

Explanation

In some contexts, such as trials involving young children or cognitively impaired participants or participants who are unable to reliably self-report for other reasons, it may be necessary for a proxy—someone other than a trial participant, to report the participant’s outcomes on their behalf as though they are the patient.^{10 86}

Proxy reports should be used only when necessary. The European Medicines Agency states that ‘in general proxy reporting should be avoided, unless the use of such “proxy raters” may be the only effective means of obtaining information that might otherwise be lost.’^{16 43} The US FDA also discourages the use of proxy-reported outcomes to inform labelling claims, recommending observer reports for observable phenomenon only (eg, vomiting, but not nausea) instead.

In contexts such as cancer, dementia or palliative care, it is reasonable to anticipate the need for proxy response throughout all or some of the trial. Previous studies have shown varying levels of agreement between participant and proxy ratings, dependent on the variable being measured, the quality, duration and stability of the relationship between proxy and participant.^{87 88}

A trial protocol should indicate clearly who is eligible to provide the proxy report, with explicit administration guidelines for completion of proxy measures including how the report is to be captured, whether that same individual must be the ‘consistent rater’ across all time-points of assessment (this is preferable, for consistency), or whether varying proxy reports will be permissible. This information should also be provided for observer-reported outcomes.

Just as the measurement properties of the PRO instrument should be specified, so should the properties of measures to be used by proxy reporters.

Given known issues with patient and proxy reporter discordance,^{88–91} while patient-participants are still able to self-complete, collecting both participant and proxy-reported data enables quantification of the size and direction of any bias, that may later be adjusted for, if needed.

Further data may be gathered about the proxy (eg, age, relationship to the patient, gender, proxy literacy, relationship and exposure to the patient⁹²) as these variables may guide interpretation of results and any subgroup/sensitivity analyses. Whether proxy-reported data will be analysed separately or pooled with participant-reported data should also be detailed. Any such plans should be specified in the protocol and SAP. This information should also be provided for observer-reported outcomes.

SPRIT-18b(i)-PRO Extension

Specify PRO data collection and management strategies for minimising avoidable missing data.

Examples

Trial group: National Cancer Institute, Naples

Trial name: phase III randomised multicentre trial of carboplatin+liposomal doxorubicin versus carboplatin+paclitaxel in patients with ovarian cancer (Multicentre Italian Trials in Ovarian Cancer-2 (MITO-2))

PRO endpoints: 2°

‘Operating procedure:

- ▶ It is fundamental that the researchers take great care when collecting the questionnaires, in order to allow good compliance by the patients participating in the protocol.
- ▶ The quality of life form must be filled in by the patient herself.
- ▶ The quality of life form must be filled in before the clinical examination, and thus before the discussion with the examining doctor which may provide favourable or unfavourable information about the disease’s status.
- ▶ When supplying the form to the patient, it is important to explain how to fill it in without going into details about the contents of the questions.
- ▶ After the form has been returned, check that the patient has answered all the questions and ask her to reply to any questions she has skipped.
- ▶ The quality of life questionnaires must be filled in using a black or blue pen.⁹³

Trial group: National Cancer Institute of Canada Clinical Trials Group (NCIC CTG)

Trial name: a double-blind randomisation to letrozole or placebo for women previously diagnosed with primary breast cancer completing 5 years of adjuvant aromatase inhibitor either as initial therapy or after tamoxifen (including those in the MA.17 Study) (NCIC CTG: MA.17R)

PRO endpoints: 2°

Quality of life

‘Mandatory for NCIC CTG centres and optional for centers within other cooperative groups:

Patient is able (ie, sufficiently fluent) and willing to complete the two quality of life questionnaires in either English or French. The baseline assessment must have been completed prior to randomisation. Inability

(illiteracy in English or French, loss of sight or other equivalent reason) to complete questionnaires will not make the patient ineligible for the study. However, ability but unwillingness to complete the questionnaires will make the patient ineligible.⁹⁴

Explanation

Missing PRO data are a particular problem because data cannot be obtained retrospectively or from medical records. Missing PRO data may arise from different sources⁹⁵ and, broadly speaking, missing data can be attributed to causes that are unavoidable or avoidable. Unavoidable reasons may include if a participant has died or become too unwell to self-complete PRO instruments. Avoidable reasons may include some type of human error that could have been prevented. Examples of avoidable missing data include: staff failing to hand out a scheduled questionnaire; a participant not realising the questionnaire is double-sided so missing half of the questions; an electronic PRO device not being charged; the internet or server being down; when the PRO assessment is overly burdensome on the patient (eg, due to the burden of multiple questionnaires at one time or repetitive scales) so the patient decides not to complete it.

Although ‘unavoidable’ types of missing data are more challenging for interpretation because the missing data may be related to the measured outcome and it is impossible to accurately calculate the extent of any associated bias,^{96,97} avoidable types of missing data are also problematic.³⁰ Avoidable missing PRO data compromise the interpretability, accuracy and value of PRO findings because study power is reduced, which increases the risk of type 2 errors,⁹⁶ and because any assumptions made during the analysis about missing PRO values are not verifiable.⁹⁸

There are a range of design, implementation and reporting strategies to help minimise and address missing PRO data,³⁰ most of which can be addressed in the trial protocol. Specific recommendations related to data collection and management include: refraining from administering an excessive number of questionnaires to participants (researchers should refrain from collecting more data than they really need), using standardised and documented PRO administration procedures, engaging and educating participants in the trial by providing updates or incentives, maintaining participant records, employing active quality assurance measures (such as real-time compliance/completion monitoring, sending reminders for upcoming or missed assessments, checking completed questionnaires for missing items while the participant is still present at clinic), appointing a dedicated staff member responsible for PRO assessment at each centre, training staff about the importance of PROs as well as procedures for assessment, offering an alternative mode of administration if the participant is not able to complete the questionnaire via the primary mode (eg, completing the questionnaire over the phone if the hardcopy cannot be completed at the clinic within the acceptable time-window) and recording reasons for

missed assessments using standardised forms.³⁰ From a regulatory perspective, the FDA encourages maintaining consistency in assessment methods, however, this should be balanced with reduction of missing data. If different modes are used, they should be justified and presented in the study documentation.⁹⁹ If different modes are used, 'FDA will review the comparability of data obtained when using multiple data collection methods or administration modes within a single clinical trial to determine whether the treatment effect varies by method or mode.'⁶⁹ Additional strategies are described in the full review.³⁰

The MITO-2 and MA.17R examples above illustrate some of these strategies. MA.17R comes from the Canadian Cancer Trials Group (CCTG), a multicentre cooperative oncology group that conducts clinical trials in cancer therapy, supportive care and prevention. The CCTG requires completion of the PRO questionnaire(s) as a prerandomisation eligibility requirement (as per SPIRIT-10-PRO Extension). This flags the importance of PRO data to investigators and clinical research associates, indicating PROs are as important as other inclusion clinical criteria. It also helps to maximise compliance.

Our prior work suggests that few trials actively specify such procedures in their protocols; 46.7% HTA protocols³ and 38.5% of international ovarian cancer trials⁴ included strategies to minimise avoidable missing data.

SPIRIT-18b(ii)-PRO Elaboration

Describe the process of PRO assessment for participants who discontinue or deviate from the assigned intervention protocol.

Examples

Trial name: a multicentre, open-label, randomised, two-arm phase III trial on the effect on PFS of bevacizumab plus chemotherapy versus chemotherapy alone in patients with platinum-resistant, epithelial ovarian, fallopian tube or primary peritoneal cancer (AURELIA)

PRO endpoints: 2°

AURELIA was a multicentre, open-label, randomised, two-arm phase III trial of bevacizumab plus chemotherapy versus chemotherapy alone in patients with platinum-resistant, epithelial ovarian, fallopian tube or primary peritoneal cancer (*Reference*). The primary endpoint was PFS; quality of life was a secondary endpoint. The protocol stated that:

'On clear evidence of disease progression (PD) or toxicity, study therapy should be discontinued permanently.'

The protocol also specified postprogression treatment options: women who had been on the chemotherapy alone arm would have the option of bevacizumab alone or standard of care, while those who had been on the chemotherapy plus bevacizumab arm would receive standard of care treatment. The protocol stated that:

'In case the patient decides to prematurely discontinue study treatment ("refuses treatment"), she should be

asked if she can still be contacted for further information' and that 'after PD, patients will be followed for survival only.'

However, it lacked an explicit statement about PRO assessment postprogression, which may explain some inconsistency among sites, with some sites collecting PRO data postprogression, and other not.

As stated in the AURELIA PRO paper (*reference, p1310*), 'only questionnaires completed until PD were included in the main analyses. Questionnaires completed after PD were excluded based on the medical assumption that these patients were unlikely to be benefiting from their study treatment, may have been receiving another treatment and were therefore not relevant to the intended comparison of chemotherapy alone versus bevacizumab plus chemotherapy. However, post hoc sensitivity analyses were performed to determine the impact of questionnaires completed after PD.' The latter analyses were consistent with the main analyses, but could perhaps have been avoided, and patients saved unnecessary HRQL assessment burden, if the protocol has contained a clear statement that HRQL assessments should cease at disease progression.¹⁰⁰

Trial group: Australasian Leukaemia & Lymphoma Group (ALLG)

Trial name: BLAM-A phase IIb study of blinatumomab+cytarabine (AraC) and methotrexate in adult B-precursor acute lymphoblastic leukaemia (ALLG ALL8)

PRO endpoints: 2°

ALLG ALL8 is a BLAM-A phase IIb study of blinatumomab+cytarabine (AraC) and methotrexate in adult B-precursor acute lymphoblastic leukaemia (B-ALL). It is a single-arm study which aims to demonstrate preliminary evidence of the benefit of frontline blinatumomab in combination with cytarabine and methotrexate in adult B-ALL, and to demonstrate the ability of this combination to attain deep (MRDnegative) remissions and hence to reduce the need for allogeneic stem cell transplantation. The ALLG ALL8 protocol specifies that the decision for allogeneic stem cell transplantation is left up to the investigator, and that patients who proceed to allogeneic stem cell transplant discontinue the assigned intervention protocol. Being recommended for allogeneic stem cell transplant is therefore a withdrawal criteria. However, patients do not have to withdraw—it is optional, the patient's choice. The trial team is in fact interested in the experience of patients who proceed to allogeneic stem cell transplant. So the protocol states:

Subjects who proceed to allogeneic stem cell transplant who have not withdrawn consent to the study should have FACT-Leu (HRQL questionnaire) assessments performed 6-monthly.

This provides the trial team with the opportunity to document the experience of patients who proceed to allogeneic stem cell transplant in terms of the range of

common symptoms and aspects of well-being assessed by FACT-Leu.¹⁰¹

Explanation

A clear plan for collection of PROs for trial participants who withdraw early from a study or who discontinue the intervention helps minimise bias,¹⁰² ensures that staff collect all required PRO data in a standardised and timely way, and may assist ethical appraisal of the study.^{10 102}

Often, participants can provide valuable PRO data even after stopping the assigned intervention protocol, whether due to personal choice and/or clinical recommendation, as illustrated in the one-arm phase II trial ALLG ALL8 trial. However, this does not hold in all contexts (as illustrated in the AURELIA trial), a randomised phase III trial, in which participants whose cancer had progressed on their assigned treatments, were then often switched to alternative treatment.

Providing a clear description of the process of PRO assessment for participants who discontinue or deviate from the assigned intervention protocol and how the data will be used, enables all staff to follow a standardised procedure to collect the required PRO data in timely way and to avoid collecting data that will not contribute to analysis.

Correspondingly, the SAP should be clear on how such data will be handled. In the case where postdiscontinuation/deviation PRO data are useful, the SAP should state the study objective that these data will address and how they will be analysed.

Participants should also be aware of this process, so a simple and clear description of whether or not they will be asked to continue to complete PRO questionnaires after stopping or changing the treatment they were initially allocated to should be included in the participant information sheet (PIS).

SPIRIT-20a-PRO Elaboration

State PRO analysis methods, including any plans for addressing multiplicity/type I (α) error.

Example

Trial name: a phase III, randomised double-blind, placebo-controlled study of the efficacy and safety of two doses of tofacitinib (Cp-690,550) in subjects with active psoriatic arthritis and an inadequate response to at least one tumour necrosis factor inhibitor (OPAL BEYOND)

PRO endpoints: 1°, 2°, exploratory

Analysis of primary endpoint

'There are two primary endpoints in A3921125 and two doses of tofacitinib each of which will be compared with placebo for each endpoint. In order to control for type I error, a stepwise testing procedure will be used. This implies that a given endpoint for a given dose can only achieve significance if the prior endpoint is significant. The order of the fixed sequence for testing against placebo is as follows: tofacitinib 10 mg ACR20 response rate at 3 months, tofacitinib 5 mg ACR20 response

rate at 3 months, tofacitinib 10 mg Health Assessment Questionnaire-Disability Index (HAQ-DI) at 3 months, tofacitinib 5 mg HAQ-DI at 3 months. This gate-keeping or step-down approach strongly protects the type I error rate at the 0.05 (two-sided) level.

ACR20 response

For all comparative analyses, the normal approximation for the difference in binomial proportions will be used to test the superiority of each dose of tofacitinib to placebo and to generate CIs for the differences. The primary analysis will be ACR20 response rate at month 3. The ACR20 response rate will also be analysed at other timepoints as a secondary analysis.

Missing values due to a subject dropping from the study for any reason (eg, lack of efficacy or adverse event) will be handled by setting the ACR20 value to non-responsive.

HAQ-DI

The HAQ-DI score will be expressed as a change from baseline. The primary timepoint will be at month 3. The analysis will be done using a repeated measure model that includes fixed effects of treatment, visit (week 2, months 1, 2 and 3), treatment by visit interaction, geographic location and baseline value. The model will use and fit an unstructured variance-covariance matrix. Full details will be listed in the analysis plan.

Additional analyses of the HAQ-DI will include a responder analysis at month 3 where subjects with a change of 0.3 will be considered responders and subjects who dropped from the study will be considered non-responsive. Another responder analysis will be conducted using a change of 0.35 at the cutpoint for response (*Reference*). The normal approximation for the difference in binomial proportions will be used for these responder analyses.

Analysis of secondary and other endpoints

Key secondary efficacy variables are as follows: PASI75, enthesitis score, dactylitis severity score, physical function domain of SF-36 and FACIT-F at month 3. In order to strongly protect the study-wise type I error rate with respect to these key secondary endpoints and the primary endpoints, these endpoints will be tested only if all endpoints/doses for the primary endpoints are statistically significant. The order of testing is as listed above; for each endpoint, tofacitinib 10 mg will be tested versus placebo first, followed by tofacitinib 5 mg. Testing stops at the first instance in which statistical significance is not achieved.

Methods for analysing all other endpoints will be enumerated in the statistical analysis plan. Briefly, binary variables (eg, remission rates) will follow the analyses described above for binary variables (eg, ACR20) and continuous endpoints will follow the same type of analyses described above for continuous endpoints (eg, HAQ-DI). Descriptive statistics may also be calculated and displayed'.¹⁰³



Explanation

Statistical analysis of multiple domains^{10 104} and time-points implies multiple hypothesis testing, which inflates the probability of false-positive results (type I error).⁴⁶ This can be contained by prespecifying the key PRO domain(s) or overall score of interest and the principal timepoint(s) which cross-reference to SPIRIT-7-PRO Extension and SPIRIT-12-PRO Extension.

Any plans to address multiplicity, such as stepwise or sequential analyses, whereby multiple endpoints are tested in a defined sequence that contains the overall type I error to the desired level, or conventional non-hierarchical methods (eg, Bonferroni correction), should be specified a priori.¹⁶ There are many strategies and/or choices of methods that may be appropriate.¹⁰⁵ Family-wise type I error should be considered for all of the applicable endpoints of the trials together and not for the PRO endpoints separately. Some clinical trials include PROs as exploratory endpoints and no adjustment is made for multiplicity in subscale scores administered at multiple study visits. These analyses may only provide limited information on the tolerability of the intervention.

Protocols should make some reference to key considerations for the analysis of the trial PROs (including any plans for addressing multiple testing), but the detail is often more appropriately included in the SAP, usually developed after the protocol. If no adjustments for type I error are to be made, then this should be clearly stated. However, clinical trial protocols in which PROs are secondary outcomes rarely include any information about PRO statistical analyses, beyond any that are prespecified as primary or key secondary endpoints or when the sponsor is interested in achieving a product label claim. Our review of trial protocols found that fewer than 2% provided information on the statistical analysis plans to address multiplicity for the PRO endpoints.³

SPIRIT-20c-PRO Elaboration

State how missing data will be described and outline the methods for handling missing items or entire assessments (eg, approach to imputation and sensitivity analyses).

Example

Trial group: European Organisation for Research and Treatment of Cancer-Gynaecological Cancer Group, and Gynecologic Cancer Intergroup (EORTC-GCG and GCIG)

Trial name: a randomised, multicentre, phase III study of erlotinib versus observation in patients with no evidence of disease progression after first line, platinum-based chemotherapy for high-risk ovarian epithelial, primary peritoneal, or fallopian tube cancer (EORTC 55041)

PRO endpoints: 2°

Compliance

'Missing data may hamper assessment of HRQL in clinical trials. This may occur because centres do not collect

the questionnaires at the appropriate time (unit non-response), and also because patients may not reply to questions within the questionnaires (item non-response). The latter problem occurs <2% on average and should not be a problem. The former problem will be minimised by ensuring that participating centres are properly informed and motivated towards HRQL assessment.

During the study, compliance with completing HRQL questionnaires will be investigated at each timepoint. The compliance of the HRQL assessments will also be reviewed twice a year and will be a part of the descriptive report by the Data Center for the Group's plenary sessions and, if possible, be presented by the EORTC Quality of Life Group's appointed liaison person.

The compliance rate between the two arms will be compared at each timepoint using a χ^2 test. In order to adjust for the multiplicity of the tests, a Bonferroni adjustment will be made by which each test will be performed at the 0.01 significance level. Should follow-up compliance levels drop below 60% at subsequent biannual compliance reviews, then the Protocol Writing Committee would review this to either improve compliance or consider terminating the HRQL assessment in the trial.

Missing data

When performing HRQL analyses complications may arise due to large quantities of missing data. This issue has a bearing on whether a valid comparison of the treatment arms is being made.

In HRQL research there are two main types of missing data: (1) item non-response, (2) unit non-response (the whole questionnaire is missing for a patient). As item non-response occurs <2% on average in the QLQ-C30, it is not such a major problem and thus the methods described in the EORTC QLQ-C30 scoring manual for handling item non-response will be used. For missing questionnaires, it is necessary to identify both the extent of missing questionnaires and the main process of missing data. Three different types of missing data processes may exist: missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR, informative dropout mechanism). These have distinct consequences for data analysis (*Reference*).

If the missing data process is considered to be non-ignorable (MNAR), then quality of life will be compared between groups using longitudinal data modelling techniques (ie, Proc mixed in SAS with either selection models or pattern-mixture models) in combination with a logistic regression for the dropout process. If the missing data mechanism can be considered ignorable (MAR), then standard longitudinal data analysis will be used (proc mixed in SAS). If the data are MCAR, then complete case analysis can be used without biasing the results.¹⁰⁶

Explanation

Most clinical trials with PROs will have some missing PRO data,⁹⁸ yet a review of protocols found that less than half outlined statistical methods to deal with missing PRO data.³

There are two types of missing PRO data: (1) missing items, when the participant completes some but not all of the questions within a PRO instrument and (2) missing assessments, when the participant does not complete a scheduled assessment at all (ie, there are no PRO data available for analysis from the participant at that assessment timepoint). The latter type is more serious, as it potentially affects the choice of analysis method and the interpretation and generalisability of the results. The trial protocol should explain how both types of missing data will be handled in the analysis.

1. Handling missing items: many PRO scoring manuals provide guidance for handling missing items. A common approach is to impute the mean score of the completed items, if less than one-half of items comprising the scale are missing.^{107 108} This approach is possible for multi-item scales but it is not possible to impute scores for single item scales. Always consult the scoring manual to determine how to handle missing items, and cite the manual in the protocol. Missing items of PRO instruments that are underpinned by modern psychometrics techniques (Item Response Theory or Rasch Measurement Theory) are naturally handled without requiring imputation.^{34 109}
2. Handling missed assessments: missing assessments present a major problem for PRO analyses; lead to a loss of power and wider CIs from a lack of precision,^{110 111} and potentially to biased results. The risk of bias depends on the underlying cause of why data are missing (called the ‘missing data mechanism’), and this in turn should influence the choice of statistical analysis methods.

In order to gain insights into the mechanism of missing PRO data, it is helpful to ascertain and record reasons for missed PRO assessments during trial conduct. The protocol needs to describe how these reasons will be collected in a standardised manner. Typically, a standard form is used (which can be included in the online supplemental appendix along with the PRO questionnaires). Also, the form can be used to collect additional data (referred to as ‘auxiliary data’) related to ‘missingness’ or the PRO,³⁰ which should be specified in the protocol.

The trial protocol should also provide a summary of how missing data will be described and handled in the analysis,¹⁰ and state that comprehensive details about the planned analysis will be provided in a subsequent SAP.¹¹² The SISAQOL Consortium have developed a taxonomy of research objectives that can be matched with appropriate statistical methods for PRO analysis, standardised statistical terminology relating to missing data and are determining appropriate ways to manage missing data, currently focused in an oncology setting. A simulation study was done to assess whether it was possible to have a threshold to define substantial missing data.¹¹³ Although no agreement was reached for a threshold, the simulation study showed that the effect of missing data rates on PRO findings depends

on the type of missing data (ie, informative or non-informative missing data). It was recommended that collecting reasons for missing data is key in assessing the effect of missing data for PRO findings.³⁴ Additionally, SISAQOL is developing a set of macros to describe patterns of missing data, and to evaluate imputation methods for use in sensitivity analysis.^{34 36 38}

METHODS: MONITORING SPIRIT-22-PRO Extension

State whether or not PRO data will be monitored during the study to inform the clinical care of individual trial participants and, if so, how this will be managed in a standardised way. Describe how this process will be explained to participants, for example, in the PIS and consent form.

Example

Trial name: the use of an electronic patient-reported outcome measure in the management of patients with advanced chronic kidney disease (RePROM)

PRO endpoints: 2° Protocol

‘If, when reviewing the completed EQ-5D-5L questionnaire, the research nurse becomes concerned for the well-being of the participant, they should discuss their concerns with the participant directly, working in partnership to determine the best course of action. With the participant’s permission, the research nurse may need to consult with the PI and/or treating clinician to address these concerns. In exceptional circumstances, the research nurse may consult with the PI and/or treating clinician without the permission of the participant if they are concerned for the participant’s safety’.

Patient information sheet

‘If the study research nurse becomes concerned for your well-being during a study review, they will discuss their concerns with you, to determine the best course of action. With your permission, the research nurse may need to consult with a senior member of the study and/or your treating clinician to address these concerns. In exceptional circumstances, the research nurse may need to do this without your prior permission if they are concerned for your safety’.

Consent form

‘I understand that if the study research nurse becomes concerned for my well-being during a study review, they will discuss their concerns with me to determine the best course of action. With my permission, the research nurse may need to consult with a senior member of the study and/or my treating clinician to address these concerns. In exceptional circumstances, the research nurse may need to do this without my prior permission if they are concerned for my safety.’¹¹⁴

**Box regulatory/HTA perspective**

Regulatory agencies, such as the Medicines and Healthcare products Regulatory Agency, European Medicines Agency and FDA are placing more focus on capturing the patient experience when developing drugs.^{16 43 55 133} However, poorly defined PRO objectives have hindered the utility of PROs in regulatory decisions. Accurate and well-defined PRO methods can provide the patients' perspective on the impact of a treatment on disease-related symptoms and symptomatic adverse events. Efforts to improve PRO clinical trial standards are welcomed and the SPIRIT-PRO extension provides an additional resource for drug developers to consider in their development programmes. Alongside these guidelines and other regulatory guidance documents, the importance of seeking early scientific advice directly from the regulatory authorities and health technology assessment bodies such as NICE cannot be understated, where advice can be given on the acceptability of a particular approach. Therefore, raising PRO standards is key to the successful integration of PROs in drug development programmes, ensuring that the impact of medicines on a disease can be captured from the patient's perspectives.

Explanation

To protect participant safety, PRO data may be monitored during a study for signs of psychological distress or physical symptoms that may require an immediate response: so-called 'PRO Alerts'.¹¹⁵ Examples of PRO data that may raise concern include signs of psychological distress, poor physical well-being or high symptom burden presenting as extreme scores on questionnaires. Concerns can also arise when additional information is provided by the participant (eg, through free text report), or in discussions between the participant and research staff.¹¹⁵ The nature of some studies may mean that participants are more at risk than in others. In studies where prior risk assessment deems the probability of PRO alerts being generated by participants to be minimal, PRO data may not be reviewed until the end of the trial for pragmatic reasons, however, concerning PRO data, may still arise during the course of the study.^{11 116} If monitoring is not planned this should be stated. If monitoring is planned, steps for how PRO alerts will be dealt with should be included in the trial protocol, to provide immediate reassurance to concerned trial staff about how to proceed and promote appropriate clinical management and transparency considering professional obligations to patient care. Any arising interventions should be recorded. Evidence suggests the absence of such information leads to inconsistent handling of concerning data including administration of non-prespecified interventions to aid the trial participant which risks co-intervention bias that is, bias caused by 'any intervention other than the experimental manoeuvre that alters the frequency of a trial's outcome of interest'.^{12 116 117} Identifying participants at risk and in need of urgent attention through PRO monitoring is an ethical issue as is acknowledging that information identified in the completion of a PRO too may be shared with the clinical team when the need arises. Information about how the trial staff will respond to concerns, or alternative

Box patient perspective

What is the question that doctors and nurses ask their patients more than any other? "How are you feeling?" That is why we need PROs in clinical research. They allow patients to answer that question in a systematic and measurable way, which will benefit others and from which we may well benefit ourselves at some point.

For patients, including PROs in health and social care research studies is vital in assessing whether or not our health and/or well-being are improving. Well-designed PRO instruments will help assess our general health and emotional mood, our ability to complete our daily tasks and our self-measured levels of pain and/or fatigue.

PROs are very different from the clinical measures used to assess the effectiveness of new drugs or other treatments, yet for us, it is the PROs that measure 'tolerability' and thus the real-life 'effectiveness' of the drug or intervention as a medicine or treatment.

The ultimate measure of the performance of any health service must be in whether or not it helps people recover from an acute illness, live well with a chronic condition and face the end of life with dignity—and people's own reports on their own condition are the only valid way to gauge success. So if a drug or treatment is to be trialled for use in healthcare delivery, it is essential that PROs are included in the success criteria.

It is equally essential therefore that participants on the study understand the importance of completing PRO assessments and to understand how and why the data are collected. This should not be too onerous for researchers to explain to patients. Patients who choose to participate in clinical trials do so because they wish to benefit others and if possible, to benefit themselves. PROs are a means of doing both—provided that the reporting and recording are not too much of a burden. "We must do all that we can to make patient-reported outcome assessment feasible and credible. If we fail in our task we will have left out the heart of all healthcare research: the patient".¹⁴¹

And when we do complete our PROs on your research study, please let us know what happens to the study, what you know now that you did not know before, and how that will be used to help people. Because that is why we participate in research; it is to help people like us.

support mechanisms where monitoring is not taking place, should be provided to participants (in participant information and consent documentation). This information provided in the patient information sheet (PIS) will manage participant expectations and ensure transparent and explicit communication about the intended use of PRO study data in fulfilment of contemporary data protection laws, for example, the European Union General Data Protection Regulation.¹¹⁸

There is no regulatory requirement from the US FDA that PRO data measuring symptomatic side effects be monitored and alerts for these items created. However, it is good practice to remind participants at each PRO assessment whether their data are or is not being monitored in real time. In the case PRO data are not being monitored, participants should be reminded to speak to clinical staff if they are experiencing a specific problem, symptom or side effect.

Protocol template

The PRO protocol template (online supplemental file 2) aims to support protocol writing for pharmaceutical

companies, funders, clinicians and international trials groups by providing PRO content that can be incorporated directly into the relevant sections of existing clinical trial protocols or retained as a dedicated PRO section within the protocol.

Supplementary trial documents

Online supplemental 2 outlines additional items recommended for inclusion in other trial documentation, such as the SAP, PIS and training and guidance documents for staff.^{10 119} This is not an exhaustive list and further PRO content may be warranted in training materials and patient facing documents dependant on the trial. We recommend input from PRO experts working in conjunction with the clinical team, trials unit or Contract Research Organisation and patient partners involved in the co-design of research with regulatory input as required to optimise the protocol and supplementary resources.

DISCUSSION

This paper provides a detailed rationale, implementation instructions and real-world examples to assist investigators to develop the PRO-specific components of clinical trial protocols, in accordance with the 16 items of the SPIRIT-PRO Extension. This SPIRIT-PRO Extension E&E paper is recommended for use alongside the original SPIRIT 2013 and SPIRIT-PRO guidelines.^{12 10} The mission of the SPIRIT-PRO Group is to improve the design and standardisation of PRO components of clinical trials and thereby ensure high-quality PRO data to inform patient-centred care. To further facilitate uptake of the SPIRIT-PRO items, we have provided a PRO-specific protocol template covering all the SPIRIT-PRO items. This can be used in two ways: either incorporated item-by-item into relevant sections of existing clinical trial protocols or retained whole as a dedicated PRO section within a trial protocol. The use of a template should support investigators to address all required SPIRIT-PRO checklist items comprehensively and meaningfully, in conjunction with the real-world examples provided for each SPIRIT-PRO item in this manuscript.

The overall aim of the SPIRIT-PRO Extension and E&E is to improve the completeness, quality and transparency of PRO sections of clinical trial protocols, where PROs are a primary or key secondary outcome. We also recommend use of the guidelines to support development of protocols where PROs data are exploratory in nature, including single-arm trials with PRO endpoints. Many of the SPIRIT-PRO items may also provide useful prompts about PRO content for cohort studies and other non-randomised designs. The SPIRIT-PRO guidelines,¹⁰ and E&E paper aim to facilitate development of high-quality PRO protocol content, which will ultimately also facilitate the review of protocols by research ethics committees/IRBs, scientific review groups and funders. Improved PRO protocol content has been associated with more complete reporting which will help facilitate the critical

appraisal of final trial reports and results³ and use of PRO data to inform patient-centred care. Several SPIRIT-PRO items correspond to items on the CONSORT-PRO checklist.^{10 19} This is particularly important since reviews of PRO reporting indicate that, where published PRO trial data were available, there was often considerable delay between publication of primary outcomes and the PRO results and standards of reporting were poor.^{4 5 7 9 120–123}

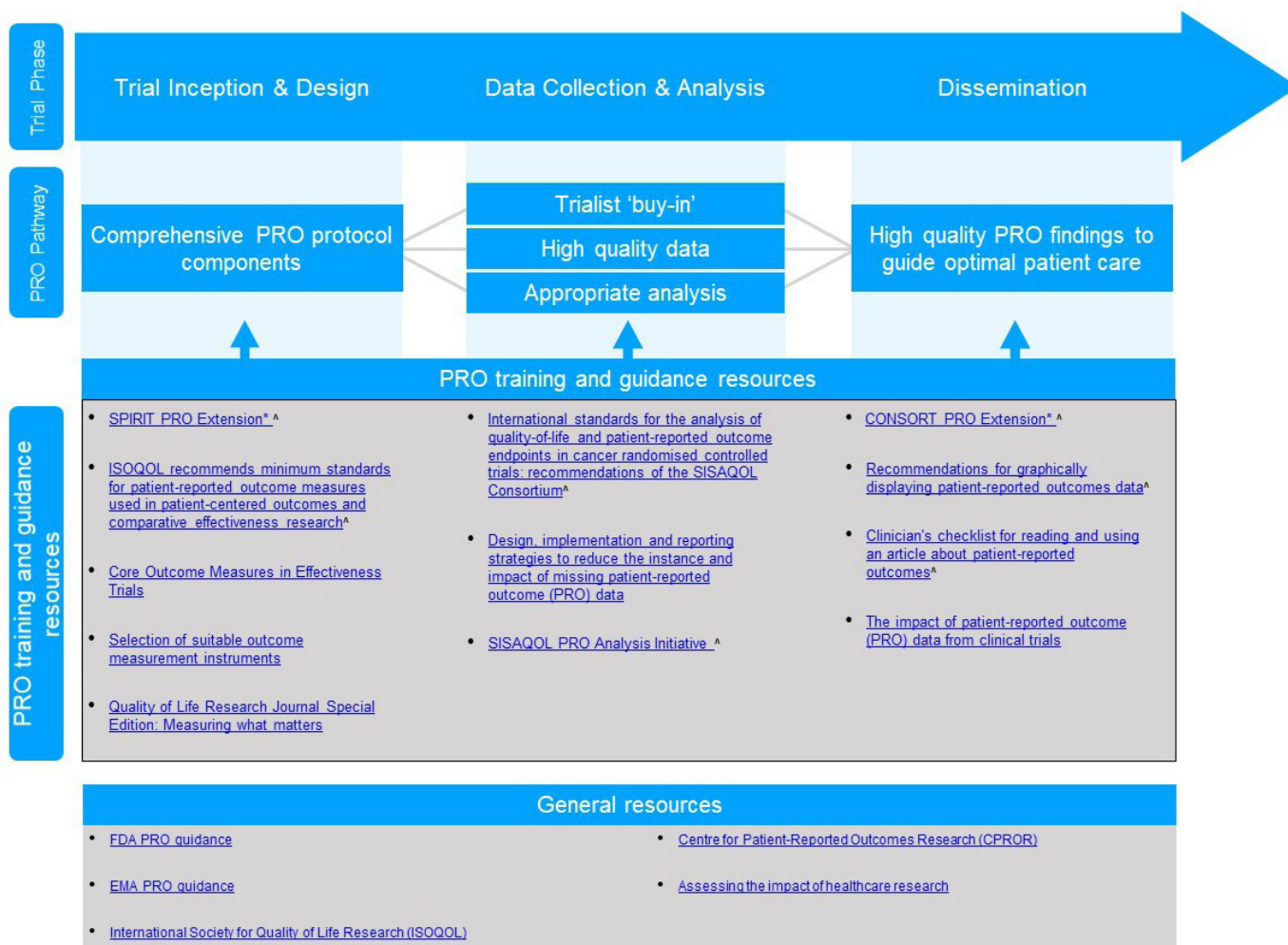
Worryingly, a recent review of cancer trials suggested that 49568 participants were involved in studies that failed to publish their PRO data and that poor reporting was associated with suboptimal PRO protocol content.⁹ This finding is consistent with findings from Schandelmaier *et al*, which demonstrated that 52% of cancer trials specified HRQL outcomes in their protocols, however, only 20% reported any HRQL data in associated publications.¹²⁰ Non-reporting of PRO findings is widespread,^{7 9 120–128} meaning patient-centred information may not be available to benefit patients, clinicians and regulators. Non-reporting of these important patient data is unethical and is a waste of limited healthcare research resources.^{129–131}

In the EU Clinical Trials Regulation (536/2014), there will be a requirement for results of all primary endpoint and patient relevant secondary endpoints to be reported within 12 months of the end of the study.¹³² The provision of a protocol template alongside example excerpts from trial protocols will help facilitate protocol developers understand how to write high-quality PRO protocol content and support more complete reporting of results.

In a companion paper,¹⁵ we also present tools for patient advocates involved in the co-design of trial protocols or the review of protocols through roles on ethics committees or funding committees with PRO endpoints to further optimise study design and facilitate patient involvement. It is essential that patients' experiences, perspectives, needs and priorities are captured and meaningfully incorporated. The SPIRIT-PRO Group and regulatory agencies strongly support the early and continued involvement of patients and members of the public in trial design and conduct.^{133 134}

The next steps for the SPIRIT-PRO Initiative are to promote uptake and use of the guidelines and implementation tools and development of ethical guidelines for IRBs and ethics committees. The SPIRIT website (www.spirit-statement.org) and PROlearn, a free resource on the optimal use of PROs in research and routine practice (www.bham.ac.uk/prolearn) provides the latest resources and information on the initiative, including a list of supporters. We invite international stakeholders to assist in the evaluation of the SPIRIT Statement and E&E paper by using the documents and providing feedback to inform future revisions.

SPIRIT-PRO forms part of a growing toolbox to promote the optimal use of PROs in trials including guidance for the selection of measures,⁶¹ design (SPIRIT-PRO),¹⁰ analysis (SISAQOL),³⁴ reporting (CONSORT-PRO)¹³ and presentation of results (figure 5).^{135–137} These tools are currently being disseminated by the PROTEUS



* It is recommended that these are used in conjunction with the general SPIRIT and CONSORT guidelines, as well as additional SPIRIT and CONSORT extensions as relevant to study design
^ Resources currently being promoted by the PROTEUS Initiative

Figure 5 Resources to promote high-quality patient-reported outcomes (PROs) trial design, analysis and dissemination.⁹ CONSORT, Consolidated Standards of Reporting Trials; EMA, European Medicines Agency; FDA, Food and Drug Administration; SPIRIT, Standard Protocol Items: Recommendations for Interventional Trials.

Consortium (Patient-Reported Outcomes Tools: Engaging Users & Stakeholders) who aim to partner with key patient, clinician, research and regulatory groups around the world to promote the uptake and use of these methodological tools to optimise the assessment and reporting of PROs in clinical trials (<https://more.bham.ac.uk/proteus/>). Patient and public involvement in all of these activities can help ensure that PRO selection, study implementation and application is transparent, relevant and acceptable. Consistent with this philosophy, patient partners have been involved in all aspects of the development of the SPIRIT-PRO Extension.^{10 138 139}

Through widespread uptake and support, the potential to improve the completeness and quality of trial protocols and the efficiency of their review can be fully realised. Ultimately, high-quality PRO results can help ensure that important patient-centred evidence on the efficacy, safety and tolerability of interventions is available to inform shared decision-making, labelling claims, clinical guidelines and health policy.

Author affiliations

¹Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, Birmingham, UK

²Birmingham Health Partners Centre for Regulatory Science and Innovation, University of Birmingham, Birmingham, UK

³NIHR Biomedical Research Centre, Institute of Translational Medicine, Birmingham, UK

⁴NIHR Surgical Reconstruction and Microbiology Research Centre, Birmingham, UK

⁵NIHR Applied Research Collaboration West Midlands, University of Birmingham, Birmingham, UK

⁶Faculty of Science, School of Psychology, University of Sydney, Sydney, New South Wales, Australia

⁷Faculty of Medicine and Health, NHMRC Clinical Trials Centre, University of Sydney, Sydney, New South Wales, Australia

⁸Women's College Research Institute, Department of Medicine, University of Toronto, Toronto, Ontario, Canada

⁹University of North Carolina, Chapel Hill, North Carolina, USA

¹⁰Oncology Digital Health, AstraZeneca, Gaithersburg, Maryland, USA

¹¹Cancer Outcomes Research Program, University of North Carolina, Chapel Hill, North Carolina, USA

¹²US Food and Drug Administration, Silver Spring, Maryland, USA

¹³NIHR Bristol Biomedical Research Centre, University of Bristol, Bristol, UK

- ¹⁴Division of Surgery, Head and Neck, University Hospitals Bristol NHS Foundation Trust, Bristol, UK
- ¹⁵Department of Quality of Life, European Organisation for Research and Treatment of Cancer, Brussels, Belgium
- ¹⁶Clinical Trials Research Unit, University of Leeds, Leeds, UK
- ¹⁷Department of Oncology, Queen's University Cancer Research Institute, Kingston, Ontario, Canada
- ¹⁸Medicines and Healthcare Products Regulatory Agency, London, UK
- ¹⁹Global Biometrics & Data Management—Statistics, Pfizer Inc, New York City, New York, USA
- ²⁰University of Warwick, Coventry, UK
- ²¹Department of Health Sciences Research, Mayo Clinic, Scottsdale, Arizona, USA
- ²²School of Population and Global Health, McGill University, Montreal, Quebec, Canada
- ²³Patient-Centered Outcomes Research Institute, Washington, DC, USA
- ²⁴JAMA, Chicago, Illinois, USA
- ²⁵Boehringer Ingelheim International GmbH, Ingelheim, Germany
- ²⁶Warwick Research in Nursing, University of Warwick, Warwick Medical School, Coventry, UK
- ²⁷Independent Advisor, Manchester, UK
- ²⁸Daiichi Sankyo Europe GmbH, Munchen, Germany
- ²⁹National Cancer Institute, Rockville, Maryland, USA
- ³⁰Patient-Centered Outcomes Research, UCB Pharma, Brussels, Belgium
- ³¹Value Evidence and Outcomes—Patient Centered Outcomes, GSK, Collegeville, Pennsylvania, USA
- ³²Center for Observational Real-world Evidence (CORE), Patient-Centered Endpoints & Strategy, Merck & Co Inc, Kenilworth, New Jersey, USA
- ³³Cancer Research Institute, Queen's University, Kingston, Ontario, Canada
- ³⁴Department of Health Services, University of Washington, Seattle, Washington, USA
- ³⁵Modus Outcomes, Lyon, France
- ³⁶Outcomes Research Consulting, Sarasota, Florida, USA
- ³⁷Johnson and Johnson, Janssen Global Services LLC, High Wycombe, UK
- ³⁸National Cancer Research Institute Consumer Forum, London, UK
- ³⁹EMD Serono Inc, Healthcare Business of Merck KGaA, Darmstadt, Germany
- ⁴⁰Leeds Institute of Medical Research at St James's, University of Leeds, Leeds, UK
- ⁴¹University of California, Irvine, California, USA

Twitter Grace Turner @graceturner

Acknowledgements The authors would like to thank Trish Groves, MRCPsych. Kluetz P, MD, US Food and Drug Administration, Jeanette Kusel, MSc, National Institute for Health and Care Excellence, Laura-Lee Johnson, PhD, US Food and Drug Administration, Joanna Coast, PhD, University of Bristol and Doug Altman, DSc for their contribution to the SPIRIT-PRO. The SPIRIT-PRO Group gratefully acknowledge the additional contributions as detailed in reference [10], Appendix (refer to online supplemental file 2) made by the SPIRIT-PRO Executive, the ISOQOL Best Practices for PROs in Randomised Clinical Trials Protocol Checklist Taskforce, the international stakeholders responsible for stakeholder survey distribution and stakeholders who completed the stakeholder survey, the Delphi panellists and the SPIRIT-PRO International Consensus Meeting Participants.

Contributors MC and MK had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. MC and MK are coauthors of the SPIRIT-PRO Group. Concept and design: MC, MK. Acquisition, analysis or interpretation of data: MC, MK, RM-B, OA, DK, AS, A-WC, EB, JBe, ABe, VB, JBI, ABo, JBr, MB, LC, JCC, HD, ACD, CE, LF, RMG, IG, KH, BK-K, LM, SM, TM, LN, JN, DO'C, MP, DP, GP, AReg, ARet, DR, JS, RS, GT, AV, GV, MvH, AW, LW. Drafting of the manuscript: MC, DK, RM-B, AS, A-WC, MK+section writers GV, AReg, DR, ABe, SM, LW, MP, JBr, GT, ARet, AW. Critical revision of the manuscript for important intellectual content: MC, MK, RM-B, OA, DK, AS, A-WC, EB, JBe, ABe, VB, JBI, ABo, JBr, MB, LC, JCC, HD, ACD, CE, LF, RMG, IG, KH, BK-K, LM, SM, TM, LN, JN, DO'C, MP, DP, GP, AReg, ARet, DR, JS, RS, GT, AV, GV, MvH, AW, LW. Obtained funding: MC, DK, MK, Doug Altman, JBI, JBr, MB, Joanna Coast, HD, MvH, J Ives, RM-B, GP, L Roberts, AS. Supervision: MC.

Funding This SPIRIT-PRO Extension was funded by Macmillan Cancer Support (grant 5592105) and the University of Birmingham and was sponsored by the University of Birmingham. Development of the PRO protocol template was funded by an unrestricted educational research grant from UCB Pharma. MC receives funding from the National Institute for Health Research (NIHR) Birmingham Biomedical Research Centre, the NIHR Surgical Reconstruction and Microbiology

Research Centre and NIHR ARC West Midlands at the at the University of Birmingham and University Hospitals Birmingham NHS Foundation Trust, Health Data Research UK, Innovate UK (part of UK Research and Innovation), Macmillan Cancer Support, UCB Pharma. MK is supported by the Australian government through Cancer Australia. RM-B is supported by the Australian Government by a National Health and Medical Research Council (NHMRC) research fellowship. JBI is supported by the NIHR Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust. She is also a NIHR Senior Investigator. Work was funded by an unrestricted grant from UCB Pharma, and a participant (TM) contributed as a coauthor and member of the industry advisory group.

Disclaimer The study funders/sponsors had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript or decision to submit the manuscript for publication.

Competing interests All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and declare: SPIRIT-PRO group members were reimbursed for travel/subsistence at the consensus meeting. MC is Director of the Birmingham Health Partners Centre for Regulatory Science and Innovation, Director of the Centre for the Centre for Patient Reported Outcomes Research and is a National Institute for Health Research (NIHR) Senior Investigator. MC has received personal fees from Astellas, Takeda, Merck, Daiichi Sankyo, Glaukos, GSK and the Patient-Centred Outcomes Research Institute (PCORI) outside the submitted work. RM-B reports non-financial support from University of Birmingham. OA, DK and ARet report grants from NIHR. OA and DK report grants from Birmingham Biomedical Research Centre (BRC). OA reports grants from UCB Pharma and also receives funding from the Health Foundation and declares personal fees from Gilead Sciences Ltd. DK and ARet report grants from Innovate UK and Macmillan Cancer Support. DK reports grants from Kidney Research UK, NIHR SRMRC at the University of Birmingham and University Hospitals Birmingham NHS Foundation Trust, personal fees from Merck, GSK. EB declares personal fees from Navigating Cancer, Sivan Healthcare, CareVive Systems and AstraZeneca. Bell reports other from AstraZeneca, is an employee with stock ownership and/or stock options in the company. JCC reports other from Pfizer Inc., and is an employee and a stockholder of Pfizer Inc. IG is a fully paid employee of Boehringer Ingelheim International GmbH. CE is Chair of the Government of Canada Interagency Advisory Panel on Research Ethics. LM reports non-financial support from Daiichi-Sankyo and Cell & Gene Therapy Catapult. Morel reports other from UCB. LN reports other from GlaxoSmithKline, including employment and ownership of stock in GSK. RS reports personal fees from BioMed Central and Pfizer, other from NHS England, NHSx, NDC, NCRI, NIHR, MRC CTU, GeL, Glasgow CTU, UCLH, LSHTM, Cancer Research UK, Macmillan, Warwick University, Warwick CTU and University of Birmingham. AW reports grants from NIHR and Innovate UK. All other authors have completed the ICMJE uniform disclosure form and declare no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years, no other relationships or activities that could appear to have influenced the submitted work.

Patient consent for publication Not required.

Ethics approval Ethical approval was provided by the University of Birmingham Ethical Review Board (ERN_16-0819).

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Melanie Calvert <http://orcid.org/0000-0002-1856-837X>

Rebecca Mercieca-Bebber <http://orcid.org/0000-0003-3708-9099>
 Jane Blazeby <http://orcid.org/0000-0002-3354-3330>
 Amanda Hunn <http://orcid.org/0000-0003-2156-7279>
 Grace Turner <http://orcid.org/0000-0002-9783-9413>

REFERENCES

- Chan A-W, Tetzlaff JM, Altman DG, et al. Spirit 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med* 2013;158:200–7.
- Chan A-W, Tetzlaff JM, Gøtzsche PC, et al. Spirit 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;346:e7586.
- Kyte D, Duffy H, Fletcher B, et al. Systematic evaluation of the patient-reported outcome (pro) content of clinical trial protocols. *PLoS One* 2014;9:e110229.
- Mercieca-Bebber R, Friedlander M, Kok P-S, et al. The patient-reported outcome content of international ovarian cancer randomised controlled trial protocols. *Qual Life Res* 2016;25:2457–65.
- Kyte D, Retzer A, Ahmed K. Systematic evaluation of patient-reported outcome protocol content and reporting in cancer trials. *JNCI: Journal of the National Cancer Institute*, 2019: 111. 1170–8.
- Kyte D, Ives J, Draper H, et al. Current practices in patient-reported outcome (pro) data collection in clinical trials: a cross-sectional survey of UK trial staff and management. *BMJ Open* 2016;6:e012281.
- Mercieca-Bebber R, Friedlander M, Calvert M, et al. A systematic evaluation of compliance and reporting of patient-reported outcome endpoints in ovarian cancer randomised controlled trials: implications for generalisability and clinical practice. *J Patient Rep Outcomes* 2017;1:5.
- Rivera SC, Kyte DG, Aiyegbusi OL, et al. The impact of patient-reported outcome (pro) data from clinical trials: a systematic review and critical analysis. *Health Qual Life Outcomes* 2019;17:156–66.
- Kyte DG, Retzer A, Keeley T. Systematic evaluation of patient-reported outcome (pro) protocol content and reporting in cancer clinical trials: the EPIC study. *Quality of Life Research* 2017;26:55–6.
- Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols. *JAMA* 2018;319:483–94.
- Kyte D, Ives J, Draper H, et al. Inconsistencies in quality of life data collection in clinical trials: a potential source of bias? interviews with research nurses and Trialists. *PLoS One* 2013;8:e76625.
- Mercieca-Bebber R, Calvert M, Kyte D, et al. The administration of patient-reported outcome questionnaires in cancer trials: interviews with trial coordinators regarding their roles, experiences, challenges and training. *Contemp Clin Trials Commun* 2018;9:23–32.
- Moher D, Schulz KF, Simera I, et al. Guidance for developers of health research reporting guidelines. *PLoS Med* 2010;7:e1000217.
- Calvert M, Kyte D, Duffy H, et al. Patient-Reported outcome (pro) assessment in clinical trials: a systematic review of guidance for trial protocol writers. *PLoS One* 2014;9:e110216.
- Rivera S, Stephens R, Mercieca-Bebber R. "Give Us The Tools!" - Development of knowledge transfer tools to support the involvement of patient partners in clinical trial protocol development with patient-reported outcomes, in accordance with SPIRIT-PRO guideline. *BMJ*. Under review.
- FDA. Guidance for industry, patient-reported outcome measures: use in medical product development to Support labeling claims, 2009. Available: <https://www.fda.gov/media/77832/download> [Accessed 25 Nov 2019].
- FDA-NIH Biomarker Working Group. Best (biomarkers, endpoints, and other tools) resource, 2016. Available: <https://www.ncbi.nlm.nih.gov/books/NBK326791/>
- Curtis LHA, Weinfurt K. *Choosing and specifying endpoints and outcomes: introduction. in: rethinking clinical trials: a living textbook of pragmatic clinical trials*, 2019.
- Calvert M, Blazeby J, Altman DG. Reporting of patient-reported outcomes in randomized trials: the CONSORT pro extension. *JAMA* 2013;309:814–22.
- Kluetz PG, Chingos DT, Basch EM. Patient-Reported Outcomes in Cancer Clinical Trials: Measuring Symptomatic Adverse Events With the National Cancer Institute's Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). *Am Soc Clin Oncol Educ Book* 2016;35:67–73.
- ClinicalTrials.gov. Glossary of common site terms, 2020. Available: <https://clinicaltrials.gov/ct2/about-studies/glossary> [Accessed 12 Mar 2020].
- FDA. Drug Development Tool Glossary [updated 08/15/2017], 2020. Available: <https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/ddt-glossary> [Accessed 03 Jul 2020].
- Fayers PM, Hopwood P, Harvey A, et al. Quality of life assessment in clinical trials--guidelines and a checklist for protocol writers: the U.K. Medical Research Council experience. MRC Cancer Trials Office. *Eur J Cancer* 1997;33:20–8.
- Moher D, Hopewell S, Schulz KF, et al. Consort 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c869.
- Hemming K, Taljaard M, McKenzie JE, et al. Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration. *BMJ* 2018;14:k1614.
- SPIRIT-Statement, 2020. Available: <https://www.spirit-statement.org/> [Accessed 27 Apr 2020].
- Kennedy RH, Francis EA, Wharton R, et al. Multicenter randomized controlled trial of conventional versus laparoscopic surgery for colorectal cancer within an enhanced recovery programme: EnROL. *J Clin Oncol* 2014;32:1804–11.
- Staniszewska S, Brett J, Simera I, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. *BMJ* 2017;358:j3453.
- Furman RR, Sharman JP, Coutre SE, et al. Idelalisib and rituximab in relapsed chronic lymphocytic leukemia. *N Engl J Med* 2014;370:997–1007.
- Mercieca-Bebber R, Palmer MJ, Brundage M, et al. Design, implementation and reporting strategies to reduce the instance and impact of missing patient-reported outcome (pro) data: a systematic review. *BMJ Open* 2016;6:e010938.
- Trog 11.03 a randomised phase III trial of high dose palliative radiotherapy (HDPRT) versus concurrent chemotherapy + HDPRT (C-HDPRT) in patients with good performance status, locally Advanced/Small volume metastatic non small cell lung cancer (NSCLC) not suitable for radical chemo-radiotherapy (P_LUNG GP), 2012. Available: <https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=363229&showOriginal=true&isReview=true> [Accessed 05 Dec 2019].
- RATE-AF trial protocol. Available: <https://www.birmingham.ac.uk/research/activity/mds/trials/bctu/trials/portfolio-v/RateAF/index.aspx> [Accessed 21 Nov 2019].
- EMA. Ich E9 (R1) addendum on estimands and sensitivity analysis in clinical trials, 2019. Available: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf
- 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:e83–96.
- Fleming TR. Clinical trials: discerning hype from substance. *Ann Intern Med* 2010;153:400–6.
- 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:e459–69.
- Fiero MH, Roydhouse JK, Vallejo J, et al. US food and drug administration review of statistical analysis of patient-reported outcomes in lung cancer clinical trials Approved between January, 2008, and December, 2017. *Lancet Oncol* 2019;20:e582–9.
- Bottomley A, Pe M, Sloan J, 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:e510–4.
- SWOG-9208 Quality of Life and Health Status in Patients With Stage I or Stage II Hodgkin's Disease. Available: <https://www.clinicaltrials.gov/ct2/show/NCT00002563> [Accessed 25 Nov 2019].
- Brooks MM, Jenkins LS, Schron EB. Quality of life at baseline: is assessment after randomization valid? *Medical Care* 1998;36:1515–9.
- Blümle A, Meerpoel JJ, Rücker G, et al. Reporting of eligibility criteria of randomised trials: cohort study comparing trial protocols with subsequent articles. *BMJ* 2011;342:d1828.
- Calvert M, Kyte D, von Hildebrand M, et al. Putting patients at the heart of health-care research. *The Lancet* 2015;385:1073–4.
- EMA. Appendix 2 to the Guideline on the evaluation of anticancer medicinal products in man. The use of patient-reported outcome (pro) measures in oncology studies, 2016. Available: https://www.ema.europa.eu/en/documents/other/appendix-2-guideline-evaluation-anticancer-medicinal-products-man_en.pdf

- 44 Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *Jama* 1995;273:59–65.
- 45 Kluetz PG, Slagle A, Papadopoulos EJ, et al. Focusing on core patient-reported outcomes in cancer clinical trials: symptomatic adverse events, physical function, and disease-related symptoms. *Clin Cancer Res* 2016;22:1553–8.
- 46 Sloan JA, Dueck AC, Erickson PA, et al. Analysis and interpretation of results based on patient-reported outcomes. *Value Health* 2007;10(Suppl 2):S106–15.
- 47 Winther KH, Watt T, Björner JB, et al. The chronic autoimmune thyroiditis quality of life selenium trial (catalyst): study protocol for a randomized controlled trial. *Trials* 2014;15:115.
- 48 Steinemann DC, Raptis DA, Lurje G, et al. Cosmesis and body image after single-port laparoscopic or conventional laparoscopic cholecystectomy: a multicenter double blinded randomised controlled trial (SPOCC-trial). *BMC Surg* 2011;11:24.
- 49 King MT, Link EK, Whelan TJ, et al. Quality of life after breast-conserving therapy and adjuvant radiotherapy for non-low-risk ductal carcinoma in situ (big 3-07/TROG 07.01): 2-year results of a randomised, controlled, phase 3 trial. *Lancet Oncol* 2020;21:685–98.
- 50 Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61:102–9.
- 51 King MT. A point of minimal important difference (mid): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res* 2011;11:171–84.
- 52 Regnault A, Hamel J-F, Patrick DL. Pooling of cross-cultural PRO data in multinational clinical trials: how much can poor measurement affect statistical power?. *Qual Life Res* 2015;24:273–7.
- 53 Rochette A, Korner-Bitensky N, Bishop D, et al. Study protocol of the YOU CALL--WE CALL TRIAL: impact of a multimodal support intervention after a "mild" stroke. *BMC Neurol* 2010;10:3.
- 54 Hurt CN, Nixon LS, Griffiths GO, et al. Scope1: a randomised phase II/III multicentre clinical trial of definitive chemoradiation, with or without cetuximab, in carcinoma of the oesophagus. *BMC Cancer* 2011;11:466.
- 55 Kluetz PG, O'Connor DJ, Soltys K. Incorporating the patient experience into regulatory decision making in the USA, Europe, and Canada. *Lancet Oncol* 2018;19:e267–74.
- 56 Calvert M, Kyte D, Price G, et al. Maximising the impact of patient reported outcome assessment for patients and society. *BMJ* 2019;364:k5267.
- 57 Comet initiative. Available: <http://www.comet-initiative.org/>
- 58 Patrick DL, Burke LB, Powers JH, et al. Patient-Reported outcomes to support medical product labeling claims: FDA perspective. *Value Health* 2007;10(Suppl 2):S125–37.
- 59 Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1--eliciting concepts for a new PRO instrument. *Value Health* 2011;14:967–77.
- 60 Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2--assessing respondent understanding. *Value Health* 2011;14:978–88.
- 61 Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res* 2013;22:1889–905.
- 62 Prinsen CAC, Vohra S, Rose MR, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" - a practical guideline. *Trials* 2016;17:449.
- 63 Valderas JM, Ferrer M, Mendivil J, et al. Development of EMPRO: a tool for the standardized assessment of patient-reported outcome measures. *Value Health* 2008;11:700–8.
- 64 Rothrock NE, Kaiser KA, Cella D. Developing a valid patient-reported outcome measure. *Clin Pharmacol Ther* 2011;90:737–42.
- 65 Mercieca-Bebber R, King MT, Calvert MJ, et al. The importance of patient-reported outcomes in clinical trials and strategies for future optimization. *Patient Relat Outcome Meas* 2018;9:353–67.
- 66 Ahmed Ali U, Issa Y, Bruno MJ, et al. Early surgery versus optimal current step-up practice for chronic pancreatitis (escape): design and rationale of a randomized trial. *BMC Gastroenterol* 2013;13:49.
- 67 Collinson FJ, Gregory WM, McCabe C, et al. The StAR trial protocol: a randomised multi-stage phase II/III study of sunitinib comparing temporary cessation with allowing continuation, at the time of maximal radiological response, in the first-line treatment of locally advanced/metastatic renal cancer. *BMC Cancer* 2012;12:598.
- 68 Standard vs modified drug therapy (StAR), 2015. Available: <https://njl-admin.nihr.ac.uk/document/download/2024304> [Accessed 03 Mar 2020].
- 69 FDA. Methods to Identify What is Important to Patients & Select, Develop or Modify Fit-for-Purpose Clinical Outcomes Assessments, 2018. Available: <https://www.fda.gov/drugs/news-events-human-drugs/patient-focused-drug-development-guidance-methods-identify-what-important-patients-and-select>
- 70 Rutherford C, Costa D, Mercieca-Bebber R, et al. Mode of administration does not cause bias in patient-reported outcome results: a meta-analysis. *Qual Life Res* 2016;25:559–74.
- 71 Coons SJ, Gwaltney CJ, Hays RD, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (pro) measures: ISPOR ePRO good research practices Task force report. *Value Health* 2009;12:419–29.
- 72 Eremenco S, Coons SJ, Paty J, et al. Pro data collection in clinical trials using mixed modes: report of the ISPOR pro mixed modes good research practices Task force. *Value Health* 2014;17:501–16.
- 73 Lonial S, Dimopoulos M, Palumbo A, et al. Elotuzumab therapy for relapsed or refractory multiple myeloma. *N Engl J Med* 2015;373:621–31.
- 74 Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (pro) measures: report of the ISPOR Task force for translation and cultural adaptation. *Value Health* 2005;8:94–104.
- 75 Wild D, Eremenco S, Mear I, et al. Multinational trials--recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: the ISPOR patient-reported outcomes translation and linguistic validation good research practices Task force report. *Value Health* 2009;12:430–40.
- 76 Luckett T, Goldstein D, Butow PN, et al. Psychological morbidity and quality of life of ethnic minority patients with cancer: a systematic review and meta-analysis. *Lancet Oncol* 2011;12:1240–8.
- 77 Slade A, Retzer A, Ahmed K. Systematic review of the use of translated patient reported outcome measures in cancer trials. Submitted, under review.
- 78 Eortc translation procedure manual. Available: https://www.eortc.org/app/uploads/sites/2/2018/02/translation_manual_2017.pdf [Accessed 25 Nov 2019].
- 79 Herdman M, Fox-Rushby J, Badia X. 'Equivalence' and the translation and adaptation of health-related quality of life questionnaires. *Qual Life Res* 1997;6:237–47.
- 80 Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Qual Life Res* 1998;7:323–35.
- 81 Breugelmans R. Dangers in using translated medical questionnaires: the importance of conceptual equivalence across languages and cultures in patient-reported outcome measures. *Chest* 2009;136:1175–7.
- 82 Tennant A, Penta M, Tesio L, et al. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model. *Med Care* 2004;42:37–48.
- 83 Regnault A, Herdman M. Using quantitative methods within the Universalist model framework to explore the cross-cultural equivalence of patient-reported outcome instruments. *Qual Life Res* 2015;24:115–24.
- 84 Hypher RE, Brandt AE, Risnes K, et al. Paediatric goal management training in patients with acquired brain injury: study protocol for a randomised controlled trial. *BMJ Open* 2019;9:e029273.
- 85 Surr CA, Walwyn REA, Liley-Kelly A, et al. Evaluating the effectiveness and cost-effectiveness of dementia care Mapping™ to enable person-centred care for people with dementia and their carers (DCM-EPIC) in care homes: study protocol for a randomised controlled trial. *Trials* 2016;17:300.
- 86 Bausewein C, Daveson BA, Currow DC, et al. EAPC White Paper on outcome measurement in palliative care: Improving practice, attaining outcomes and delivering quality services - Recommendations from the European Association for Palliative Care (EAPC) Task Force on Outcome Measurement. *Palliat Med* 2016;30:6–22.
- 87 Evans CJ, Benalia H, Preston NJ, et al. The selection and use of outcome measures in palliative and end-of-life care research: the MORECare international consensus workshop. *J Pain Symptom Manage* 2013;46:925–37.

- 88 Jones JM, McPherson CJ, Zimmermann C, *et al.* Assessing Agreement Between Terminally Ill Cancer Patients' Reports of Their Quality of Life and Family Caregiver and Palliative Care Physician Proxy Ratings. *J Pain Symptom Manage* 2011;42:354–65.
- 89 Elliott MN, Beckett MK, Chong K, *et al.* How do proxy responses and proxy-assisted responses differ from what Medicare beneficiaries might have reported about their health care? *Health Serv Res* 2008;43:833–48.
- 90 Howland M, Allan KC, Carlton CE, *et al.* Patient-rated versus proxy-rated cognitive and functional measures in older adults. *Patient Relat Outcome Meas* 2017;8:33–42.
- 91 Evans CJ, Benalia H, Preston NJ, *et al.* The selection and use of outcome measures in palliative and end-of-life care research: the MORECare international consensus workshop. *J Pain Symptom Manage* 2013;46:925–37.
- 92 Roydhouse JK, Gutman R, Keating NL, *et al.* The association of proxy care engagement with proxy reports of patient experience and quality of life. *Health Serv Res* 2018;53:3809–24.
- 93 Pignata S, Scambia G, Ferrandina G, *et al.* Carboplatin plus paclitaxel versus carboplatin plus pegylated liposomal doxorubicin as first-line treatment for patients with ovarian cancer: the MITO-2 randomized phase III trial. *JCO* 2011;29:3628–35.
- 94 Lemieux J, Brundage MD, Parulekar WR, *et al.* Quality of life from Canadian cancer trials Group MA.17R: a randomized trial of extending adjuvant letrozole to 10 years. *J Clin Oncol* 2018;36:563–71.
- 95 Palmer MJ, Mercieca-Bebber R, King M, *et al.* A systematic review and development of a classification framework for factors associated with missing patient-reported outcome data. *Clinical Trials* 2018;15:95–106.
- 96 Fairclough DL, Peterson HF, Cella D, *et al.* Comparison of several model-based methods for analysing incomplete quality of life data in cancer clinical trials. *Stat Med* 1998;17:781–96.
- 97 Fairclough DL, Peterson HF, Chang V. Why are missing quality of life data a problem in clinical trials of cancer therapy? *Stat Med* 1998;17:667–77.
- 98 Bell ML, Fiero M, Horton NJ, *et al.* Handling missing data in RCTs; a review of the top medical journals. *BMC Med Res Methodol* 2014;14:118.
- 99 FDA. FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency 2020 [updated 21/09/2020], 2020. Available: <https://www.fda.gov/media/136238/download> [Accessed 22 Oct 2020].
- 100 Pujade-Lauraine E, Hilpert F, Weber B, *et al.* Bevacizumab combined with chemotherapy for platinum-resistant recurrent ovarian cancer: the aurelia open-label randomized phase III trial. *J Clin Oncol* 2014;32:1302–8.
- 101 Stockler MR, Hilpert F, Friedlander M, *et al.* Patient-Reported outcome results from the open-label phase III aurelia trial evaluating bevacizumab-containing therapy for platinum-resistant ovarian cancer. *J Clin Oncol* 2014;32:1309–16.
- 102 Fairclough DL. *Design and analysis of quality of life studies in clinical trials*. CRC Press, 2010.
- 103 Gladman D, Rigby W, Azevedo VF, *et al.* Tofacitinib for psoriatic arthritis in patients with an inadequate response to TNF inhibitors. *N Engl J Med* 2017;377:1525–36.
- 104 Aaronson NK, Ahmedzai S, Bergman B, *et al.* The European organization for research and treatment of cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;85:365–76.
- 105 FDA. Multiple endpoints in clinical trials. Guidance for industry, 2017. Available: <https://www.fda.gov/media/102657/download> [Accessed 19 Mar 2020].
- 106 Vergote IB, Jimeno A, Joly F, *et al.* Randomized phase III study of erlotinib versus observation in patients with no evidence of disease progression after first-line platin-based chemotherapy for ovarian carcinoma: a European organisation for research and treatment of Cancer-Gynaecological cancer group, and gynecologic cancer intergroup study. *J Clin Oncol* 2014;32:320–6.
- 107 Fairclough DL, Cella DF. Functional assessment of cancer therapy (FACT-G): non-response to individual questions. *Qual Life Res* 1996;5:321–9.
- 108 Fayers P, Aaronson N, Bjordal K. *Bottomley a, on behalf of the EORTC quality of life group: the EORTC QLQ-C30 scoring manual*. 3rd edn, 2001.
- 109 Hardouin J-B, Conroy R, Sébille V. Imputation by the mean score should be avoided when validating a patient reported outcomes questionnaire by a Rasch model in presence of informative missing data. *BMC Med Res Methodol* 2011;11:105.
- 110 Little RJ, Cohen ML, Dickersin K, *et al.* The design and conduct of clinical trials to limit missing data. *Stat Med* 2012;31:3433–43.
- 111 McIsaac MA, Cook RJ, Poulin-Costello M. Incomplete data in randomized dermatology trials: consequences and statistical methodology. *Dermatology* 2013;226:19–27.
- 112 Gamble C, Krishan A, Stocken D, *et al.* Guidelines for the content of statistical analysis plans in clinical trials. *JAMA* 2017;318:2337–43.
- 113 Mazza GCC, Pe M, Dorme L. How much missing data is too much? Monte Carlo simulations to develop SISAQOL guidelines for missing data handling. *25th annual conference of the International Society of quality of life research*, Dublin, Ireland, 2018:1–190.
- 114 Kyte D, Bishop J, Brettell E, *et al.* Use of an electronic patient-reported outcome measure in the management of patients with advanced chronic kidney disease: the RePROM pilot trial protocol. *BMJ Open* 2018;8:e026080.
- 115 Kyte D, Draper H, Calvert M. Patient-Reported outcome alerts: ethical and logistical considerations in clinical trials. *JAMA* 2013;310:1229–30.
- 116 Kyte D, Ives J, Draper H, *et al.* Management of patient-reported outcome (pro) alerts in clinical trials: a cross sectional survey. *PLoS One* 2016;11:e0144658.
- 117 Sackett DL. Clinician-trialist rounds: 5. Cointervention bias – how to diagnose it in their trial and prevent it in yours. *Clin Trials* 2011;8:440–2.
- 118 European Parliament, Council of the European Union. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/EC (General data protection regulation)*, 2016.
- 119 Calvert M, Kyte D, Mercieca-Bebber R. Supplement 3 - Summary Recommendations for Inclusion of PRO-Specific Information in the Trial Protocol, Supplementary Documents for Training for Staff, 2018. Available: <https://jamanetwork.com/journals/jama/fullarticle/2671472>
- 120 Schandelmaier S, Conen K, von Elm E, *et al.* Planning and reporting of quality-of-life outcomes in cancer trials. *Ann Oncol* 2015;26:1966–73.
- 121 Efficace F, Fayers P, Pusic A, *et al.* Quality of patient-reported outcome reporting across cancer randomized controlled trials according to the CONSORT patient-reported outcome extension: a pooled analysis of 557 trials. *Cancer* 2015;121:3335–42.
- 122 Mercieca-Bebber R, Rouette J, Calvert M, *et al.* Preliminary evidence on the uptake, use and benefits of the CONSORT-PRO extension. *Qual Life Res* 2017;26:1427–37.
- 123 Marandino L, La Salvia A, Sonetto C, *et al.* Deficiencies in health-related quality-of-life assessment and reporting: a systematic review of oncology randomized phase III trials published between 2012 and 2016. *Ann Oncol* 2018;29:2288–95.
- 124 Lombardi P, Marandino L, De Luca E, *et al.* Quality of life assessment and reporting in colorectal cancer: a systematic review of phase III trials published between 2012 and 2018. *Crit Rev Oncol Hematol* 2020;146:102877.
- 125 Marandino L, De Luca E, Zichi C, *et al.* Quality-Of-Life assessment and reporting in prostate cancer: systematic review of phase 3 trials testing anticancer drugs published between 2012 and 2018. *Clin Genitourin Cancer* 2019;17:332–47.
- 126 Mathoulin-Pelissier S, Gourgou-Bourgade S, Bonnetain F, *et al.* Survival endpoint reporting in randomized cancer clinical trials: a review of major journals. *J Clin Oncol* 2008;26:3721–6.
- 127 Reale ML, De Luca E, Lombardi P, *et al.* Quality of life analysis in lung cancer: a systematic review of phase III trials published between 2012 and 2018. *Lung Cancer* 2020;139:47–54.
- 128 Tam VC, Hotte SJ. Consistency of phase III clinical trial Abstracts presented at an annual meeting of the American Society of clinical oncology compared with their subsequent full-text publications. *J Clin Oncol* 2008;26:2205–11.
- 129 Ioannidis JPA, Greenland S, Hlatky MA, *et al.* Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;383:166–75.
- 130 Glasziou P, Altman DG, Bossuyt P, *et al.* Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014;383:267–76.
- 131 European Commission. Clinical trials regulations (EU) NO 536/2014, 2019. Available: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/regulation5362014_qa_en.pdf [Accessed 19 Mar 2020].
- 132 EU Clinical Trials. Summaries of Clinical Trial Results for Laypersons 2018 [Recommendations of the expert group on clinical trials for the implementation of Regulation (EU) No 536/2014 on clinical trials on medicinal products for human use]. Available: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/2017_01_

- 26_summaries_of_ct_results_for_laypersons.pdf [Accessed 03 May 2020].
- 133 FDA. CDER patient-focused drug development, 2019. Available: <https://www.fda.gov/drugs/development-approval-process-drugs/cder-patient-focused-drug-development> [Accessed 23 Apr 2020].
- 134 Turner G, Aiyegbusi OL, Price G, *et al.* Moving beyond project-specific patient and public involvement in research. *J R Soc Med* 2020;113:16–23.
- 135 Snyder C, Smith K, Holzner B, *et al.* Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data. *Qual Life Res* 2019;28:345–56.
- 136 Snyder CF, Smith KC, Bantug ET, *et al.* What do these scores mean? Presenting patient-reported outcomes data to patients and clinicians to improve interpretability. *Cancer* 2017;123:1848–59.
- 137 Brundage M, Blackford A, Tolbert E, *et al.* Presenting comparative study PRO results to clinicians and researchers: beyond the eye of the beholder. *Qual Life Res* 2018;27:75–90.
- 138 Haywood K, Lyddiatt A, Brace-McDonnell SJ, *et al.* Establishing the values for patient engagement (PE) in health-related quality of life (HRQoL) research: an international, multiple-stakeholder perspective. *Qual Life Res* 2017;26:1393–404.
- 139 Frank L, Basch E, Selby JV, *et al.* The PCORI perspective on patient-centered outcomes research. *JAMA* 2014;312:1513–4.
- 140 Supremo useful links. Available: <https://www.supremo-trial.com/useful-links.asp> [Accessed 19 Mar 2020].
- 141 Sloan JA, Halyard MY, Frost MH, *et al.* The Mayo clinic manuscript series relative to the discussion, Dissemination, and operationalization of the food and drug administration guidance on patient-reported outcomes. *Value Health* 2007;10(Suppl 2):S59–63.