BMJ Open

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045105.R1
Article Type:	Communication
Date Submitted by the Author:	10-Nov-2020
Complete List of Authors:	Calvert , Melanie ; University of Birmingham, Centre for Patient Reported Outcome Research, Institute of Applied Health Research King, Madeleine; University of Sydney, Faculty of Science, School of Psychology Mercieca-Bebber, Rebecca; University of Sydney, Faculty of Medicine and Health, NHMRC Clinical Trials Centre Aiyegbusi, Olalekan ; University of Birmingham, Centre for Patient Reported Outcome Research, Institute of Applied Health Research Kyte, Derek; University of Birmingham, Centre for Patient Reported Outcome Research, Institute of Applied Health Research Slade, Anita; University of Birmingham, Centre for Patient Reported Outcome Research, Institute of Applied Health Research Chan, An-Wen; University of Toronto, Women's College Research Institute, Department of Medicine Basch, E; University of North Carolina Bell, Jill; AstraZeneca Gaithersburg, Oncology Digital Health Bennett, Antonia; University of North Carolina at Chapel Hill, Cancer Outcomes Research Program Bhatnagar, Vishal; US Food and Drug Administration Blazeby, Jane; University of Bristol, Bristol Biomedical Research Centre; University Hospitals Bristol NHS Foundation Trust, Division of Surgery, Head and Neck Bottomley, Andrew; European Organisation for Research and Treatment of Cancer Brown, Julia; University of Leeds, Clinical Trials Research Unit Brundage, Michael; Queen's University Cancer Research Institute Campbell, Lisa; Medicines and Healthcare Products Regulatory Agency Cappelleri, Joseph; Pfizer Inc, Global Biometrics & Data Management - Statistics Draper, Heather; University of Warwick Dueck, Amylou; Mayo Clinic, Department of Health Sciences Research Ells, Carolyn; Interagency Advisory Panel on Research Ethics Frank, Lori; Patient-Centered Outcomes Research Institute Golub, Robert; JAMA Griebsch, Ingolf; Boehringer Ingelheim International GmbH Haywood, Kirstie; University of Warwick Warwick Medical School, Warwick Research in Nursing Hunn, Amanda; Independent Advisor King-Kallimanis, Bellinda; US Food and Drug Administration

	Martin, Laura; Daiichi Sankyo Europe GmbH Mitchell, Sandra; National Cancer Institute Morel, Thomas; UCB Pharma, Patient-Centered Outcomes Research Nelson, Linda; GSK, Value Evidence and Outcomes - Patient Centered Outcomes Norquist, Josephine; Merck & Co Inc, Center for Observational Real-world Evidence (CORE), Patient-Centered Endpoints & Strategy O'Connor, Daniel; Medicines and Healthcare Products Regulatory Agency Palmer, Michael; Queen's University, Cancer Research Institute Patrick, Donald; University of Washington Price, Gary; University of Birmingham, Centre for Patient Reported Outcome Research, Primary Care and Clinical Sciences Regnault, Antoine; Modus Outcomes Retzer, Ameeta; University of Birmingham, Centre for Patient Reported Outcome Research, Institute for Applied Health Research Revicki, Dennis; Outcomes Research Consulting Scott, Jane; Janssen Global Services LLC, Johnson and Johnson Stephens, Richard; National Cancer Research Institute Consumer Forum Turner, Grace; University of Birmingham, Centre for Patient Reported Outcome Research, Institute of Applied Health Research Valakas, Antonia; EMD Serono Inc, Healthcare business of Merck KGaA, Darmstadt, Germany Velikova, Galina; University of Leeds, Leeds Institute of Medical Research at St James's von Hildebrand, Maria; University of Birmingham, Centre for Patient Reported Outcome Research, Institute of Applied Health Research Walker, Anita; University of Birmingham, Centre for Patient Reported Outcome Research, Institute of Applied Health Research Wenzel, Lari; University of California Irvine
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Health services research
Keywords:	STATISTICS & RESEARCH METHODS, EDUCATION & TRAINING (see Medical Education & Training), Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Clinical trials < THERAPEUTICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in BMJ Open and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Research Methods & Reporting

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

Authors:

Calvert M, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, Birmingham Health Partners Centre for Regulatory Science and Innovation, NIHR Biomedical Research Centre, NIHR Surgical Reconstruction and Microbiology Research Centre and NIHR Applied Research Collaboration West Midlands, University of Birmingham, England

King MT; University of Sydney, Faculty of Science, School of Psychology, Sydney, Australia

Mercieca-Bebber R, University of Sydney, Faculty of Medicine and Health, NHMRC Clinical Trials Centre, Sydney, Australia

Aiyegbusi OL, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, Birmingham Health Partners Centre for Regulatory Science and Innovation, NIHR Biomedical Research Centre and NIHR Applied Research Collaboration West Midlands, University of Birmingham, England

Kyte D, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, England

Slade A, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, England

Chan A-W, Women's College Research Institute, Department of Medicine, University of Toronto, Ontario, Canada

Basch E, University of North Carolina, Chapel Hill

Bell JA, Oncology Digital Health, AstraZeneca, Gaithersburg, Maryland

Bennett AV, Cancer Outcomes Research Program, University of North Carolina, Chapel Hill

Bhatnagar V, US Food and Drug Administration, Silver Spring, Maryland

Blazeby J, Bristol Biomedical Research Centre, University of Bristol, Bristol, England

Bottomley A, European Organization for Research and Treatment of Cancer, Brussels, Belgium

Brown J, Clinical Trials Research Unit, University of Leeds, Leeds, England

Brundage M, Queen's Cancer Research Institute, University, Kingston, Ontario, Canada

Campbell L, Medicines and Healthcare Products Regulatory Agency, London, England

Cappelleri JC, Global Biometrics & Data Management – Statistics, Pfizer Inc, New York, NY

Draper H, University of Warwick, Coventry, England

Dueck A, Department of Health Sciences Research, Mayo Clinic, Scottsdale, Arizona

Ells C, School of Population and Global Health, McGill University, Montreal, Canada

Frank L, Patient-Centered Outcomes Research Institute, Washington, DC

Golub RM, JAMA, Chicago, Illinois

Griebsch I, Boehringer Ingelheim International GmbH, Germany

Haywood K, Warwick Research in Nursing, Warwick Medical School, University of Warwick, Coventry, England

Hunn A, Independent Advisor, UK

King-Kallimanis BL, US Food and Drug Administration, Silver Spring, Maryland

Martin L, Daiichi Sankyo Europe GmbH, Munich, Germany

Mitchell S, National Cancer Institute, Rockville, Maryland

Morel T, Patient-Centered Outcomes Research, UCB Pharma, Brussels, Belgium

Nelsen L, Value Evidence and Outcomes - Patient Centered Outcomes, GSK, Collegeville, PA, USA

Norquist JM, Center for Observational and Real-world Evidence (CORE), Patient-Centered Endpoints & Strategy, Merck & Co., Inc., Kenilworth, NJ, USA

O'Connor D, Medicines and Healthcare Products Regulatory Agency, London, England

Palmer MJ, Cancer Research Institute, Queen's University, Kingston, Ontario, Canada

Patrick D, University of Washington, Seattle

Price G, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, Birmingham, England

Regnault A, Modus Outcomes, Lyon, France

Retzer A, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, England

Revicki D, Outcomes Research Consulting, Sarasota, Florida, USA

Scott J, Janssen Global Services, Johnson and Johnson, High Wycombe, England

Stephens R, National Cancer Research Institute Consumer Forum, London, England

Turner G, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, NIHR Surgical Reconstruction and Microbiology Research Centre, University of Birmingham, England

Valakas A, EMD Serono Inc, a Healthcare business of MerckKGaA, Darmstadt, Germany, Rockland, MA, USA

Velikova G, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds, England

von Hildebrand M, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, Birmingham, England

Walker A, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, Birmingham, England

Wenzel L, University of California, Irvine

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 SPIRIT-PRO Extension. The SPIRIT-PRO Extension is a sixteen-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.

Article Summary

Strengths and limitations of this study

- The SPIRIT-PRO Extension aims to improve the completeness and transparency of trial protocols where patient-reported outcomes (PROs) are a primary or key secondary outcomes and was developed following EQUATOR 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 multi-stakeholder 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.

Background

Clinical trial protocols are essential documents intended to include the study rationale, intervention, trial design methods, study processes, outcomes, sample size, data 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 facilities external appraisal of the scientific, methodological, and ethical rigor 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 (Supplement 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 sparce 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 SPIRIT 2013 Statement (Standard Protocol Items: Recommendations for Interventional Trials). 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. In addition, we describe the development of a new PRO protocol template for use in protocol development. Additional information and resources

regarding the SPIRIT Initiative are available on the SPIRIT website (www.spirit-statement.org).

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	Please specify which page(s) of the protocol this item is addressed on.
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)			
	2b	All items from the World Health Organization Trial Registration Data Set			
Protocol version	3	Date and version identifier			
Funding	4	Sources and types of financial, material, and other support			
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	SPIRIT5a-PRO Elaboration	Specify the individual(s) responsible for the PRO content of the trial protocol.	
	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, end-point 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 summarize PRO findings in relevant studies.	
	6b	Explanation for choice of comparators			
Objectives	7	Specific objectives or hypotheses	SPIRIT7-PRO Extension	State specific PRO objectives or hypotheses (including relevant PRO concepts/domains).	
Trial design	8	Description of trial design, including type of trial (e.g., parallel group, crossover, factorial, single group), allocation ratio, and framework (e.g., superiority, equivalence, noninferiority, exploratory)			
Methods: Participants, Interventions, and Outcomes					

Study setting	9	Description of study settings (e.g., community clinic, academic hospital) and list of countries where data will be collected; reference to where list of study sites can be obtained			
Eligibility criteria	10	Inclusion and exclusion criteria for participants; if applicable, eligibility criteria for study centres and individuals who will perform the interventions (e.g., surgeons, psychotherapists)	SPIRIT10-PRO Extension	Specify any PRO-specific eligibility criteria (e.g., language/reading requirements or pre-randomisation 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.	
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 (e.g., 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 (e.g., 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 (e.g., systolic blood pressure), analysis metric (e.g., change from baseline, final value, time to event), method of aggregation (e.g., median, proportion), and time point 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 (e.g., overall health-related quality of life, specific domain, specific symptom and, for each one, the analysis metric (e.g., change from baseline, final value, time to event) and the principal time point or period of interest.	
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 et al)[1 2]	SPIRIT13-PRO Extension	Include a schedule of PRO assessments, providing a rationale for the time points, and justifying if the initial assessment is not pre-randomisation. Specify time windows, whether PRO collection is prior to clinical assessments, and, if using multiple questionnaires, whether order of administration will be standardized.	
Sample size	14	Estimated number of participants needed to achieve study objectives and how it was	SPIRIT14-PRO Elaboration	When a PRO is the primary end point, state the required sample size (and how it was	

		determined, including clinical and statistical assumptions supporting any sample size calculations		determined) and recruitment target (accounting for expected loss to follow-up). If sample size is not established based on the PRO end point, then discuss the power of the principal PRO analyses.	
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 (e.g., computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (e.g., 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 (e.g., 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 (e.g., 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					
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (e.g., duplicate measurements, training of assessors) and description of study instruments (e.g., 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 (e.g., 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.	
			SPIRIT-18a (ii)-PRO Extension	Include a data collection plan outlining the permitted mode(s) of administration (e.g., paper, telephone, electronic, other) and setting (e.g., clinic, home, other).	
			SPIRIT-18a (iii)-PRO Extension	Specify whether more than 1 language version will be used and state whether translated versions have been developed using currently recommended methods.	
			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.	
	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 minimizing 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.	

Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (e.g., double data entry; range checks for data values); reference to where details of data management procedures can be found, if not in the protocol			
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.	
	20b	Methods for any additional analyses (e.g., subgroup and adjusted analyses)			
	20c	Definition of analysis population relating to protocol nonadherence (e.g., as randomized analysis) and any statistical methods to handle missing data (e.g., multiple imputation)	SPIRIT20c-PRO Elaboration	State how missing data will be described and outline the methods for handling missing items or entire assessments (e.g., approach to imputation and sensitivity analyses).	
Methods: Monitoring					
	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			
Harms	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 standardized way. Describe how this process will be explained to participants; eg, in the participant information sheet and consent form.	
Auditing	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			
Protocol amendments	25	Plans for communicating important protocol modifications (e.g., changes to eligibility criteria,			

		outcomes, analyses) to relevant parties (e.g., investigators, research ethics committees/institutional review boards, trial participants, trial registries, journals, regulators)			
Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorized 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, health care 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 authorized 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			

The development of the SPIRIT-PRO Extension followed the Enhancing Quality and Transparency of Health Research (EQUATOR) Network's methodological framework for guideline development,[13] and has been published elsewhere.[10] Ethical approval was provided by the University of Birmingham Ethical Review Board (ERN_16-0819). 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 three-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 of 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.

Glossary

Concept: "The specific measurement goal (i.e., 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."[15]

Domain: "A sub concept represented by a score of an instrument that measures a larger concept comprised of multiple domains. For example, psychological function is a larger concept with multiple domains (emotional and cognitive function) that are measured by relevant items."[15]

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 combined[16] (e.g. change from baseline at 6 weeks in mean fatigue score).[17]

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".[18]

Important or key secondary PROs/end points: Some PRO measures (particularly health-related quality-of-life measures) are multidimensional, producing several domain-specific outcome scales; e.g., 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 (CONSORT PRO Extension 2b),[18] 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 (e.g. using the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE™)).[19]

Instrument: "A means to capture data (e.g., 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."[15]

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.[20]

Item: "an individual question, statement, or task (and its standardized response options) that is evaluated by the patient to address a particular concept".[15]

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 health care professional".[21]

Outcome*: the variable to be measured. It is the measurable characteristic that is influenced or affected by an individuals' baseline state or an intervention as in a clinical trial or other exposure[16] (e.g. a fatigue score).

Patient-reported outcome (PRO): 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.[15 18] PROs are assessed by self-reported questionnaires, referred to as PRO measures (PROMs) or instruments.[16]

Primary outcome: the most important outcome in a trial, pre-specified 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".[15]
Secondary outcomes: outcomes pre-specified in the protocol to assess additional effects of the intervention; some PROs may be identified as important or key secondary outcomes.
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials.[1 2]
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 time point whereby the result would still be deemed to be clinically relevant.[22]

* 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 (i.e. the outcome) by fully specifying four components: measurement variable (e.g. fatigue "in the past week" as measured by the QLQ-C30), analysis metric (e.g., change in fatigue from baseline, final fatigue value, time to clinically important increase in fatigue (and "event"), method of aggregation (e.g., median fatigue, proportion of patients with severe fatigue, proportion of patients with clinically important change in fatigue), and time point. 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 Explanation and Elaboration 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 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, whilst others may wish to add PRO content to existing sections of the protocol.

Modelled after other reporting guidelines,[2 23 24] 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 25] 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: Multicenter 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]"[26]

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.[27] 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 United Kingdom (UK) National Institute for Health Research (NIHR) Health Technology Assessment programme explicitly described who was responsible for the PRO component (Supplement 1).[3]

SPIRIT-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 3, Randomized, Double-Blind, Placebo-Controlled Study Evaluating the Efficacy and Safety of Idelalisib (GS-1101) in Combination with Rituximab for Previously Treated Chronic Lymphocytic Leukemia

PRO endpoints: 2°

"3.5. Endpoint Selection Rationale

3.5.2.2. Health-Related Quality of Life (HRQL)
Direct patient reporting of outcomes using standardized methods has become an increasingly important component of therapeutic assessment. Evaluation of patient-reported outcomes (PROs) is particularly relevant in patients who cannot be cured of disease [Reference].
PRO questionnaires have been previously used in CLL [chronic lymphocytic leukaemia] 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 vs 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 anemia, 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 co-morbid 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 3 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."[28]

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 (e.g. 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 (e.g. 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.[29] This information is also relevant to research ethics committees/IRBs and funders responsible for reviewing the scientific integrity and ethical aspects of the trial.

SPIRIT-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 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, high dose palliative radiotherapy (HDPRT) versus concurrent chemotherapy and HDPRT (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 six 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 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.
- 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"[30]

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 (RATE-AF)

PRO endpoints: 1°, 2°, exploratory

"3.1 Hypothesis

Null Hypothesis for primary outcome: No difference in patient-reported quality of life (measured using the physical functioning domain of the SF36 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 SF36 questionnaire).
3.2 Primary objective

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

3.3 Secondary objectives

- Generic and AF-specific patient-reported HRQL using the SF-36 global and domain-specific scores, the AFEQT overall score and the EQ-5D-5L summary index and visual analogue scale at six and twelve months.
- Echocardiographic left-ventricular ejection fraction (LVEF) and diastolic function (E/e' and composite of diastolic indices) at twelve months.
- Functional assessment, including 6-minute walking distance achieved, change in European Heart Rhythm Association (EHRA) class and cognitive function at six and twelve months.
- Change in B-type natriuretic peptide (BNP) levels as a surrogate for total cardiac strain at six months.
- Change in heart rate from baseline and group comparison using 24-hour ambulatory ECG."[31]

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 time-point (PICOT) and the estimand framework.[32] 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 SISAQOL Consortium,[33] 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.[33] Although the SISAQOL recommendations were published for oncology trials, the principles apply more broadly. Pre-specification of objectives and hypotheses encourages identification of key PRO domains and time-points. 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 health-related quality of life (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 time points during a trial, such as baseline, end of treatment, then a series of longer-term follow-ups. Pre-specification of objectives and

hypotheses – focussing on the most important PRO domains and time points – 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.[18] 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 (e.g. "to compare the effect of treatment A versus treatment B on fatigue") rather than in terms of a particular direction of effect.[2 34] In contrast, the PRO hypothesis states the predicted effect of the interventions on the trial outcomes (e.g. "patients allocated to treatment A will have less fatigue than those allocated to treatment B"). [35-37]

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 (e.g. language/reading requirements or pre-randomisation completion of PRO).[29] 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 (CARES-SF) and Cover Sheet must be completed prior to registration and randomization on SWOG-9133.[38]

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.[29] 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, e.g. anxiety/emotional well-being.[39] 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 randomization should be defined and should not be too long.

It may not always be possible to collect PROs from all study participants, e.g. 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,[40] and data on use of other language or culturally appropriate PRO instruments is often missing from the protocol.[13 41]

Where the needs of specific groups have been identified (e.g. not fluent in English) but not accommodated in the study protocol (e.g. 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.[18]

SPIRIT-12-PRO Extension: Specify the PRO concepts/domains used to evaluate the intervention (e.g., overall health-related quality of life, specific domain, specific symptom) and, for each one, the analysis metric (e.g., change from baseline, final value, time to event) and the principal time point 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 initial rate control therapy (RATE-AF)

PRO endpoints: 1°, 2°, exploratory

"12.1.1 Primary Outcome
Patient-reported quality of life (HRQL) - SF-36 physical component summary score at six months

12.1.2 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 six and twelve months
- AFEQT overall score at six and twelve months"[31]

Trial name: A Phase 3, Randomized, Double-Blind, Placebo-Controlled Study Evaluating the Efficacy and Safety of Idelalisib (GS-1101) in Combination with Rituximab for Previously Treated Chronic Lymphocytic Leukemia

PRO endpoints: 2°

"Change in HRQL domain and symptom scores based on the Functional Assessment of Cancer Therapy: Leukemia (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 randomization 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 randomization 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."[28]

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 (e.g., 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 (e.g., change from baseline, final value, time to event); the method of aggregation, which refers to the summary measure format for each study group (e.g., mean, proportion with score > 2); and the specific measurement time point 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 time-points (i.e. where greatest impact of interventions are expected), and develops that concept further by encouraging protocol contributors to think about how these PRO domains and time-points will be analysed, i.e. the analysis metric.[33] To ensure transparency and credibility of the analysis, it is recommended that there is pre-specification of the PRO concepts/domains, analysis metric(s) and time-point(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.[33 42] The selected key domains, time-points, 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, organizing 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, i.e. direct impact of the disease and treatment (e.g. symptoms such as pain, fatigue, nausea, rash and anxiety) or more distal, "knock-on" effects, (e.g. functional status and global quality of life), as illustrated for ovarian cancer in (Figure 1, inspired by the Wilson and Cleary model[43]). 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.[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 pre-specified alongside the levels of statistical and clinical significance for the study and any responder definition in use.[15] Time-points for analysis should be chosen to best address the research question, whilst 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.[15 45]

The example, Idelalisib and Rituximab Improve Progression-Free Survival Over Rituximab Alone in Unfit Patients with Relapsed Chronic Lymphocytic Leukemia: A Phase 3 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 time-point (6 months) as the primary outcome, with other SF-36 domains, questionnaires and time-points specified as secondary outcomes.

Figure 1 Proximal and distal effects of therapy on patient symptoms and quality of life adapted from Wilson and Cleary, 1995.[43]

SPIRIT-13-PRO Extension: Include a schedule of PRO assessments, providing a rationale for the time points, and justifying if the initial assessment is not pre-randomisation. 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 high dose palliative radiotherapy (HDPRT) versus concurrent chemotherapy + HDPRT (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).[30]

PRO endpoints: 1°, 2°

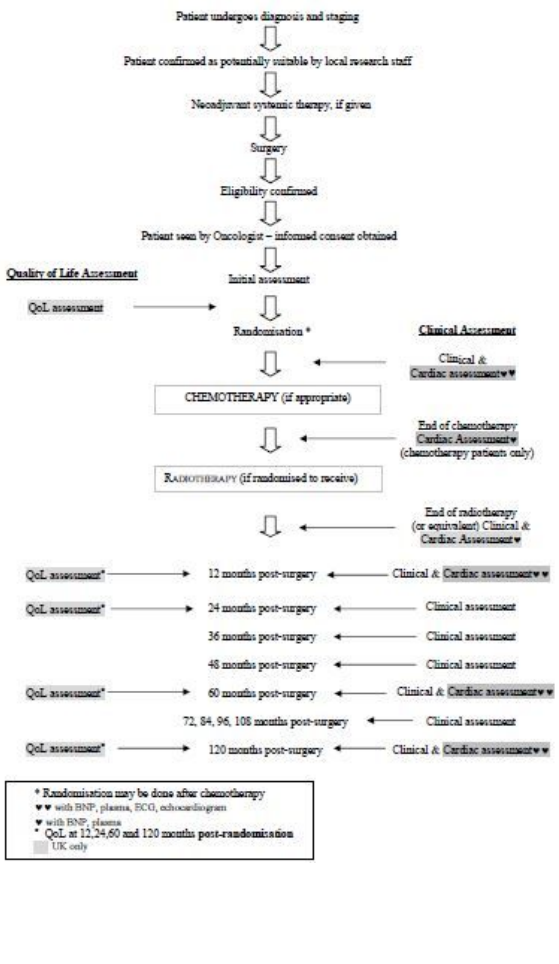
Assessment Time Point	Beginning of Acceptable time limit	End of Acceptable time limit
1. Pre-Registration	Date of signing consent	Date patient made aware of treatment arm to which they have been assigned
2. Pre-Treatment	Date patient made aware of treatment arm to which they have been assigned	Date of 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

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))

PRO endpoints: 2°

Visits[a] / Investigations	Patients Involved	Screening	Post (+/-) chemo pre (+/-) RT	Post (+/-) RT	1 yr	2 yr	3 yr	4 yr	5 yr	6 yr	7 yr	8 yr	9 yr	10 yr	Recurrence[b]
		Baseline 1	2	3	4	5	6	7	8	9	10	11	12	13	
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[c]	All	X													
Tumour paraffin block at recurrence if available[c]	All														X
Acute/ Late morbidity[d]	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 (d)	If consented to QOL sub study	X			X	X			X					X	

(d) 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.[46]



Explanation

A clear and concise schedule of PRO assessments can: assist trial staff to be organised and prepared for participant visits, inform study participants about the methods and expectations of trial participation, and facilitate review of participant burden by research ethics committees/IRBs.[29] The scheduled PRO assessments should provide the data required to address the study's PRO objectives. When selecting appropriate time-points 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 (e.g. 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 pre-randomisation 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 post-baseline, or on Day 30 +/- 3 days post-surgery. Time-windows for each scheduled PRO assessment require an unambiguous reference point, to ensure that PRO data collection captures clinically relevant time points 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, e.g. 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 (i.e. no net effect on estimates of interest, such as the impact of interventions on PROs) or systematic (i.e. causing bias).

SPIRIT-14-PRO Elaboration: When a PRO is the primary end point, 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 end point, 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 four points on the 0 to 100 ThyPRO composite scale, corresponding to a small to moderate effect. In previously obtained data, the standard deviation of ThyPRO-scores (sigma level) was 20 points. With a correlation between observations on the 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 multicenter 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 cosmesis and body image score (CBIS) is defined as an improvement of 20% of the cosmesis score (8 points). Given the reported standard deviation of the CBIS between 4-6 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 (alpha) of 0.05 and a power of 0.90. Estimating a 10% dropout rate, which is common in randomized controlled trials, 55 patients will be randomized per arm."[48]

Trial group: TROG

Trial name: A randomised phase III study of radiation doses and fractionation schedules for ductal carcinoma in situ (DCIS) 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 standard deviations of a continuous scale such as fatigue or physical symptoms, with 80% power at a two-sided alpha 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.[45] Ideally, the criteria for clinical significance (e.g. 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.[15]

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 (e.g., 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 (e.g. SF-36, Stroke Impact Scale, Quality of Life Index) and the tool chosen was a compromise between psychometric properties and adequacy of content for mild stroke. The 32 item questionnaire Quality of Life Index (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, socio-economic, psychological/spiritual and family), is evaluated in terms of satisfaction and importance on a six-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: a = 0.90) [Reference] and thus should be responsive to therapy-induced change [References]. A one point difference was observed in the first six months post-stroke descriptive follow-up (n = 63) for an effect size of 0.33 [References]. A two-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 multi-centre 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 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 wellbeing.[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.[15 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 (COSMIN) 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 (e.g. 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 standardized 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.[18]

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 (e.g., 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.[29]

SPIRIT-18a(ii)-PRO Extension: Include a data collection plan outlining the permitted mode(s) of administration (e.g., paper, telephone, electronic, other) and setting (e.g., clinic, home, other).

Examples

Trial name: Early surgery versus optimal current step-up practice for chronic pancreatitis: a multi-centre randomised control trial (ESCAPE)

PRO endpoints: 1°, 2°

"The Izbicki pain score will be assessed every two 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 Izbiki pain score forms and return envelops. Patients will be contacted by telephone every two 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 less than 3 minutes. 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 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 (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™ (FACT-G and 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 two-weekly questionnaires which participants complete at home from the 24-week time-point, where the participant consents to this, reminders for completion are sent by email or text message to the participant by the research team at 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 pre-planned 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.[29] Of note, the FDA has previously recommended that there is a back-up plan for electronic PRO data collection (e.g., web-, phone-, 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 (ISPOR) 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 standardized 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).

Further, 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 3, Randomized, Open Label Trial of Lenalidomide/dexamethasone With or Without Elotuzumab in Relapsed or Refractory Multiple Myeloma (ELOQUENT – 2)

PRO endpoints: 2°, exploratory

"5.8 Outcomes Research Assessments
5.8.1 HRQL Assessments
To assess the impact of treatment, subject's quality of life will be measured using 3 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 to 10 on any single item is generally considered to be clinically significant.[*Reference*] Pre-testing was carried out in the UK, Norway, Sweden, Denmark, and Germany. Field testing of the module has been conducted in a range of Phase 3 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- 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 generalizability 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 National Institute for Health Research 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, multi-national 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.[18 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, pre-specification 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 Pediatric Acquired Brain Injury - a Randomized 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 EF 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 (rs=0.88 for teachers, .82 for parents), internal consistency (Cronbach's αs=0.80– 0.98), and moderate correlations have been detected between teacher and parent ratings (rs=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 (i.e. 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."[15 42] The US FDA also discourages the use of proxy-reported outcomes to inform labelling claims, recommending observer reports for observable phenomenon only (e.g. 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,

dependant 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 (e.g. age, relationship to the patient, gender, proxy literacy, relationship and exposure to the patient,[92]) 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.

SPIRIT-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 randomized multicentre trial of Carboplatin + Liposomal Doxorubicin vs Carboplatin + Paclitaxel in patients with ovarian cancer (MITO-2 (Multicentre Italian Trials in Ovarian Cancer-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 randomization to Letrozole or placebo for women previously diagnosed with primary breast cancer completing five 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°

"4.9 Quality of Life
Mandatory for NCIC CTG centres and optional for centers within other cooperative groups:
Patient is able (i.e. 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 randomization. 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 sources[95] 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; or when the PRO assessment is overly burdensome on the patient (e.g. 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.[29] 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,[29] 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 (e.g.

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.[29] 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.[29]

The MITO-2 and MA.17R examples above illustrate some of these strategies. MA.17R comes from the Canadian Cancer Trials Group (CCTG), a multi-centre 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 pre-randomisation 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 maximize compliance.

Our prior work suggests that few trials actively specify such procedures in their protocols; 46.7% HTA protocols[3] and 38.5% of international ovarian cancer trials[4] 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 multi-centre, open-label, randomised, two-arm Phase III trial on the effect on progression free survival 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 multi-centre, 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 progression-free survival (PFS); quality of life was a secondary endpoint. The protocol stated that:

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

The protocol also specified post-progression 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 post-progression, which may explain some inconsistency among sites, with some sites collecting PRO data post-progression, and other not.

As stated in the AURELIA PRO paper (*ref, 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 minimize bias,[102] ensures that staff collect all required PRO data in a standardized 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 post-discontinuation/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 3, randomized double-blind, placebo-controlled study of the efficacy and safety of 2 doses of Tofacitinib (Cp-690,550) in subjects with active psoriatic arthritis and an inadequate response to at least one TNF inhibitor (OPAL BEYOND)

PRO endpoints: 1°, 2°, exploratory

"9.2.1. Analysis of Primary Endpoint

There are two primary endpoints in A3921125 and two doses of tofacitinib each of which will be compared to placebo for each endpoint. In order to control for Type I error a step-wise 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 10mg ACR20 response rate at 3 months, tofacitinib 5mg ACR20 response rate at 3 months, tofacitinib 10mg HAQ-DI at 3 months, tofacitinib 5mg HAQ-DI at 3 months. This gate-keeping or step-down approach strongly protects the Type I error rate at the 0.05 (2-sided) level…

9.2.1.1. 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 confidence intervals for the differences. The primary analysis will be ACR20 response rate at Month 3. The ACR20 response rate will also be analysed at other time points as a secondary analysis. Missing values due to a subject dropping from the study for any reason (e.g., lack of efficacy or adverse event) will be handled by setting the ACR20 value to nonresponsive.

9.2.1.2. 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 nonresponsive. 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.

9.2.2. 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 10mg will be tested vs placebo first, followed by tofacitinib 5mg. 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 (e.g., remission rates) will follow the analyses described above for binary variables (e.g., ACR20) and continuous endpoints will follow the same type of analyses described above for continuous endpoints (e.g., 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).[45] This can be contained by pre-specifying the key PRO domain(s) or overall score of interest and the principal time-point(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 end points are tested in a defined sequence that contains the overall type I error to the desired level, or conventional non-hierarchical methods (e.g. Bonferroni correction), should be specified *a* priori.[15] 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 pre-specified 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 (e.g., 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°

"10.3.2 Compliance
Missing data may hamper assessment of HRQL in clinical trials. This may occur because centers 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 less than 2% on average and should not be a problem. The former problem will be minimized by ensuring that participating centers are properly informed and motivated towards HRQL assessment. During the study, compliance with completing HRQL questionnaires will be investigated at each time point. 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 time point using a chi-square 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 bi-annual compliance reviews, then the Protocol Writing Committee would review this to either improve compliance or consider terminating the HRQL assessment in the trial…

10.4.3 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 nonresponse (the whole questionnaire is missing for a patient). As item non-response occurs less than 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 modeling techniques (i.e. 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 (i.e. there are no PRO data available for analysis from the participant at that assessment time point). 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.[33 109]

2) Handling missed assessments: missing assessments present a major problem for PRO analyses; lead to a loss of power and wider confidence intervals 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 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,[29] 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 (i.e., 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.[33] Additionally, SISAQOL is developing a set of macros to describe patterns of missing data, and to evaluate imputation methods for use in sensitivity analysis.[33 35 37]

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; e.g., in the participant information sheet 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 wellbeing 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 wellbeing 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 wellbeing 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]

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 (for example, 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 i.e., bias caused by "any intervention other than the experimental maneuver 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 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 is 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.

Box Regulatory/HTA perspective

Regulatory agencies, such as the MHRA, EMA and FDA are placing more focus on capturing the patient experience when developing drugs.[15 42 55 119] 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.

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 wellbeing 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 health-care 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 is 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 health-care research: the patient.’[120]

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 didn’t know before, and how that will be used to help people. Because that’s why we participate in research; it’s to help people like us.

Protocol template

The PRO protocol template [journal to link to template] 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

Supplement 2 outlines additional items recommended for inclusion in other trial documentation, such as the SAP, participant information sheet, and training and guidance documents for staff.[10 121] 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 (CRO) 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.[1 2 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-centered 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 results[3] and use of PRO data to inform patient-centered care. Several SPIRIT-PRO items correspond to items on the CONSORT (Consolidated Standards of Reporting Trials) -PRO checklist.[10 18] 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 9 122-125] Worryingly, a recent review of cancer trials suggested that 49,568 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.[122] Non-reporting of PRO findings is widespread,[7 9 122-130] 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 health-care research resources.[131-133] 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.[134] 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,[135] 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.[119 136]

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 institutional review boards 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),[33] reporting (CONSORT-PRO),[13] and presentation of results (Figure 2).[137-139] These tools are currently being disseminated by the PROTEUS 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 methodologic 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.[10 140 141]

Figure 2 Resources to promote high quality PRO trial design, analysis and dissemination. Adapted from [9]

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.

Article Information

Corresponding Author: Melanie Calvert, PhD, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, Edgbaston, Birmingham B15 2TT, England (m.calvert@bham.ac.uk).

In memory of Doug Altman (1948-2018), statistician, pioneer, luminary. “To maximise the benefit to society, you need to not just do research, but do it well”.

Contributorship: Drs MJ Calvert and MT King 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. Drs MJ Calvert and MT King are cochairs of the SPIRIT-PRO Group.

Concept and design: MJ Calvert, MT King.

Acquisition, analysis, or interpretation of data: MJ Calvert, MT King, R Mercieca-Bebber, OL Aiyegbusi, D Kyte, A Slade, A-W Chan, E Basch, JA Bell, AV Bennett, V Bhatnagar, J Blazeby, A Bottomley, J Brown, M Brundage, L Campbell, JC Cappelleri, H Draper, A Dueck, C Ells, L Frank, RM Golub, I Griebsch, K Haywood, BL King-Kallimanis, L Martin, S Mitchell, T Morel, L Nelson, JM Norquist, D O’Connor, MJ Palmer, D Patrick, G Price, A Regnault, A Retzer, D Revicki, J Scott, R Stephens, G Turner, A Valakas, G Velikova, M von Hildebrand, A Walker, L Wenzel.

Drafting of the manuscript: MJ Calvert, D Kyte, R Mercieca-Bebber, A Slade, A-W Chan, MT King.+ section writers G Velikova, A Regnault, D Revicki, AV Bennett, S Mitchell, L Wenzel, MJ Palmer, J Brown, G Turner, A Retzer, A Walker

Critical revision of the manuscript for important intellectual content: MJ Calvert, MT King, R Mercieca-Bebber, OL Aiyegbusi, D Kyte, A Slade, A-W Chan, E Basch, JA Bell, AV Bennett, V Bhatnagar, J Blazeby, A Bottomley, J Brown, M Brundage, L Campbell, JC Cappelleri, H Draper, A Dueck, C Ells, L Frank, RM Golub, I Griebsch, K Haywood, BL King-Kallimanis, L Martin, S Mitchell, T Morel, L Nelson, JM Norquist, D O’Connor, MJ Palmer, D Patrick, G Price, A Regnault, A Retzer, D Revicki, J Scott, R Stephens, G Turner, A Valakas, G Velikova, M von Hildebrand, A Walker, L Wenzel.

Obtained funding: MJ Calvert, D Kyte, MT King, D Altman, J Blazeby, J Brown, M Brundage, J Coast, H Draper, M von Hildebrand, J Ives, R Mercieca-Bebber, G Price, L Roberts, A Slade.

Supervision: MJ Calvert

Conflict of Interest Disclosures: 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. Calvert 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. Calvert has received personal fees from Astellas, Takeda, Merck, Daiichi Sankyo,

Glaukos, GSK and the Patient-Centered Outcomes Research Institute (PCORI) outside the submitted work. Mercieca-Bebber reports non-financial support from University of Birmingham. Aiyegbusi, Kyte, and Retzer, reports grants from National Institute of Health Research (NIHR). Aiyegbusi and Kyte report grants from Birmingham Biomedical Research Centre (BRC). Aiyegbusi reports grants from UCB Pharma and also receives funding from the Health Foundation and declares personal fees from Gilead Sciences Ltd. Kyte and Retzer reports grants from Innovate UK and Macmillan Cancer Support. Kyte 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. Basch 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. Cappelleri reports other from Pfizer Inc., and is an employee and a stockholder of Pfizer Inc. Griebsch is a fully paid employee of Boehringer Ingelheim International GmbH. Ells is Chair of the Government of Canada Interagency Advisory Panel on Research Ethics. Martin reports non-financial support from Daiichi-Sankyo and Cell & Gene Therapy Catapult. Morel reports other from UCB. Nelson reports other from GlaxoSmithKline, including employment and ownership of stock in GSK. Stephens 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. Walker 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 three years, no other relationships or activities that could appear to have influenced the submitted work.

Funding/Support: 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

Calvert 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. King is supported by the Australian government through Cancer Australia. Mercieca-Bebber is supported by the Australian Government by a National Health and Medical Research Council (NHMRC) research fellowship. Blazeby is supported by the NIHR Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust. She is also an NIHR Senior Investigator.

Role of the Funder/Sponsor: 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. Work was funded by an unrestricted grant from UCB

Pharma, and a participant (Thomas Morel) contributed as a co-author and member of the industry advisory group.

Disclaimer: The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, the NIHR, the Department of Health, or any other employer or institution. Views of authors, Delphi participants, and stakeholder participants are individual views and may not represent the views of the broader stakeholder group or host institution. This work reflects the views of the authors and should not be construed in any way to represent the views or policies of the US Food and Drug Administration, Medicines and Healthcare products Regulatory Agency (MHRA), or any other employer or institution.

Additional Contributions and Acknowledgements: With thanks to 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] eAppendix in Supplement 1 made by the SPIRIT-PRO Executive, the ISOQOL Best Practices for PROs in Randomized Clinical Trials Protocol Checklist Taskforce, the international stakeholders responsible for stakeholder survey distribution and stakeholders who completed the stakeholder survey, the Delphi panellists andthe SPIRIT-PRO International Consensus Meeting Participants.

Data sharing

References to protocols available in the public domain are provided. Permissions to access unpublished protocols may be sought via request to Melanie Calvert: m.calvert@bham.ac.uk. Requests will be forwarded to the relevant research team for consideration.

We would like to thank the trial groups, the Trans Tasman Radiation Oncology Group, the Australasian Leukaemia and Lymphoma Group, Breast International Group (BIG), Canadian Cancer Trials Group (CCTG), Cancer Trials Ireland, European Organisation for Research and Treatment of Cancer (EORTC), International Breast Cancer Study Group (IBCSG) and Scottish Cancer Trials Breast Group (SCTBG); and the individual investigators who granted permission for us to publish excerpts from their trial protocols; and those trial teams that made their protocols publicly available through the published domain.

summaries, extracts and/or, abstracts of the Contribution and convert or allow conversion into any format including without limitation audio, iii) create any other derivative work(s) based in whole or part on the on the Contribution, iv) to exploit all subsidiary rights to exploit all subsidiary rights that currently exist or as may exist in the future in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above. All research articles will be made available on an open access basis (with authors being asked to pay an open access fee—see http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse). The terms of such open access shall be governed by a Creative Commons licence—details as to which Creative Commons licence will apply to the research article are set out in our worldwide licence referred to above.

References

1. Chan A-W, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. *Annals of Internal Medicine* 2013;158(3):200-07. doi: 10.7326/0003-4819-158-3-201302050-00583
2. Chan A-W, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;8(346) doi: 10.1136/bmj.e7586
3. 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(10):e110229. doi: 10.1371/journal.pone.0110229 [published Online First: 2014/10/22]
4. Mercieca-Bebber R, Friedlander M, Kok PS, et al. The patient-reported outcome content of international ovarian cancer randomised controlled trial protocols. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2016;25(10):2457-65. doi: 10.1007/s11136-016-1339-x [published Online First: 2016/06/14]
5. Kyte D, Retzer A, Ahmed K, et al. Systematic Evaluation of Patient-Reported Outcome Protocol Content and Reporting in Cancer Trials. *JNCI: Journal of the National Cancer Institute* 2019 doi: 10.1093/jnci/djz038
6. 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(10):e012281. doi: 10.1136/bmjopen-2016-012281 [published Online First: 2016/10/05]
7. 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(1):5. doi: 10.1186/s41687-017-0008-3 [published Online First: 2017/01/01]
8. 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 Out* 2019;17(1):156-56. doi: 10.1186/s12955-019-1220-z

9. Kyte DG, Retzer A, Keeley T, et al. Systematic evaluation of Patient-Reported Outcome (PRO) protocol content and reporting in cancer clinical trials: the EPiC study. *Quality of Life Research* 2017;26(1):55-56.
10. Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for Inclusion of Patient-Reported Outcomes in Clinical Trial Protocols: The SPIRIT-PRO Extension. *JAMA* 2018;319(5):483-94. doi: 10.1001/jama.2017.21903 [published Online First: 2018/02/08]
11. 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(10):e76625. doi: 10.1371/journal.pone.0076625 [published Online First: 2013/10/15]
12. 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. *Contemporary Clinical Trials Communications* 2018;9:23-32. doi: 10.1016/j.conctc.2017.11.009
13. Moher D, Schulz KF, Simera I, et al. Guidance for Developers of Health Research Reporting Guidelines. *PLOS Medicine* 2010;7(2):e1000217. doi: 10.1371/journal.pmed.1000217
14. 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(10):e110216. doi: 10.1371/journal.pone.0110216
15. FDA. Guidance for Industry, Patient-Reported Outcome Measures: Use in Medical Product Development to Suport Labeling Claims 2009 [Available from: https://www.fda.gov/media/77832/download accessed 25/11/2019.
16. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource 2016 [Available from: https://www.ncbi.nlm.nih.gov/books/NBK326791/.
17. Curtis L HA, Weinfurt K. Choosing and Specifying Endpoints and Outcomes: Introduction. In: Rethinking Clinical Trials: A Living Textbook of Pragmatic Clinical Trials. 2019 [Available from: DOI: 10.28929/009 accessed 12/03/2020.

18. Calvert M, Blazeby J, Altman DG, et al. Reporting of Patient-Reported Outcomes in Randomized Trials: The CONSORT PRO Extension. *JAMA* 2013;309(8):814-22. doi: 10.1001/jama.2013.879
19. Kluetz PG, Chingos DT, Basch EM, et al. 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. doi: 10.14694/edbk_159514 [published Online First: 2016/06/02]
20. ClinicalTrials.gov. Glossary of Common Site Terms [Available from: https://clinicaltrials.gov/ct2/about-studies/glossary accessed 12/03/2020 2020.
21. FDA. Drug Development Tool Glossary [updated 08/15/2017. Available from: https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/ddt-glossary accessed 03/07/2020 2020.
22. 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. *European Journal of Cancer* 1997;33(1):20-28. doi: https://doi.org/10.1016/S0959-8049(96)00412-1
23. 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. doi: 10.1136/bmj.c869
24. 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;363 doi: 10.1136/bmj.k1614
25. SPIRIT-Statement [Available from: https://www.spirit-statement.org/ accessed 27/04/2020 2020.
26. 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. *Journal of Clinical Oncology* 2014;32(17):1804-11. doi: 10.1200/jco.2013.54.3694
27. 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. doi: 10.1136/bmj.j3453 [published Online First: 2017/08/05]

28. Furman RR, Sharman JP, Coutre SE, et al. Idelalisib and Rituximab in Relapsed Chronic Lymphocytic Leukemia. *New England Journal of Medicine* 2014;370(11):997-1007. doi: 10.1056/NEJMoa1315226
29. 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(6):e010938. doi: 10.1136/bmjopen-2015-010938 [published Online First: 2016/06/18]
30. 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 from: https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=363229&showOriginal=true&isReview=true accessed 05/12/2019.
31. RATE-AF Trial Protocol [Available from: https://www.birmingham.ac.uk/research/activity/mds/trials/bctu/trials/portfolio-v/RateAF/index.aspx accessed 21/11/2019.
32. EMA. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials 2019 [Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf.
33. 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. *The Lancet Oncology* 2020;21(2):e83-e96. doi: 10.1016/S1470-2045(19)30790-9
34. Fleming TR. Clinical trials: discerning hype from substance. *Ann Intern Med* 2010;153(6):400-6. doi: 10.7326/0003-4819-153-6-201009210-00008 [published Online First: 2010/09/22]
35. 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. *The Lancet Oncology* 2018;19(9):e459-e69. doi: https://doi.org/10.1016/S1470-2045(18)30418-2

36. 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(10):e582-e89. doi: 10.1016/s1470-2045(19)30335-3 [published Online First: 2019/10/04]
37. 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. *The Lancet Oncology* 2016;17(11):e510-e14. doi: https://doi.org/10.1016/S1470-2045(16)30510-1
38. SWOG-9208 Quality of Life and Health Status in Patients With Stage I or Stage II Hodgkin's Disease [Available from: https://www.clinicaltrials.gov/ct2/show/NCT00002563 accessed 25/11/2019.
39. Brooks MM, Jenkins LS, Schron EB, et al. Quality of Life at Baseline: Is Assessment After Randomization Valid? *Medical Care* 1998;36(10):1515-19.
40. Blümle A, Meerpohl JJ, Rücker G, et al. Reporting of eligibility criteria of randomised trials: cohort study comparing trial protocols with subsequent articles. *BMJ* 2011;342:d1828. doi: 10.1136/bmj.d1828
41. Calvert M, Kyte D, von Hildebrand M, et al. Putting patients at the heart of health-care research. *The Lancet* 2015;385(9973):1073-74. doi: https://doi.org/10.1016/S0140-6736(15)60599-2
42. 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 from: https://www.ema.europa.eu/en/documents/other/appendix-2-guideline-evaluation-anticancer-medicinal-products-man_en.pdf.
43. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *Jama* 1995;273(1):59-65. [published Online First: 1995/01/04]
44. 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(7):1553-8. doi: 10.1158/1078-0432.CCR-15-2035 [published Online First: 2016/01/14]

45. 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. doi: 10.1111/j.1524-4733.2007.00273.x [published Online First: 2007/11/13]
46. Supremo useful links [Available from: https://www.supremo-trial.com/useful-links.asp accessed 19/03/20 2020.
47. Winther KH, Watt T, Bjørner JB, et al. The chronic autoimmune thyroiditis quality of life selenium trial (CATALYST): study protocol for a randomized controlled trial. *Trials* 2014;15(1):115. doi: 10.1186/1745-6215-15-115
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 Surgery* 2011;11(1):24. doi: 10.1186/1471-2482-11-24
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. *The Lancet Oncology* 2020 doi: https://doi.org/10.1016/S1470-2045(20)30085-1
50. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology* 2008;61(2):102-09. doi: https://doi.org/10.1016/j.jclinepi.2007.03.012
51. King MT. A point of minimal important difference (MID): a critique of terminology and methods. *Expert Review of Pharmacoeconomics & Outcomes Research* 2011;11(2):171-84. doi: 10.1586/erp.11.9
52. Regnault A, Hamel JF, Patrick DL. Pooling of cross-cultural PRO data in multinational clinical trials: how much can poor measurement affect statistical power? *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2015;24(2):273-7. doi: 10.1007/s11136-014-0765-x [published Online First: 2014/09/06]

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 Neurology* 2010;10(1):3. doi: 10.1186/1471-2377-10-3
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. doi: 10.1186/1471-2407-11-466 [published Online First: 2011/11/01]
55. Kluetz PG, O'Connor DJ, Soltys K. Incorporating the patient experience into regulatory decision making in the USA, Europe, and Canada. *The Lancet Oncology* 2018;19(5):e267-e74. doi: https://doi.org/10.1016/S1470-2045(18)30097-4
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. doi: 10.1136/bmj.k5267 [published Online First: 2019/01/27]
57. COMET Initiative [Available from: http://www.comet-initiative.org/ accessed 25/11/19.
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. doi: 10.1111/j.1524-4733.2007.00275.x [published Online First: 2007/11/13]
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(8):967-77. doi: 10.1016/j.jval.2011.06.014 [published Online First: 2011/12/14]
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(8):978-88. doi: 10.1016/j.jval.2011.06.013 [published Online First: 2011/12/14]

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. *Quality of Life Research* 2013;22(8):1889-905. doi: 10.1007/s11136-012-0344-y
62. Prinsen CA, 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(1):449. doi: 10.1186/s13063-016-1555-2 [published Online First: 2016/09/14]
63. Valderas JM, Ferrer M, Mendívil J, et al. Development of EMPRO: A Tool for the Standardized Assessment of Patient-Reported Outcome Measures. *Value in Health* 2008;11(4):700-08. doi: 10.1111/j.1524-4733.2007.00309.x
64. Rothrock NE, Kaiser KA, Cella D. Developing a Valid Patient-Reported Outcome Measure. *Clinical Pharmacology & Therapeutics* 2011;90(5):737-42. doi: 10.1038/clpt.2011.195
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 related outcome measures* 2018;9:353-67. doi: 10.2147/PROM.S156279
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 Gastroenterology* 2013;13(1):49. doi: 10.1186/1471-230X-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(1):598. doi: 10.1186/1471-2407-12-598
68. Standard vs Modified Drug Therapy (STAR) 2015 [Available from: https://njl-admin.nihr.ac.uk/document/download/2024304 accessed 03/05/20 2020.
69. FDA. Methods to Identify What is Important to Patients & Select, Develop or Modify Fit-for-Purpose Clinical Outcomes Assessments 2018 [Available from: 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. *Quality of Life Research* 2016;25(3):559-74. doi: 10.1007/s11136-015-1110-8

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 in Health* 2009;12(4):419-29. doi: https://doi.org/10.1111/j.1524-4733.2008.00470.x

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(5):501-16. doi: 10.1016/j.jval.2014.06.005 [published Online First: 2014/08/17]

73. Lonial S, Dimopoulos M, Palumbo A, et al. Elotuzumab Therapy for Relapsed or Refractory Multiple Myeloma. *New England Journal of Medicine* 2015;373(7):621-31. doi: 10.1056/NEJMoa1505654

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 in Health* 2005;8(2):94-104. doi: 10.1111/j.1524-4733.2005.04054.x

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 in Health* 2009;12(4):430-40. doi: http://dx.doi.org/10.1111/j.1524-4733.2008.00471.x

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. *The Lancet Oncology* 2011;12(13):1240-48. doi: https://doi.org/10.1016/S1470-2045(11)70212-1

77. Slade A, Retzer A, Ahmed K, et al. Systematic Review of the Use of Translated Patient Reported Outcome Measures in Cancer Trials. Submitted, under review.

78. EORTC Translation Procedure Manual [Available from: https://www.eortc.org/app/uploads/sites/2/2018/02/translation_manual_2017.pdf accessed 25/11/2019.

79. Herdman M, Fox-Rushby J, Badia X. 'Equivalence' and the translation and adaptation of health-related quality of life questionnaires. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 1997;6(3):237-47. doi: 10.1023/a:1026410721664 [published Online First: 1997/04/01]

80. Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 1998;7(4):323-35. doi: 10.1023/a:1024985930536 [published Online First: 1998/06/04]

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(4):1175-77. doi: https://doi.org/10.1378/chest.09-1684

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 - The PRO-ESOR project. *Medical Care* 2004;42(1):37-48. doi: 10.1097/01.mlr.0000103529.63132.77

83. Regnault A, Herdman M. Using quantitative methods within the Universalist model framework to explore the cross-cultural equivalence of patient-reported outcome instruments. *Quality of Life Research* 2015;24(1):115-24. doi: 10.1007/s11136-014-0722-8

84. Hypher R, Brandt A, 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. doi: 10.1136/bmjopen-2019-029273

85. Surr CA, Walwyn REA, Lilley-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(1):300. doi: 10.1186/s13063-016-1416-z

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. *Palliative medicine* 2016;30(1):6-22. doi: 10.1177/0269216315589898 [published Online First: 2015/06/13]
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. *Journal of pain and symptom management* 2013;46(6):925-37. doi: https://doi.org/10.1016/j.jpainsymman.2013.01.010
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. *Journal of pain and symptom management* 2011;42(3):354-65. doi: 10.1016/j.jpainsymman.2010.11.018 [published Online First: 2011/04/02]
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 services research* 2008;43(3):833-48. doi: 10.1111/j.1475-6773.2007.00820.x [published Online First: 2008/05/06]
90. Howland M, Allan KC, Carlton CE, et al. Patient-rated versus proxy-rated cognitive and functional measures in older adults. *Patient related outcome measures* 2017;8:33-42. doi: 10.2147/prom.s126919 [published Online First: 2017/03/30]
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. *Journal of pain and symptom management* 2013;46(6):925-37. doi: 10.1016/j.jpainsymman.2013.01.010 [published Online First: 2013/05/01]
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 services research 2018;53(5):3809-24. doi: 10.1111/1475-6773.12980 [published Online First: 2018/05/29]

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. *J Clin Oncol* 2011;29(27):3628-35. doi: 10.1200/jco.2010.33.8566 [published Online First: 2011/08/17]
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. *Journal of Clinical Oncology* 2018;36(6):563-71. doi: 10.1200/JCO.2017.75.7500
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. *Clin Trials* 2018;15(1):95-106. doi: 10.1177/1740774517741113 [published Online First: 2017/11/11]
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. *Statistics in medicine* 1998;17(5-7):781-96. doi: 10.1002/(sici)1097-0258(19980315/15)17:5/7<781::aid-sim821>3.0.co;2-o [published Online First: 1998/04/29]
97. Fairclough DL, Peterson HF, Chang V. Why are missing quality of life data a problem in clinical trials of cancer therapy? *Statistics in medicine* 1998;17(5-7):667-77. doi: 10.1002/(sici)1097-0258(19980315/15)17:5/7<667::aid-sim813>3.0.co;2-6 [published Online First: 1998/04/29]
98. Bell ML, Fiero M, Horton NJ, et al. Handling missing data in RCTs; a review of the top medical journals. *BMC Medical Research Methodology* 2014;14(1):118. doi: 10.1186/1471-2288-14-118
99. FDA. FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency 2020 [updated 21/09/2020. Available from: https://www.fda.gov/media/136238/download accessed 22/10/2020 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. *Journal of Clinical Oncology* 2014;32(13):1302-08. doi: 10.1200/jco.2013.51.4489

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. *Journal of Clinical Oncology* 2014;32(13):1309-16. doi: 10.1200/jco.2013.51.4240
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. *New England Journal of Medicine* 2017;377(16):1525-36. doi: 10.1056/NEJMoa1615977
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. *JNCI: Journal of the National Cancer Institute* 1993;85(5):365-76. doi: 10.1093/jnci/85.5.365
105. FDA. Multiple Endpoints in Clinical Trials. Guidance for Industry 2017 [Available from: https://www.fda.gov/media/102657/download accessed 19/03 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. *Journal of Clinical Oncology* 2013;32(4):320-26. doi: 10.1200/JCO.2013.50.5669
107. Fairclough DL, Cella DF. Functional Assessment of Cancer Therapy (FACT-G): Non-response to individual questions. 1996;5(3):321-29. doi: 10.1007/bf00433916
108. Fayers P, Aaronson N, Bjordal K, et al. Bottomley A, on behalf of the EORTC Quality of Life Group: The EORTC QLQ–C30 Scoring Manual 3rd edition. 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 Medical Research Methodology* 2011;11(1):105. doi: 10.1186/1471-2288-11-105

110. Little RJ, Cohen ML, Dickersin K, et al. The design and conduct of clinical trials to limit missing data. 2012;31(28):3433-43. doi: 10.1002/sim.5519
111. McIsaac MA, Cook RJ, Poulin-Costello M. Incomplete Data in Randomized Dermatology Trials: Consequences and Statistical Methodology. *Dermatology* 2013;226(1):19-27. doi: 10.1159/000346247
112. Gamble C, Krishan A, Stocken D, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *JAMA* 2017;318(23):2337-43. doi: 10.1001/jama.2017.18556
113. Mazza G CC, Madeline Pe, Dorme L, Oliver K, Bottomley A, Sloan J, Dueck A, . 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.
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(10):e026080. doi: 10.1136/bmjopen-2018-026080
115. Kyte D, Draper H, Calvert M. Patient-Reported Outcome Alerts: Ethical and Logistical Considerations in Clinical Trials. *JAMA* 2013;310(12):1229-30. doi: 10.1001/jama.2013.277222
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(1):e0144658. doi: 10.1371/journal.pone.0144658
117. Sackett DL. Clinician-trialist rounds: 5. Cointervention bias – how to diagnose it in their trial and prevent it in yours. *Clinical Trials* 2011;8(4):440-42. doi: 10.1177/1740774511410995
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. FDA. CDER Patient-Focused Drug Development 2019 [Available from: https://www.fda.gov/drugs/development-approval-process-drugs/cder-patient-focused-drug-development accessed 23/04 2020.
120. 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. doi: 10.1111/j.1524-4733.2007.00267.x [published Online First: 2007/11/13]
121. Calvert M, Kyte D, Mercieca-Bebber R, et al. Supplement 3 - Summary Recommendations for Inclusion of PRO-Specific Information in the Trial Protocol, Supplementary Documents for Training for Staff: 2018; [Available from: https://jamanetwork.com/journals/jama/fullarticle/2671472.
122. Schandelmaier S, Conen K, von Elm E, et al. Planning and reporting of quality-of-life outcomes in cancer trials. *Annals of Oncology* 2015;26(9):1966-73. doi: 10.1093/annonc/mdv283
123. 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(18):3335-42. doi: 10.1002/cncr.29489
124. Mercieca-Bebber R, Rouette J, Calvert M, et al. Preliminary evidence on the uptake, use and benefits of the CONSORT-PRO extension. *Quality of Life Research* 2017;26(6):1427-37. doi: 10.1007/s11136-017-1508-6
125. 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. *Annals of Oncology* 2018;29(12):2288-95. doi: https://doi.org/10.1093/annonc/mdy449
126. 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. *Critical Reviews in Oncology/Hematology* 2020;146:102877. doi: https://doi.org/10.1016/j.critrevonc.2020.102877
127. 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. *Clinical Genitourinary Cancer* 2019;17(5):332-47.e2. doi: https://doi.org/10.1016/j.clgc.2019.07.007

128. Mathoulin-Pelissier S, Gourgou-Bourgade S, Bonnetain F, et al. Survival End Point Reporting in Randomized Cancer Clinical Trials: A Review of Major Journals. *Journal of Clinical Oncology* 2008;26(22):3721-26. doi: 10.1200/JCO.2007.14.1192
129. 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. doi: https://doi.org/10.1016/j.lungcan.2019.10.022
130. 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. *Journal of Clinical Oncology* 2008;26(13):2205-11. doi: 10.1200/JCO.2007.14.6795
131. Ioannidis JPA, Greenland S, Hlatky MA, et al. Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet* 2014;383(9912):166-75. doi: 10.1016/S0140-6736(13)62227-8
132. Glasziou P, Altman DG, Bossuyt P, et al. Reducing waste from incomplete or unusable reports of biomedical research. *The Lancet* 2014;383(9913):267-76. doi: 10.1016/S0140-6736(13)62228-X
133. European Commission. Clinical Trials Regulations (EU) No 536/2014 2019 [Available from: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/regulation5362014_qa_en.pdf accessed 19/03/2020.
134. 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 from: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/2017_01_26_summaries_of_ct_results_for_laypersons.pdf accessed 03/05/20 2020.
135. Rivera S, Stephens R, Mercieca-Bebber R, et al. “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

136. Turner G, Aiyegbusi OL, Price G, et al. Moving beyond project-specific patient and public involvement in research. *Journal of the Royal Society of Medicine* 2020;113(1):16-23. doi: 10.1177/0141076819890551

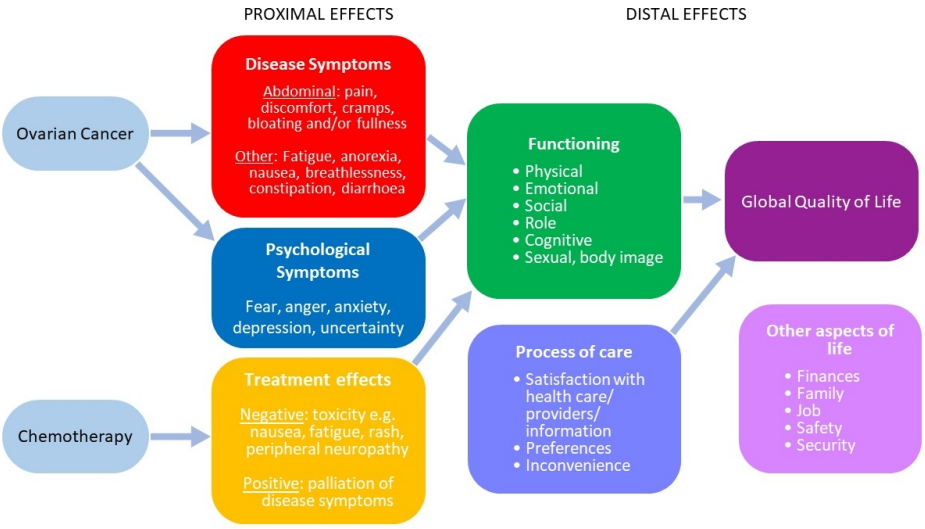
137. Snyder C, Smith K, Holzner B, et al. Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data. *Quality of Life Research* 2019;28(2):345-56. doi: 10.1007/s11136-018-2020-3

138. 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(10):1848-59. doi: 10.1002/cncr.30530

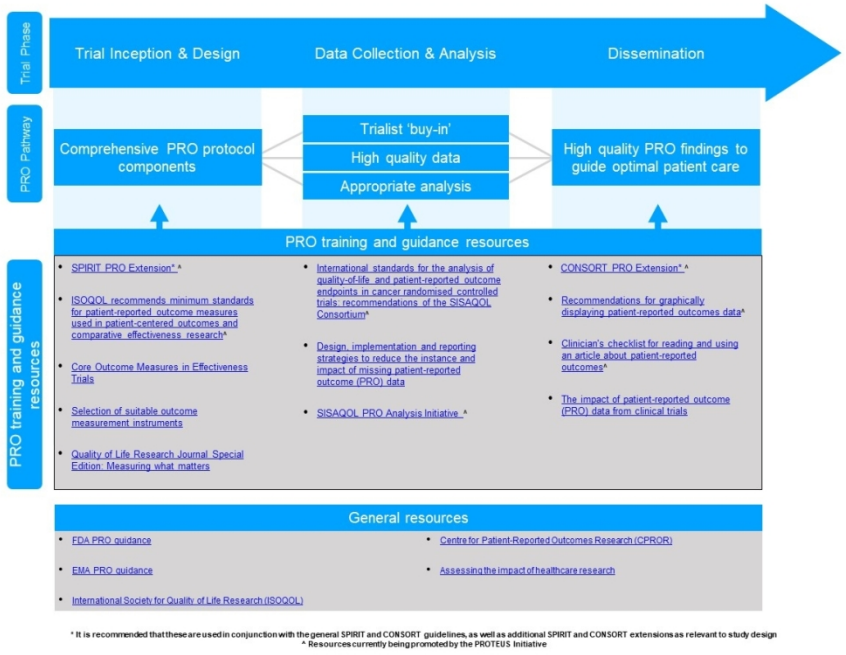
139. Brundage M, Blackford A, Tolbert E, et al. Presenting comparative study PRO results to clinicians and researchers: beyond the eye of the beholder. *Quality of Life Research* 2018;27(1):75-90. doi: 10.1007/s11136-017-1710-6

140. 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. *Quality of Life Research* 2017;26(6):1393-404. doi: 10.1007/s11136-016-1465-5

141. Frank L, Basch E, Selby JV, et al. The PCORI Perspective on Patient-Centered Outcomes Research. *JAMA* 2014;312(15):1513-14. doi: 10.1001/jama.2014.11100



338x190mm (96 x 96 DPI)



361x270mm (96 x 96 DPI)

REBUTTAL LETTER

Reviewer: 1

Recommendation:

Comments:
Calvert et al. report a document that supports the implementation of the SPIRIT-PRO recommendations. The authors provided examples from different trials for each of the checklist items, and a PRO protocol template. I find this document extremely useful as it gives tools for the design of clinical trials that include PROs as primary or secondary aims. Yet, given that the document is an extension of the SPIRIT-PRO, there is a significant amount of overlap with that document. Although the overlap is expected, it does reduce the originality of the publication.
Response: As with other E&E papers it is only natural that there will be overlap with the original publication. We need to ensure consistency with the original checklist. The JAMA paper describes the rationale and methods for the SPIRIT-PRO checklist, and includes the 16 final SPIRIT-PRO checklist items. This E&E aims to support the implementation of the guidance. As such, we have added examples (not included in the original manuscript) and further evidence to support each of the items co-authored by leading international experts including: regulators, PRO methodologists, trialists, patient partners. In addition, the manuscript includes a brand new protocol template to support implementation.

Furthermore, the document is focused on the implementation of specific guidance around the conduct of clinical trials and may not be of interest to the majority of the readership of the BMJ.
Response: A recent review of the Australian New Zealand Clinical Trials Registry suggested that 455 clinical trials include a PRO endpoint as the primary or secondary outcome, and similarly 26,337 trials registered on ClinicalTrials.gov between 2007-2013 included a PRO endpoint . Both reviews found that the use of PRO endpoints is increasing over time. Therefore we believe this manuscript would be of relevance to a large proportion of BMJO readers. Although the main focus of our manuscript is on clinical trials, we note that the methods proposed should be considered for broader PRO research. The guidance applies to all clinical disciplines and as such we believe this will be of major interest to BMJ Open readers.

The following are suggestions to improve the manuscript

There are several limitations about the SPIRIT-PRO extension manuscript, including the date of the last search for the systematic review.
Response: The systematic review was undertaken as part of the original development of the SPIRIT-PRO guidance (published in 2018), to identify items to inform the Delphi and consensus exercise, and finalise the content of SPIRIT-PRO guidelines. As such, it does not make sense to update the review. The SPIRIT-PRO Extension is the latest international consensus based guideline, and is endorsed by the EQUATOR network. We only mention the systematic search by way of background and to highlight the rigorous methodology use to develop the SPIRIT-PRO guidelines. The current manuscript does not address the development of SPIRIT-PRO, but rather provides support for its implementation.
It may be important to mention these limitations in this manuscript as well.
Response: As reasoned above we do not believe this is a limitation.

The methods used to develop this manuscript are short and do not give a clear overview of how the examples were found, chosen, and how those decisions were made.
Response: Thank you we have restructured the methods section. Protocol selection is described as follows (page 20):
Protocol excerpts for each checklist item were obtained from public websites, journals, trial investigators, and industry sponsors. In addition, protocols which adhered well to SPIRIT-PRO guidance were identified through a review of trial protocols (ref EPIC) and via international trials groups known to the coauthors. For those protocols unavailable in the public domain permission was sought to publish.
Real-world examples, quoted verbatim, were selected to reflect how key elements could be appropriately described in a trial protocol. Some examples illustrate a specific component of a checklist item, while others encompass all key recommendations for an item. Empirical data and references to support each SPIRIT-PRO item are provided. Reference numbers cited in the original quoted text are denoted by [Reference] to distinguish them from references cited in this E&E paper.

Consider expanding in the introduction section the need for this manuscript. Why is SPIRIT-PRO not enough? Perhaps some examples of where SPIRIT-PRO has been used and been unclear may help clarify some of the subjectivity of the SPIRIT-PRO extension recommendations.
Response: The introduction already includes justification for the E&E paper. We have noted that the protocol template is a new addition (page 5). The original SPIRIT-PRO manuscript describes the background, rationale and methodology for developing the SPIRIT-POR guidance, as well as the final 16 items. This E&E manuscript addresses how those items may be implemented in clinical trial protocols, and provides supportive resources to assist researchers and investigators to do this in a rigorous, complete and high-quality manner.
This E&E paper aims to promote understanding of the guidelines, provide examples from a range of different trials and facilitate uptake of the recommended checklist items. In addition, we describe the development of a new PRO protocol template for use in protocol development.
We have also added additional evidence and supporting reference (page 5)-
A recent review of cancer portfolio trials illustrates this point; PRO protocol content was frequently inadequate and PRO data from trials including 49 568 participants remaining unpublished.

A key aspect of the use of PRO is participant burden. The authors mentioned this potential burden and the importance of addressing it at the protocol stage. Yet, there is also the need to assess PRO related burden during the trial conduct. Could the authors recommend ways to achieve this?
Response: Thank you for noting this point. We believe that patient and public involvement is key here and we have included further text written by our patient partners on page 22. Qualitative work alongside the trial could help assess participant burden but is beyond the scope of the SPIRIT-PRO protocol guidance. 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.

Additional Questions:
The BMJ uses compulsory open peer review. Your name and institution will be included with your comments when they are sent to the authors. If the manuscript is accepted, your review, name and institution will be published alongside the article.

If this manuscript is rejected from The BMJ, it may be transferred to another BMJ journal along with your reviewer comments. If the article is selected for publication in another BMJ journal, depending on the editorial policy of the journal your review may also be published. You will be contacted for your permission before this happens.

For more information, please see our peer review terms and conditions.

Please confirm that you understand and consent to the above terms and conditions.: I consent to the publication of this review

Please enter your name: Juan P. Brito

Job Title: Associate Professor of Medicine

Institution: Mayo clinic

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?:

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests (please see BMJ policy) please declare them here:

SPIRIT-PRO

PROtocol Template

About this PROtocol Template

Patient-reported outcome (PRO) data from clinical trials can provide valuable evidence to inform shared decision making, labeling claims, clinical guidelines, and health policy; however, the PRO content of clinical trial protocols is often suboptimal.

To address this issue an international, consensus-based, PRO-specific protocol guidance (the SPIRIT-PRO Extension) was developed and published in 2018:

> Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for Inclusion of Patient-Reported Outcomes in Clinical Trial Protocols: The SPIRIT-PRO Extension. *JAMA.* 2018;319(5):483–494. doi:https://doi.org/10.1001/jama.2017.21903

This PROtocol template aims to promote implementation and use of the SPIRIT-PRO Extension for trials where PROs are a primary or key secondary outcome; however content is also recommended for use where PROs are exploratory outcomes. This PROtocol template was designed as a stand-alone document which may be incorporated into any clinical trial protocol. When using this addendum please cite as:

Journal to add E&E paper reference please

Nothing in this template should be construed to represent or warrant that persons using this template have complied with all applicable laws and regulations. All individuals and organizations using this template have the responsibility for complying with the applicable laws and regulations or regulatory requirements for the relevant jurisdiction.

We recommend integration of key SPIRIT-PRO information within relevant sections of the protocol (e.g., rationale, schedule of assessments, objectives, endpoints, and statistical analysis). Additional notes from the SPIRIT-PRO group and the industry advisory group (IAG) have also been provided where necessary.

In addition we recommend a separate PRO specific section of the protocol which provides further background information, justification for selection of measures, details on psychometric properties of measures and data collection procedures. The protocol template aims to serve as a guide and sections can be moved to best fit with existing trial templates, however we recommend the use of the SPIRIT-PRO checklist to ensure all content has been covered (Page 18). Efforts should be made by protocol writers to avoid unnecessary repetition of content.

Protocol writers can confirm that they have successfully adhered to the SPIRIT-PRO guideline using the checklist available here:
https://jamanetwork.com/journals/jama/article-abstract/2671472

Protocol writers are encouraged to read and consider other relevant resources which are beyond the scope of the SPIRIT-PRO Extension as detailed below:

References and useful resources:

Evidence-based recommendations for the minimum content of a clinical trial protocol

Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, Hróbjartsson A, Mann H, Dickersin K, Berlin J, Doré C, Parulekar W, Summerskill W, Groves T, Schulz K, Sox H, Rockhold FW, Rennie D, Moher D. SPIRIT 2013 Statement: Defining standard protocol items for clinical trials. *Ann Intern Med* 2013;158:200-207.

Chan A-W, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin J, Dickersin K, Hróbjartsson A, Schulz KF, Parulekar WR, Krleža-Jerić K, Laupacis A, Moher D. SPIRIT 2013 Explanation and Elaboration: Guidance for protocols of clinical trials. *BMJ* 2013;346:e7586.

Enhancing the Quality and Transparency Of health Research (EQUATOR) Network: https://www.equator-network.org/

Estimands

FDA-ASCO Public Workshop: 2019 Clinical Outcome Assessments in Cancer Clinical Trials Fourth Annual Workshop. King-Kallimanis, B. Systematically defining research objectives and framing questions using the estimand framework. https://www.fda.gov/drugs/news-events-human-drugs/fda-asco-public-workshop-2019-clinical-outcome-assessments-cancer-clinical-trials-fourth-annual [accessed 19/12/19]

E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). October 2017. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM582738.pdf

Akacha M, Bretz F, Ruberg S. Estimands in clinical trials - broadening the perspective. Stat Med. 2017;36(1):5-19.

Akacha M, Bretz F, Ohlssen D, Rosenkranz G, Schmidli H. Estimands and Their Role in Clinical Trials. Statistics in Biopharmaceutical Research. 2017;9(3):268-71.

Permutt T. A taxonomy of estimands for regulatory clinical trials with discontinuations. Statistics in Medicine. 2016;35(17):2865-75.

Bell ML, Floden L, Rabe BA, Hudgens S, Dhillon HM, Bray VJ, Vardy JL. Analytical approaches and estimands to take account of missing patient-reported data in longitudinal studies. *Patient Relat Outcome Meas*. 2019;10:129–140. Published 2019 Apr 16. doi:10.2147/PROM.S178963

Patient Reported adverse events

If the trial is assessing patient reported adverse event and symptom monitoring consider how this links to adverse reporting in the protocol.

Kluetz PG, Chingos DT, Basch EM, Mitchell SA. 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. doi: 10.14694/EDBK_159514.

Tolerability

If PROs are being used to assess tolerability this should be described.

Kluetz PG, Kanapuru B, Lemery S, Johnson LL, Fiero MH, Arscott K, Barbachano Y, Basch E, Campbell M, Cappelleri JC, Cella D, Cleeland C, Coens C, Daniels S, Denlinger CS, Fairclough DL, Hillard JR, Minasian L, Mitchell SA, O'Connor D, Patel S, Rubin EH, Ryden A, Soltys K, Sridhara R, Thanarajasingam G, Velikova G, Coons SJ. Informing the Tolerability of Cancer Treatments Using Patient-Reported Outcome Measures: Summary of an FDA and Critical Path Institute Workshop. Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2018;21(6):742-7.

Analysis

The Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data (SISAQOL) Consortium aims to develop recommendations to standardize the analysis of PRO data in cancer randomized controlled trials. https://event.eortc.org/sisaqol/

Coens C, Pe M, Dueck AC, Sloan J, Basch E, Calvert M, Campbell A, Cleeland C, Cocks K, Collette L, Devlin N, Dorme L, Flechtner HH, Gotay C, Griebsch I, Groenvold M, King M, Kluetz PG, Koller M, Malone DC, Martinelli F, Mitchell SA, Musoro J, O'Connor D, Oliver K, Piault-Louis E, Piccart M, Quinten C, Reijneveld JC, Schürmann C, Smith AW, Soltys KM, Taphoorn M, Velikova G, Bottomley A. International standards for the analysis of quality-of-life and patient-reported outcome endpoints in cancer randomised controlled trials: recommendations of the SISAQOL Consortium. The Lancet Oncology 2020;21(2):e83-e96.

Bottomley A, Pe M, Sloan J, Basch E, Bonnetain F, Calvert M, Campbell A, Cleeland C, Cocks K, Collette L, Dueck AC, Devlin N, Flechtner HH, Gotay C, Greimel E, Griebsch I, Groenvold M, Hamel JF, King M, Kluetz PG, Koller M, Malone DC, Martinelli F, Mitchell SA, Moinpour CM, Musoro J, O'Connor D, Oliver K, Piault-Louis E, Piccart M, Pimentel FL, Quinten C, Reijneveld JC, Schürmann C, Smith AW, Soltys KM, Sridhara R, Taphoorn M, Velikova G, Coens C. Moving forward toward standardizing analysis of quality of life data in randomized cancer clinical trials. Clinical trials (London, England). 2018;15(6):624-30.

Pe M, Dorme L, Coens C, Basch E, Calvert M, Campbell A, Cleeland C, Cocks K, Collette L, Dirven L, Dueck AC, Devlin N, Flechtner HH, Gotay C, Griebsch I, Groenvold M, King M, Koller M, Malone DC, Martinelli F, Mitchell SA, Musoro J, O'Connor D, Oliver K, Piault-Louis E, Piccart M, Pimentel FL, Quinten C, Reijneveld JC, Schürmann C, Sloan J, Velikova G, Bottomley. Statistical analysis of patient-reported outcome data in randomised controlled trials of locally advanced and metastatic breast cancer: a systematic review. The Lancet Oncology. 2018;19(9):e459-e69.

Bottomley A, Pe M, Sloan J, Basch E, Bonnetain F, Calvert M, Campbell A, Cleeland C, Cocks K, Collette L, Dueck AC, Devlin N, Flechtner HH, Gotay C, Greimel E, Griebsch I, Groenvold M, Hamel JF, King M, Kluetz PG, Koller M, Malone DC, Martinelli F, Mitchell SA, Moinpour CM, Musoro J, O'Connor D, Oliver K, Piault-Louis E, Piccart M, Pimentel FL, Quinten C, Reijneveld JC, Schürmann C, Smith AW, Soltys KM, Taphoorn M, Velikova G,

Coens C. Analysing data from patient-reported outcome and quality of life endpoints for cancer clinical trials: a start in setting international standards. The Lancet Oncology. 2016;17(11):e510-e4.

Other resources

Cappelleri JC, Zou KH, Bushmakin AG, Alvir JMJ, Alemayehu D, Symonds T. Patient-Reported Outcomes: Measurement, Implementation and Interpretation. Boca Raton, Florida: Chapman & Hall/CRC; 2013.

de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine. Cambridge, UK: Cambridge University Press; 2011.

Fairclough DL. Design and Analysis of Quality of Life Studies in Clinical Trials. 2nd ed. Boca Raton, Florida: Chapman & Hall/CRC; 2010.

Fayers FM, Machin D. Quality of Life: The Assessment, Analysis and Interpretation of Patient-reported Outcomes. 3rd ed. Chichester, England: John Wiley & Sons Ltd.; 2016.

Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A Practical Guide to Their Development and Use. 5th ed. New York, NY: Oxford University Press; 2015.

Missing Data

Mercieca-Bebber R, Palmer MJ, Brundage M, Calvert M, Stockler MR, King MT. 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(6):e010938.

Little RJ, D'Agostino R, Cohen ML, et al. The Prevention and Treatment of Missing Data in Clinical Trials. New England Journal of Medicine 2012;367(14):1355-60.

Little RJ, Cohen ML, Dickersin K, et al. The design and conduct of clinical trials to limit missing data. Stat Med 2012;31(28):3433-43. doi: 10.1002/sim.5519

Patient Public Involvement (PPI) and Patient Experience

Kluetz PG, O'Connor DJ, Soltys K. Incorporating the patient experience into regulatory decision making in the USA, Europe, and Canada. The Lancet Oncology. 2018;19(5):e267-e74.

Table of Contents

Title page

PRO content author:

Affiliation(s), telephone:

Specify the individual(s) responsible for the PRO content of the trial protocol. (SPIRIT-5a-PRO Elaboration)

Explanation: Providing information (e.g. name, affiliation, contact details) on who wrote the PRO-specific aspects of the trial protocol promotes transparency and accountability and identifies the appropriate point of contact for resolution of any PRO specific queries. When patients have actively contributed to this process, this should be documented as per recent guidance for the reporting of patient and public involvement.[7]

Additional notes: *The PRO author should be part of the protocol writing committee. The trial study coordinator should ensure that PROs are harmonized with all the other clinical endpoints.*

1. Protocol Summary

The protocol synopsis is a short (1 to 2 pages) summary of the key points of the protocol (including PRO-specific information). This section of the protocol should be completed after the main text to ensure consistency with the main text.

1.3. Schedule of Activities (SoA)

Include a schedule of PRO assessments, providing a rationale for the time points, and justifying if the initial assessment is not pre-randomization. Specify time windows, whether PRO collection is prior to clinical assessments, and, if using multiple questionnaires, whether order of administration will be standardized (SPIRIT-13-PRO Extension).
Additional notes: *The PRO schedule of assessments should be summarised in a table alongside all other trial assessments to provide trial staff with a single point of reference. Ideally PRO assessments should appear in the table in the order which assessments will take place. If more than one PRO measure is being used, each may be specified on a separate line if they are administered at different time points.*

Procedure	Screening (up to X days before Day 1)	Intervention Period [Days or Weeks, etc.]									Follow-up (X days after last dose)	Notes
		–1 (time window)	1 (time window)	2 (time window)	3 (time window)	4 (time window)	5 (time window)	6 (time window)	7 (time window)	8 (time window) E/D *		E/D = Early Discontinuation
List PRO assessments consistent with desired order of completion in study visits (more than one line may be necessary). Further details should be provided in Section 4.2. (SPIRIT 13 PRO Extension)												

2. Introduction

2.1. Study Rationale

2.1.1. Summary of PRO-specific rationale

Describe the PRO-specific research question and rationale for PRO assessment and summarize PRO findings in relevant studies. (SPIRIT-6a-PRO Extension)
Additional note: *As with other clinical information, the PRO-specific rationale should be summarized here and detailed information including summary of PRO findings in relevant studies should be provided in Section 4.2. Scientific Rationale for Study Design or in the PRO-specific section 8.1.*

Explanation: Inclusion of PROs in a trial requires careful consideration and planning. A clearly defined question helps with selection of measures and specification of hypotheses and analyses. When the PRO is a secondary or exploratory outcome, a brief rationale may be adequate.[6]

3. Objectives and Endpoints

State specific PRO objectives (including relevant PRO concepts/domains). (SPIRIT-7-PRO Extension)
Additional notes: *The level of detail on the PRO endpoints should also be similar to the other clinical endpoints to ensure harmonization. State whether specific PRO domains will be used for confirmatory or descriptive or exploratory purposes (SISAQOL recommendation statement (RS)1).[4] If primary or secondary objectives are confirmatory objectives (draw conclusions about treatment efficacy), there is a need to specify whether the objective is to show superiority, equivalence or non-inferiority (SISAQOL RS2).[4]*

Explanation: Pre-specification of objectives encourages identification of key PRO domains and time points, reducing the risk of multiple statistical testing and selective reporting of PROs based on statistically significant results.[8]

Specify the PRO *endpoint: the* concepts/domains used to evaluate the intervention (e.g., overall health-related quality of life, specific domain, specific symptom) and, for each one, the analysis metric (e.g., change from baseline, final value, time to event) and the principal time point or period of interest. (SPIRIT-12-PRO Extension)
Additional note: *Utility measures (e.g. EQ-5D) that inform cost-utility evaluations should also be included here.*

Explanation: The PRO concepts/domains and time points for assessment should closely align with the trial objectives and hypotheses. Because of the risk of multiple statistical testing, the domain(s) and principal time point(s) for analyses should be specified *a priori*.[8 9]

Objectives (including PRO objectives)	[Endpoints]
Primary	
1.	
Secondary	
2.	
Tertiary/Exploratory	
3.	

4. Study Design

4.1. Scientific Rationale for Study Design (PRO-specific) [Note this section could also be combined into the PRO-specific section 8.1]

Describe the PRO-specific research question and rationale for PRO assessment and summarize PRO findings in relevant studies. (SPIRIT-6a-PRO Extension)
Additional notes: *Detailed information should be provided here. Information not provided in Section 2.1.1. Summary of PRO-specific rationale should be presented here. Indicate how PRO evaluation aligns with the overall trial design, so it addresses the research objectives.*

Explanation: Inclusion of PROs in a trial requires careful consideration and planning. A clearly defined question helps with selection of measures and specification of hypotheses and analyses. When the PRO is a secondary outcome, a brief rationale may be adequate.[6]

5. Study Population

Specify any PRO-specific eligibility criteria (e.g., language/reading requirements or pre-randomization[1] completion of PRO). (SPIRIT-10-PRO Extension)
If PROs will not be collected from the entire study sample, provide a rationale and describe the method for obtaining the PRO subsample. (SPIRIT-10-PRO Extension)
Additional notes: *Efforts should be made to provide translated PRO versions where needed to promote inclusivity in PRO completion. Specify reasons if translated versions are not provided and certain populations are excluded. PRO-specific eligibility criteria should be included alongside other eligibility criteria to ensure research personnel have a single point of information.*

Explanation: Any PRO-specific eligibility criteria should be considered at the design stage of the trial and clearly specified in the protocol. In large trials, sufficient power may be achieved by collecting PROs from a representative subset of participants, while in some trials it may not be possible to collect PROs in the entire population (e.g., because of non-availability of validated questionnaires in all languages)[10]; in such instances, the rationale for the sampling method should be described.

6. Study Intervention

6.1. PRO capture method involving a medical device

Notes: *Some ePRO systems may meet the requirement for a medical device (for example by providing actionable alerts). If the ePRO fulfils the definition of a medical device, then the protocol should refer to the relevant regulatory directive for that jurisdiction.*

7. Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal from the Study

7.1. Discontinuation of Study Intervention

Describe the process of PRO assessment for participants who discontinue or deviate from the assigned intervention protocol. (SPIRIT-18b (ii)-PRO Elaboration)
Additional note: *Participants who withdraw from the study should be provided pre-paid packaging to return PRO assessments and ePRO devices if applicable.*

Explanation: A clear plan for collection of PROs for trial participants who withdraw early from a study or who discontinue the intervention helps minimize bias,[11] ensures that staff collect all required PRO data in a standardized and timely way, and may assist ethical appraisal of the study.[6]

8. Study Assessments and Procedures

8.1. Efficacy Assessments

8.1.1. Patient Reported Outcomes

Specify the PRO Hypotheses (SPIRIT-7-PRO Extension)
Additional note: *Also specify whether the objective is to support superiority, equivalence or non-inferiority.*

Explanation: Pre-specification of objectives and hypotheses encourages identification of key PRO domains and time points, reducing the risk of multiple statistical testing and selective reporting of PROs based on statistically significant results.[8]

Justify the PRO instrument to be used and describe domains, number of items, recall period, instrument scaling and scoring (e.g., 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. (SPIRIT-18a (i)-PRO Extension).
Additional note: *Interpretation guidelines should include reference to pre-defined clinical relevance thresholds, preferably specific to the PRO population.*

Explanation: The selection of PROs to be used in a clinical trial requires careful consideration. Ideally, the measure should be validated in the target population.[12] Consideration should be given to the number of questionnaires to be used, acceptability of the questions, and the likely patient burden (e.g., time taken for completion, cognitive burden, emotional burden). Justification for the measures selected will help trial personnel understand why specific measures are being used.[13] Questionnaires should be used in accordance with any existing user manuals to promote data quality and ensure standardized scoring, and any deviations should be described.[6]

Include a schedule of PRO assessments, providing a rationale for the time points, and justifying if the initial assessment is not pre-randomization. Specify time windows, whether PRO collection is prior to clinical assessments, and, if using multiple questionnaires, whether order of administration will be standardized. (SPIRIT-13-PRO Extension)

Additional notes: *State the mode of administration (e.g., paper and/or electronic capture). Baseline PROs should be completed prior to randomization allocation. The windowing for the first on treatment assessment should ensure the recall period does not overlap with day 0 of the trial (e.g., if recall is 7 days the minimum window should be trial day 7). The order of PRO administration should be the same for the entire duration of the study. Specify when patients are still expected to complete a PRO assessment (e.g., whether PROs will be collected after disease progression or not). This will allow calculation of completion rates (SISAQOL recommendation statement (RS) 19-20).[4] This information can be summarised in section 1.3 with further detail provided in this PRO-specific section.*

Explanation: Provision of an easy-to-follow schedule will assist staff and may help reduce missing data.[1] Collecting PRO data prior to randomization helps ensure an unbiased baseline assessment, and if specified as an eligibility criterion, ensures data completeness. This is important because baseline PRO data are often used as a covariate in analyses and are essential to calculating change from baseline. Completion of PROs prior to clinical assessments (as these may influence patient responses) and standardization of the order of questionnaire administration are advised to help reduce measurement error.[14] Allowable time windows for each scheduled PRO assessment should be specified to ensure that PRO data collection captures the effect of the clinical event(s) of interest.[6]

Include a data collection plan outlining the permitted mode(s) of administration (e.g., paper, telephone, electronic, other) and setting (e.g., clinic, home, other). (SPIRIT-18a (ii)-PRO Extension)

Additional notes: *Specify whether participants would be completing assessments at the site. If so, they should be completed prior to any other assessments, procedures, or discussions with their care team. Participants should also be given time and a quiet space to complete the assessments. If electronic PRO collection tools are unavailable, damaged or lost, describe procedure to follow for the collection of PROs according to the SoA.*

Explanation: It is important that both research personnel and trial participants understand how, when, and where PRO data will be collected in the study. Increasingly, electronic PRO assessment is undertaken in trials, so evidence of equivalence between different modes of administration should be considered.[15] If electronic PRO measures contain only minor modifications with respect to the paper-based versions, usability testing and cognitive debriefing may provide sufficient evidence of equivalence.[15 16] The setting for PRO data collection should be described and standardized across trial intervention groups and sites.[6]

Specify PRO data collection and management strategies for minimizing avoidable missing data. (SPIRIT-18b (i)-PRO Extension)
Additional notes: *Signpost to statistical analysis plan (SAP) as required. Collect reasons for missed PRO assessments during the trial. These could be documented in the case report forms (CRF). SISAQOL, CONSORT-PRO Extension and PROTEUS may be consulted for further guidance on missing data.*[2 4 5]

Explanation: Missing data are a particular problem for PROs because participants with the poorest outcomes in a trial often are those who do not complete planned PRO assessments, and data cannot be obtained retrospectively beyond the time frame of interest or from medical records. This is a potentially significant source of bias and may reduce trial power.[17] It is important to note that not all missing PRO data are avoidable: participants have the right to decide not to complete questionnaires. Common reasons for avoidable missing PRO data are administrative errors, lack of explanation of the importance of PRO data, and overly burdensome questionnaires. Addressing these in the protocol should help minimize avoidable missing data. Examples of protocol content include ensuring that PRO end points and hypotheses are clearly defined and scientifically compelling, providing a rationale for PRO assessment, clearly specifying the PRO assessment time points, defining acceptable PRO assessment time windows, aligning PRO assessment time points to clinic visits (if clinically informative), minimizing participant burden, and specifying the importance of complete PRO data.[1]

Specify whether more than one language version will be used and state whether translated versions have been developed using currently recommended methods. (SPIRIT-18a (iii)-PRO Extension)

Explanation: Multinational trials, or national trials involving participants with different languages, require measures that have been translated and culturally adapted where needed using appropriate methodology.[18 19] This may influence the selection of measure to be used because inclusion of a wide range of participants can help ensure the generalizability of trial results. Plans to use translated versions should be specified in the protocol, citing references when available.[6]

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. (SPIRIT-18a (iv)-PRO Extension)
Additional note: *If a proxy reporting is used, this should ideally be the same person throughout the trial.*

Explanation: In some contexts, such as trials involving young children or cognitively impaired participants, it may be necessary for someone other than a trial participant to respond on that participant's behalf. Note that proxy is different from someone assisting a person to respond to questionnaires (e.g., a nurse reads questions to a patient and writes

down their actual answers). Clear justification and specification of proxy reporting in the protocol allows external reviewers to assess potential bias and facilitates trial reporting in accordance with CONSORT-PRO.[2]

8.2. Safety Assessments

8.2.1 Monitoring of PRO data

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 standardized way. Describe how this process will be explained to participants; e.g., in the participant information sheet and consent form. (SPIRIT-22-PRO Extension)
Additional note: *If PROs are used in place of or as an add-on to standard solicited AE recording (at visits/phone calls the investigators asks if there have been any problems since last visit) they should be recognised as part of the safety monitoring procedures of the trial.*

Explanation: Evidence suggests that monitoring and management of PRO alerts (psychological distress or physical symptoms evident from PRO responses that may require an immediate response) vary across and within trials.[13 18 20] To protect the interests of trial participants and minimize potential bias, it is important to specify plans for monitoring.[21] If monitoring is not planned (e.g., in a low-risk study in which alerts are not anticipated), this should also be briefly stated in the protocol, the participant information sheet, and the consent form. Alternative support mechanisms for participants should be outlined.[6]

9. Statistical considerations

When a PRO is the primary end point, 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 end point, then discuss the power of the principal PRO analyses. (SPIRIT-14-PRO Elaboration)
Additional note: *Enough information on secondary, tertiary and even exploratory PRO endpoints should be provided to justify their inclusion.*

Explanation: In studies in which PROs are the primary outcome or end point, the target sample size will generally be based on an a priori sample size calculation for that end point.[9] Ideally, the criteria for clinical significance (e.g., minimal important difference, responder definition) should be specified when known.[22 23] If PROs are a secondary end point, researchers should specify whether the sample size provides sufficient power to test the principal PRO hypotheses.[9]

9.1. Statistical analyses

State PRO analysis methods, including any plans for addressing multiplicity/type I (α) error. (SPIRIT-20a-PRO Elaboration)
Additional notes: *Also state if there are no plans to address multiplicity/type I (α) error. Pre-define clinical relevance thresholds for the domains (SISAQOL recommendation statement (RS) 3-5).[4] Methods should align with specified endpoints and reflect superiority, equivalence or non-inferiority objectives (SISAQOL RS 2).[4] A power analysis can be conducted to assess whether a clinical relevant difference as specified in the objective can be detected reliably with the given sample size for the PRO population (SISAQOL RS 2).[4] If sensitivity analysis is planned, this should be pre-specified (SISAQOL RS 32).[4]*

Explanation: Many questionnaires, such as health-related quality-of-life measures, are multidimensional and therefore may yield several summary scores (e.g., multiple domains and an overall score). Furthermore, PROs are usually assessed at multiple time points. Statistical analysis of all domains and time points implies multiple hypothesis testing, which inflates the probability of false-positive results (type I error).[9] This can be contained by pre-specifying the key PRO domain(s) or overall score of interest and the principal time point(s). Any plans to address multiplicity, such as stepwise or sequential analyses, whereby multiple end points are tested in a defined sequence that contains the overall type I error to the desired level, or conventional non-hierarchical methods (e.g., Bonferroni correction), should be specified a priori.[8] The protocol should either fully address these issues or provide a summary with reference to where full details can be found (e.g., in the statistical analysis plan).[6]

9.1.1. Missing PRO data

State how missing data will be described and outline the methods for handling missing items or entire assessments (e.g., approach to imputation and sensitivity analyses). (SPIRIT-20c-PRO Elaboration)
Additional notes: *State whether PRO data is missing at the item level or at entire PRO assessment level and whether intermittent or lost to follow-up. In addition, state how missing PRO data is recorded and categorised. SISAQOL recommendations may be consulted for further guidance on missing data.[4]*

Explanation: There are 2 levels of missing PRO data: (1) patient completion of some but not all items within an instrument and (2) absence of the entire PRO assessment. Whether and how missing items should be imputed is usually specified in an instrument's scoring algorithm. When entire PRO assessments are missed, analysis requires assumptions about why those data were missing (i.e., the missing data mechanism). There are a range of statistical approaches, each with specific assumptions. Inappropriate method selection may lead to potentially biased and misleading results.[1 11] The protocol should acknowledge and summarize these issues, with full details provided in the statistical analysis plan.[6]

9.1.2. Other Analyse(s)

Note: *If psychometric analyses of the PRO measure(s) or any additional research questions are planned with the PRO data collected, this should be acknowledged in the protocol, with reference to statistical analysis plan (SAP), Calvert et al 2018 (JAMA Supplement 3).*[6]

10. Supporting Documentation and Operational Considerations

10.1. Appendix 1: Regulatory, Ethical, and Study Oversight Considerations

10.1.1. Publication Policy

Notes: *It is important to publish PRO data. There is evidence that this is not done enough or well. Publication of PRO data should be done according to CONSORT and CONSORT-PRO*

Note: *Whilst not a formal recommendation in the SPIRIT-PRO guidance, it is regarded as good practice to include, depending on the mode of administration, paper copies of PRO or screenshots of ePRO instruments.*
All copyrights and version information should be clearly showing in the appended PRO instruments with due acknowledgment for PROs requiring permission for use.

10.2. Appendix Copies of PROs

Glossary

Concept: "The specific measurement goal (i.e., 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."[24]
Domain: "A sub concept represented by a score of an instrument that measures a larger concept comprised of multiple domains. For example, psychological function is a larger concept with multiple domains (emotional and cognitive function) that are measured by relevant items."[24]
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 combined[25] (e.g. change from baseline at 6 weeks in mean fatigue score).[26]
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"[27]
Important or key secondary PROs/end points: Some PRO measures (particularly health-related quality-of-life measures) are multidimensional, producing several domain-specific outcome scales; e.g., 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 (CONSORT PRO Extension 2b),[27] 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 (e.g. using the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE™))[28]
Instrument: "A means to capture data (e.g., 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."[24]
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.[29]
Item: "an individual question, statement, or task (and its standardized response options) that is evaluated by the patient to address a particular concept"[24]
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 health care professional".[30]
Outcome*: the variable to be measured. It is the measurable characteristic that is influenced or affected by an individuals' baseline state or an intervention as in a clinical trial or other exposure[25] (e.g. a fatigue score).
Patient-reported outcome (PRO): 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.[24 27] PROs are assessed by self-reported questionnaires, referred to as PRO measures (PROMs) or instruments.[25]
Primary outcome: the most important outcome in a trial, pre-specified 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"[24]
Secondary outcomes: outcomes pre-specified in the protocol to assess additional effects of the intervention; some PROs may be identified as important or key secondary outcomes
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials.[31 32]
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 time point whereby the result would still be deemed to be clinically relevant.[33]

* 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 (i.e. the outcome) by fully specifying four components: measurement variable (e.g. fatigue "in the past week" as measured by the QLQ-C30), analysis metric (e.g., change in fatigue from baseline, final fatigue value, time to clinically important increase in fatigue (and "event"), method of aggregation (e.g., median fatigue, proportion of patients with severe fatigue, proportion of patients with clinically important change in fatigue), and time point. Note that using theses definitions, several endpoints can be defined from the same outcome source data, revealing the distinction and relationship between "outcome" and "endpoint" for PROs.

List of Abbreviations

CONSORT – Consolidated Standards of Reporting Trials

CRF – case report form

ePRO – electronic patient-reported outcome

PPI – Patient Public Involvement

PRO – patient-reported outcome

PRO-CTCAE – Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events

PROTEUS – Patient-Reported Outcomes Tools: Engaging Users & Stakeholders

RS – recommendation statement

SAP – statistical analysis plan

SISAQOL – Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data

SoA – Schedule of Activities

SPIRIT – Standard Protocol Items: Recommendations for Interventional Trials

References

1. 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(6):e010938. doi: 10.1136/bmjopen-2015-010938 [published Online First: 2016/06/18]
2. Calvert M, Blazeby J, Altman DG, et al. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *Jama* 2013;309(8):814-22. doi: 10.1001/jama.2013.879 [published Online First: 2013/02/28]
3. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Annals of internal medicine* 2013;158(3):200-7. doi: 10.7326/0003-4819-158-3-201302050-00583 [published Online First: 2013/01/09]
4. 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. *The Lancet Oncology* 2020;21(2):e83-e96. doi: 10.1016/S1470-2045(19)30790-9
5. Snyder C, Smith K, Holzner B, et al. Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data. *Quality of Life Research* 2019;28(2):345-56. doi: 10.1007/s11136-018-2020-3
6. Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for Inclusion of Patient-Reported Outcomes in Clinical Trial Protocols: The SPIRIT-PRO Extension. *Jama* 2018;319(5):483-94. doi: 10.1001/jama.2017.21903
7. 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. doi: 10.1136/bmj.j3453 [published Online First: 2017/08/05]
8. FDA. Guidance for industry. Patient-reported outcome measures: use in medicinal product development to support labeling claims. Silver Spring, MD: US Department of Health and Human Services Food and Drug Administration, 2009.
9. 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. doi: 10.1111/j.1524-4733.2007.00273.x [published Online First: 2007/11/13]
10. Kyte D, Retzer A, Ahmed K, et al. Systematic Evaluation of Patient-Reported Outcome Protocol Content and Reporting in Cancer Trials. *JNCI: Journal of the National Cancer Institute* 2019;111(11):1170-78. doi: 10.1093/jnci/djz038
11. Fairclough DL. Design and analysis of quality of life studies in clinical trials. New York: Chapman & Hall/CRC Press 2002.
12. Rothrock NE, Kaiser KA, Cella D. Developing a valid patient-reported outcome measure. *Clin Pharmacol Ther* 2011;90(5):737-42. doi: 10.1038/clpt.2011.195 [published Online First: 2011/10/05]
13. 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(10):e76625. doi: 10.1371/journal.pone.0076625 [published Online First: 2013/10/15]
14. Leidy NK, Revicki DA, Geneste B. Recommendations for evaluating the validity of quality of life claims for labeling and promotion. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 1999;2(2):113-27. doi: 10.1046/j.1524-4733.1999.02210.x [published Online First: 2006/05/06]
15. 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 in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 2009;12(4):419-29. doi: 10.1111/j.1524-4733.2008.00470.x [published Online First: 2009/11/11]

16. Rutherford C, Costa D, Mercieca-Bebber R, et al. Mode of administration does not cause bias in patient-reported outcome results: a meta-analysis. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2016;25(3):559-74. doi: 10.1007/s11136-015-1110-8 [published Online First: 2015/09/04]
17. Bernhard J, Cella DF, Coates AS, et al. Missing quality of life data in cancer clinical trials: serious problems and challenges. *Statistics in Medicine* 1998;17(5-7):517-32. doi: 10.1002/(sici)1097-0258(19980315/15)17:5/7<517::Aid-sim799>3.0.Co;2-s
18. 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. doi: 10.1016/j.conctc.2017.11.009 [published Online First: 2018/04/27]
19. 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 in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 2005;8(2):94-104. doi: 10.1111/j.1524-4733.2005.04054.x [published Online First: 2005/04/05]
20. 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(1):e0144658. doi: 10.1371/journal.pone.0144658 [published Online First: 2016/01/20]
21. Kyte D, Draper H, Calvert M. Patient-reported outcome alerts: ethical and logistical considerations in clinical trials. *Jama* 2013;310(12):1229-30. doi: 10.1001/jama.2013.277222 [published Online First: 2013/09/26]
22. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of clinical epidemiology* 2008;61(2):102-9. doi: 10.1016/j.jclinepi.2007.03.012 [published Online First: 2008/01/08]
23. King MT. A point of minimal important difference (MID): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res* 2011;11(2):171-84. doi: 10.1586/erp.11.9 [published Online First: 2011/04/12]
24. FDA. Guidance for Industry, Patient-Reported Outcome Measures: Use in Medical Product Development to Suport Labeling Claims 2009 [Available from: https://www.fda.gov/media/77832/download accessed 25/11/2019.
25. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource 2016 [Available from: https://www.ncbi.nlm.nih.gov/books/NBK326791/.
26. Curtis L HA, Weinfurt K. Choosing and Specifying Endpoints and Outcomes: Introduction. In: Rethinking Clinical Trials: A Living Textbook of Pragmatic Clinical Trials. 2019 [Available from: DOI: 10.28929/009 accessed 12/03/2020.
27. Calvert M, Blazeby J, Altman DG, et al. Reporting of Patient-Reported Outcomes in Randomized Trials: The CONSORT PRO Extension. *JAMA* 2013;309(8):814-22. doi: 10.1001/jama.2013.879
28. Kluetz PG, Chingos DT, Basch EM, et al. 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. doi: 10.14694/edbk_159514 [published Online First: 2016/06/02]
29. ClinicalTrials.gov. Glossary of Common Site Terms [Available from: https://clinicaltrials.gov/ct2/about-studies/glossary accessed 12/03/2020 2020.
30. FDA. Drug Development Tool Glossary [updated 08/15/2017. Available from: https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/ddt-glossary accessed 03/07/2020 2020.
31. Chan A-W, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. *Annals of Internal Medicine* 2013;158(3):200-07. doi: 10.7326/0003-4819-158-3-201302050-00583

32. Chan A-W, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;8(346) doi: 10.1136/bmj.e7586
33. 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. *European Journal of Cancer* 1997;33(1):20-28. doi: https://doi.org/10.1016/S0959-8049(96)00412-1

Supplement 1

Protocol evidence per final SPIRIT-PRO item.

SPIRIT PRO Item number and description	SPIRIT-PRO wording prior to finalisation	% HTA protocols including item	% EPiC protocols including item	% international ovarian cancer protocols including item
(5a) Specify the individual(s) responsible for the PRO content of the trial protocol		6.67%	22.78%	23.1%
(6a) Describe the PRO-specific research question and rationale for PRO assessment and summarise PRO findings in relevant studies	Describe what is currently known about PROs in this area and explain the gaps in the literature	49.33%	32.91%	42.3%
	Provide a rationale for the inclusion of PROs as appropriate to the study population, intervention, context, objectives and setting	8.00%	33.54%	57.7%
(7) State specific PRO objectives or hypotheses (including relevant PRO concepts/domains	State the PRO study objective in relation to PRO domain/s, patient population and timeframe	77.33%	73.42% (17.09% in relation to dimension, population or timeframe)	30.8%
	State the PRO hypothesis and corresponding null hypothesis and to which outcome(s) the hypothesis relates	18.67%	-	PRO hypothesis provided 19.2%
(10) Specify any PRO-specific eligibility criteria (e.g. 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 methods for obtaining the PRO subsample	If PROs will be collected in a subset of the study population or in specific centres, include a description/rationale for the sampling method	0.00%	10.76%	11.5%
	State the inclusion/exclusion criteria for PRO endpoint(s) (e.g., language/reading requirements)	45.33%	50.00%	7.7%

	Specify if PRO completion is pre-randomisation eligibility requirement	-	-	7.7%
(12) Specify the PRO concepts/domains used to evaluate the intervention (e.g. overall health-related quality of life, specific domain, specific symptom) and for each one, the analysis metric (e.g. change from baseline, final value, time to event) and the principal time point or period of interest	Describe the PRO constructs used to evaluate the intervention e.g. overall QOL, specific domain, specific symptom	-	-	73.1%
(13) Include a schedule of PRO assessments, providing a rationale for the time points, 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	Specify the timepoint(s) for PRO analysis (including the principle timepoint of interest) and provide the rationale for these	Timing specified 97.33%	Timing specified 83.54%	42.3%
	Include PRO assessments in the main protocol schedule of assessments, specifying which PRO measures (PROMs) will be used at each assessment	-	-	96.2%
	Specify if baseline PRO assessment should be completed before randomisation	-	-	53.8%
	Specify the targeted time and acceptable time windows for each PRO assessment	-	-	26.9%
	If PROs are to be completed in the clinic: specify timing of PROM delivery in relation to clinical assessments (e.g. before/whilst/after seeing clinician and/or clinical assessments)	-	-	54%
	Justify the timing of PRO assessments. Scheduled	Timing justified 6.67%	Timing justified 12.03%	23.1%

	PRO assessments should link to research questions, hypotheses, length of recall, disease/treatment natural history, planned analysis and time of comparison must be comparable for both arms			
(14) When a PRO is the primary end point, 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 end point, then discuss the power of the principal PRO analyses	If PRO is the primary endpoint, state the required PRO sample size, otherwise discuss the power of the PRO analysis	50.67%	25.95%	30.8%
(18a i) Justify the PRO instrument to be used and describe the domains, number of items, recall period, and instrument scaling and scoring (e.g. 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	Describe the PROMs including, number of items/domains, instrument scaling/scoring, reliability, content and construct validity, responsiveness, sensitivity, acceptability, recall period. Provide references as appropriate	PROM identified 100%; Justification in relation to study hypotheses 41.33%; Justified in relation to measurement properties 37.33%; Justified in relation to acceptability/patient burden 14.67%	PROM identified 63.29%; Justification in relation to study hypotheses 36.71%; Justified in relation to measurement properties 46.84%; Justified in relation to acceptability/patient burden 29.11%	Justification for measure used 84.6%
(18a ii) Include a data collection plan outlining the permitted mode(s) of administration (e.g. paper, telephone, electronic, other)	Include a pre-specified data collection plan	84.00% included brief details of PRO data collection procedures but often omitted information surrounding mode of administration,	57.59%	46.2%

and setting (e.g. clinic, home, other)		setting and proxy reporting: 8.00% included PRO data collection guidelines/training information for trial personnel.		
(18a iii) Specify whether more than one language version will be used and state whether translated versions have been developed using currently recommended methods	Provide evidence of measurement equivalence across modes (i.e., when mixing modes of PRO data collection) and/or of cross cultural validity where different language versions of questionnaires are used	-	-	7.7%
(18a iv) 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		-	-	0.0%
(18b i) Specify PRO data collection and management strategies for minimising avoidable missing data	Specify procedures for data collection and management methods to minimise missing data. E.g. checking completed PROMs (including who will check forms and how will they deal with missing PROMs or missing items).	-	-	38.5%
(18b ii) Describe the process of PRO assessment for participants who discontinue or deviate from the assigned intervention protocol	Establish process for PRO assessment at (and beyond) withdrawal for patients who withdraw early from a study or who go 'off-study'/'off treatment'	-	-	34.6%
(20a) State PRO analysis methods, including any plans for addressing multiplicity/type I (a) error	Include an a priori description of all planned PRO analyses pertaining to the study hypotheses	PRO statistical analysis plan provided? 96.00%	PRO statistical analysis plan provided? 53.16%	61.5%

	Pre-specify sequence of testing/exploratory analyses to control for multiplicity or pre-specify domains (e.g. in a regulatory trial/labelling claim) (Common in pharma trials. Involves pre-specifying domains that alpha would be spent on, or ordering the domains in priority & alpha would be spent down the list)	Plans to address multiplicity of PRO data provided? 1.33%	10.13%	7.7%
(20c) State how missing data will be described and outline the methods for handling missing items or entire assessments (e.g. approach to imputation and sensitivity analyses)	State how missing data will be described	45.33%	30.38%	-
	Describe method for handling missing assessments (e.g. approach to imputation and sensitivity analyses)	45.33%	30.38%	-
	Describe methods for scoring endpoints. Where possible, reference scoring manuals for summated scales from PROM (domain-specific &/or total) & methods for handling missing items, and methodological papers for composite endpoints (e.g. QTWiST)	-	-	53.8%
(22) 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; e.g. in the participant information sheet and consent form	Include an a priori plan for consistent/standardised management of PRO alerts (symptoms reported by patients that exceed a pre-defined level of severity) to be clearly communicated to all appropriate trial staff	10.67%	0.63%	0.0%
	Specify whether PRO forms will be used to	4.00%	3.80%	7.7%

	influence therapy or patient management (i.e. will the clinician use PRO responses to inform the patient's care?)			

Supplement 2

Summary recommendations for inclusion* of PRO specific information in the trial protocol, supplementary documents or training for staff.

SPIRIT PRO Item Description	SPIRIT 2013 / SPIRIT- PRO Item Number	Guidance/training for trial staff (eg, site initiation/face to face or online training/operations manual)	Information/ guidance for participants (eg, participant information sheet)	Statistical Analysis Plan
Administrative information				
Specify the individual(s) responsible for the PRO content of the trial protocol	✓ SPIRIT-5a-PRO Elaboration	+		
Introduction				
Describe the PRO specific research question and rationale for PRO assessment, and summarize PRO findings in relevant studies.	✓ SPIRIT-6a-PRO Extension	✓	✓[1]	
State specific PRO objectives or hypotheses (including relevant PRO concepts/domains).	✓ SPIRIT-7-PRO Extension			✓
Methods: Participants, interventions, and outcomes				
Specify any PRO-specific eligibility criteria (eg, language/reading requirements or pre-randomization completion of PRO). If PROs will not be collected in the entire study sample, provide a rationale and describe the method for obtaining the PRO subsample.	✓ SPIRIT-10-PRO Extension	✓		✓
Identify the PRO endpoint as the primary, secondary (and if so - whether a key/important secondary), or an exploratory endpoint.	✓ SPIRIT-12			✓
Specify the PRO concepts/domains used to evaluate the intervention (eg, overall HRQOL, specific domain, specific symptom) and, for each one, the analysis metric (eg, change from baseline, final value, time to event) and the principal time point or period of interest.	✓ SPIRIT-12-PRO Extension			
Include a schedule of PRO assessments, providing a rationale for the time points, and justifying if the initial assessment is not pre-randomization. Specify: time windows; whether PRO collection is prior to clinical assessments; and if using multiple questionnaires, whether order of administration will be standardized.	✓ SPIRIT-13- PRO Extension	✓		✓
Where 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 PRO endpoint, then discuss the power of the principal PRO analyses.	✓ SPIRIT-14-PRO Elaboration			✓
Methods: Data collection, management, and analysis				

Justify the PRO instrument to be used, and describe domains, number of items, recall period, instrument scaling/scoring (eg, range and direction of scores indicating a good/poor outcome). Evidence of PRO instrument measurement properties, interpretation guidelines, and patient acceptability/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.	✓ SPIRIT-18a(i)-PRO Extension			✓[2]
Provide evidence of measurement equivalence across modes (i.e. when mixing modes of PRO data collection) and/or of cross cultural validity where different language versions of questionnaires are used.				✓[3]
Outline plans for evaluation of measurement properties, if appropriate (eg, if not previously validated in the population of interest).				✓[4]
Specify the estimated time to complete each assessment, and discuss feasibility of assessment for the population.		✓	✓[5]	
Include a data collection plan outlining the permitted mode(s) of administration (eg, paper, telephone, electronic, other) and setting (eg, clinic, home, other).	✓ SPIRIT-18a(ii)-PRO Extension	✓	✓	
Specify whether more than one language version will be used, and state whether translated versions have been developed using currently recommended methods.	✓ SPIRIT-18a(iii)-PRO Extension	✓		
Where the trial context requires someone other than the trial participant to answer on their behalf (a proxy reported outcome), state and justify this. Provide/cite evidence of the validity of proxy assessment if available.	✓ SPIRIT-18a(iv)-PRO Extension	✓		
Specify who will administer the PROM (eg, a physician, nurse etc).		✓		
If it is permissible for another person to help the study participant complete the PROM, describe what type and level of assistance is acceptable.		✓	✓	
Include a plan for systematically training and contacting local site personnel to ensure that they understand the content and importance of collecting PRO data. Ideally coordinated by a lead data manager who monitors PRO completion rates in real time and communicates with sites if completion rates are suboptimal.		✓		
Specify PRO data collection and management strategies for minimising avoidable missing data.	✓ SPIRIT-18b(i)-PRO Extension	✓		
Include guidance on discussing importance of PROs with patient		✓		
Describe the process of PRO assessment for participants who discontinue or deviate from their assigned intervention protocol.	✓ SPIRIT-18b(ii)-PRO Elaboration	✓		
Specify that a named person/position at each centre (and/or centrally) be nominated to take responsibility for administration, collection and checking of PROM, specify whether this is the treating clinician or not.		✓[6]		
Specify how an electronic PRO system/database will be maintained and how the investigator will meet regulatory requirements and ensure data integrity and security.	✓ SPIRIT-19	✓	✓[7]	
Specify plan to monitor PRO compliance, including adherence to time windows.	✓ SPIRIT-19	✓	✓[8]	
Include an overview of PRO administration (data collection), and data handling/transmission and storage procedures	✓ SPIRIT-19	✓	✓[7]	

Item 39: Include an a priori description of all planned PRO analyses pertaining to the study hypotheses. Item 44: Include a priori identified summary statistics (as appropriate).	✓ SPIRIT-20a			✓
State the assumptions of PRO analyses.				✓
Include an a priori estimation of PRO effect size.	✓ SPIRIT-14			✓
Specify intention-to-treat or per-protocol PRO analyses.	✓ SPIRIT-20c			✓
Specify the minimum PRO response rate and acceptable degree of timing deviation (i.e. acceptable time windows for each PRO assessment time point) before the PRO objective is compromised. Specify the minimum PRO response rate and acceptable degree of timing deviation (i.e. acceptable time windows for each PRO assessment time point) before the PRO objective is compromised.	✓ SPIRIT-14			✓
Describe methods for scoring endpoints. Where possible, reference scoring manuals for summated scales from PROM (domain-specific and/or total) and methods for handling missing items, and methodological papers for composite endpoints (eg, QTWiST).	✓ SPIRIT-20a			✓
State PRO analysis methods including any plans for addressing multiplicity/type 1 (α) error.	✓ SPIRIT-20a-PRO Elaboration			✓
Specify the criteria for clinical significance (eg, state minimal [clinical] important difference and/or responder definition (size and duration of benefit)).	✓ SPIRIT-14			✓
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).	✓ SPIRIT-20c PRO Elaboration			✓
Monitoring				
Describe the role of the Data Monitoring Committee and Quality Assurance for PROs.	✓ SPIRIT-21a			
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 standardized way. Describe how this process will be explained to participants, eg, in the participant information sheet and consent form.	✓ SPIRIT-22 PRO Extension	✓	✓[7]	
Ethics and dissemination				
Describe informed consent procedure for PRO assessment.		✓		
Include detailed plans for regular feedback to participants via letter/newsletter on PRO aspect of study.		✓	✓	

Footnotes:*"Recommended for inclusion" indicates that >50% of the Delphi and Stakeholder Survey participants endorsed the item (or at least one item if two items were merged for the consensus meeting) for inclusion and/or the item was considered important for inclusion at the consensus meeting.
+Recommended by the Delphi panel but excluded following discussion at the consensus meeting.
[1] <50% of Delphi panellists voted to include this item, but item was included following discussion at the consensus meeting. It was felt important to include a brief rationale for PRO assessment in the PIS.
[2] Use of the PRO instrument in accordance with the user manual should be specified in the SAP.
[3] Plans for dealing with different modes of administration should be specified in the SAP. This was not supported by the Delphi but was identified as important by the consensus panel.
[4] If validation of the PRO instrument is planned as part of the trial then this should be pre-specified in the main trial SAP. If validation is planned as a separate sub-study, this should be specified a separate study protocol and SAP.
[5] Estimation of time to complete each PRO assessment should be included in guidance and training for staff and in information for trial participants.
[6]This should also be recorded in the delegation of duties log.
[7] The participant information sheet should contain information regarding the storage and security of PRO data and provide information on who will access their PRO data and for what purpose.
[8] If plans to monitor adherence to time windows, include reminders for participants eg. Via text or SMS, the relevant details should be specified in the information to participants.

BMJ Open

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045105.R2
Article Type:	Communication
Date Submitted by the Author:	21-Dec-2020
Complete List of Authors:	Calvert , Melanie ; University of Birmingham, Centre for Patient Reported Outcome Research, Institute of Applied Health Research King, Madeleine; University of Sydney, Faculty of Science, School of Psychology Mercieca-Bebber, Rebecca; University of Sydney, Faculty of Medicine and Health, NHMRC Clinical Trials Centre Aiyegbusi, Olalekan ; University of Birmingham, Centre for Patient Reported Outcome Research, Institute of Applied Health Research Kyte, Derek; University of Birmingham, Centre for Patient Reported Outcome Research, Institute of Applied Health Research Slade, Anita; University of Birmingham, Centre for Patient Reported Outcome Research, Institute of Applied Health Research Chan, An-Wen; University of Toronto, Women's College Research Institute, Department of Medicine Basch, E; University of North Carolina Bell, Jill; AstraZeneca Gaithersburg, Oncology Digital Health Bennett, Antonia; University of North Carolina at Chapel Hill, Cancer Outcomes Research Program Bhatnagar, Vishal; US Food and Drug Administration Blazeby, Jane; University of Bristol, Bristol Biomedical Research Centre; University Hospitals Bristol NHS Foundation Trust, Division of Surgery, Head and Neck Bottomley, Andrew; European Organisation for Research and Treatment of Cancer Brown, Julia; University of Leeds, Clinical Trials Research Unit Brundage, Michael; Queen's University Cancer Research Institute Campbell, Lisa; Medicines and Healthcare Products Regulatory Agency Cappelleri, Joseph; Pfizer Inc, Global Biometrics & Data Management - Statistics Draper, Heather; University of Warwick Dueck, Amylou; Mayo Clinic, Department of Health Sciences Research Ells, Carolyn; Interagency Advisory Panel on Research Ethics Frank, Lori; Patient-Centered Outcomes Research Institute Golub, Robert; JAMA Griebsch, Ingolf; Boehringer Ingelheim International GmbH Haywood, Kirstie; University of Warwick Warwick Medical School, Warwick Research in Nursing Hunn, Amanda; Independent Advisor King-Kallimanis, Bellinda; US Food and Drug Administration

	Martin, Laura; Daiichi Sankyo Europe GmbH Mitchell, Sandra; National Cancer Institute Morel, Thomas; UCB Pharma, Patient-Centered Outcomes Research Nelson, Linda; GSK, Value Evidence and Outcomes - Patient Centered Outcomes Norquist, Josephine; Merck & Co Inc, Center for Observational Real-world Evidence (CORE), Patient-Centered Endpoints & Strategy O'Connor, Daniel; Medicines and Healthcare Products Regulatory Agency Palmer, Michael; Queen's University, Cancer Research Institute Patrick, Donald; University of Washington Price, Gary; University of Birmingham, Centre for Patient Reported Outcome Research, Primary Care and Clinical Sciences Regnault, Antoine; Modus Outcomes Retzer, Ameeta; University of Birmingham, Centre for Patient Reported Outcome Research, Institute for Applied Health Research Revicki, Dennis; Outcomes Research Consulting Scott, Jane; Janssen Global Services LLC, Johnson and Johnson Stephens, Richard; National Cancer Research Institute Consumer Forum Turner, Grace; University of Birmingham, Centre for Patient Reported Outcome Research, Institute of Applied Health Research Valakas, Antonia; EMD Serono Inc, Healthcare business of Merck KGaA, Darmstadt, Germany Velikova, Galina; University of Leeds, Leeds Institute of Medical Research at St James's von Hildebrand, Maria; University of Birmingham, Centre for Patient Reported Outcome Research, Institute of Applied Health Research Walker, Anita; University of Birmingham, Centre for Patient Reported Outcome Research, Institute of Applied Health Research Wenzel, Lari; University of California Irvine
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Health services research
Keywords:	STATISTICS & RESEARCH METHODS, EDUCATION & TRAINING (see Medical Education & Training), Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Clinical trials < THERAPEUTICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in BMJ Open and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Research Methods & Reporting

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

Authors:

Calvert M, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, Birmingham Health Partners Centre for Regulatory Science and Innovation, NIHR Biomedical Research Centre, NIHR Surgical Reconstruction and Microbiology Research Centre and NIHR Applied Research Collaboration West Midlands, University of Birmingham, England

King MT; University of Sydney, Faculty of Science, School of Psychology, Sydney, Australia

Mercieca-Bebber R, University of Sydney, Faculty of Medicine and Health, NHMRC Clinical Trials Centre, Sydney, Australia

Aiyegbusi OL, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, Birmingham Health Partners Centre for Regulatory Science and Innovation, NIHR Biomedical Research Centre and NIHR Applied Research Collaboration West Midlands, University of Birmingham, England

Kyte D, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, England

Slade A, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, England

Chan A-W, Women's College Research Institute, Department of Medicine, University of Toronto, Ontario, Canada

Basch E, University of North Carolina, Chapel Hill

Bell JA, Oncology Digital Health, AstraZeneca, Gaithersburg, Maryland

Bennett AV, Cancer Outcomes Research Program, University of North Carolina, Chapel Hill

Bhatnagar V, US Food and Drug Administration, Silver Spring, Maryland

Blazeby J, Bristol Biomedical Research Centre, University of Bristol, Bristol, England

Bottomley A, European Organization for Research and Treatment of Cancer, Brussels, Belgium

Brown J, Clinical Trials Research Unit, University of Leeds, Leeds, England

Brundage M, Queen's Cancer Research Institute, University, Kingston, Ontario, Canada

Campbell L, Medicines and Healthcare Products Regulatory Agency, London, England

Cappelleri JC, Global Biometrics & Data Management – Statistics, Pfizer Inc, New York, NY

Draper H, University of Warwick, Coventry, England

Dueck A, Department of Health Sciences Research, Mayo Clinic, Scottsdale, Arizona

Ells C, School of Population and Global Health, McGill University, Montreal, Canada

Frank L, Patient-Centered Outcomes Research Institute, Washington, DC

Golub RM, JAMA, Chicago, Illinois

Griebsch I, Boehringer Ingelheim International GmbH, Germany

Haywood K, Warwick Research in Nursing, Warwick Medical School, University of Warwick, Coventry, England

Hunn A, Independent Advisor, UK

King-Kallimanis BL, US Food and Drug Administration, Silver Spring, Maryland

Martin L, Daiichi Sankyo Europe GmbH, Munich, Germany

Mitchell S, National Cancer Institute, Rockville, Maryland

Morel T, Patient-Centered Outcomes Research, UCB Pharma, Brussels, Belgium

Nelsen L, Value Evidence and Outcomes - Patient Centered Outcomes, GSK, Collegeville, PA, USA

Norquist JM, Center for Observational and Real-world Evidence (CORE), Patient-Centered Endpoints & Strategy, Merck & Co., Inc., Kenilworth, NJ, USA

O'Connor D, Medicines and Healthcare Products Regulatory Agency, London, England

Palmer MJ, Cancer Research Institute, Queen's University, Kingston, Ontario, Canada

Patrick D, University of Washington, Seattle

Price G, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, Birmingham, England

Regnault A, Modus Outcomes, Lyon, France

Retzer A, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, England

Revicki D, Outcomes Research Consulting, Sarasota, Florida, USA

Scott J, Janssen Global Services, Johnson and Johnson, High Wycombe, England

Stephens R, National Cancer Research Institute Consumer Forum, London, England

Turner G, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, NIHR Surgical Reconstruction and Microbiology Research Centre, University of Birmingham, England

Valakas A, EMD Serono Inc, a Healthcare business of MerckKGaA, Darmstadt, Germany, Rockland, MA, USA

Velikova G, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds, England

von Hildebrand M, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, Birmingham, England

Walker A, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, Birmingham, England

Wenzel L, University of California, Irvine

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 SPIRIT-PRO Extension. The SPIRIT-PRO Extension is a sixteen-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.

Article Summary

Strengths and limitations of this study

- The SPIRIT-PRO Extension aims to improve the completeness and transparency of trial protocols where patient-reported outcomes (PROs) are a primary or key secondary outcomes and was developed following EQUATOR 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 multi-stakeholder 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.

Background

Clinical trial protocols are essential documents intended to include the study rationale, intervention, trial design methods, study processes, outcomes, sample size, data 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 facilities external appraisal of the scientific, methodological, and ethical rigor 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 (Supplement 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 sparce 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 SPIRIT 2013 Statement (Standard Protocol Items: Recommendations for Interventional Trials). 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 for use in protocol development. Additional information and resources regarding the SPIRIT Initiative are available on the SPIRIT website (www.spirit-statement.org).

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	Please specify which page(s) of the protocol this item is addressed on.
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)			
	2b	All items from the World Health Organization Trial Registration Data Set			
Protocol version	3	Date and version identifier			
Funding	4	Sources and types of financial, material, and other support			
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	SPIRIT5a-PRO Elaboration	Specify the individual(s) responsible for the PRO content of the trial protocol.	
	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, end-point 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 summarize PRO findings in relevant studies.	
	6b	Explanation for choice of comparators			
Objectives	7	Specific objectives or hypotheses	SPIRIT7-PRO Extension	State specific PRO objectives or hypotheses (including relevant PRO concepts/domains).	
Trial design	8	Description of trial design, including type of trial (e.g., parallel group, crossover, factorial, single group), allocation ratio, and framework (e.g., superiority, equivalence, noninferiority, exploratory)			
Methods: Participants, Interventions, and Outcomes					

Study setting	9	Description of study settings (e.g., community clinic, academic hospital) and list of countries where data will be collected; reference to where list of study sites can be obtained			
Eligibility criteria	10	Inclusion and exclusion criteria for participants; if applicable, eligibility criteria for study centres and individuals who will perform the interventions (e.g., surgeons, psychotherapists)	SPIRIT10-PRO Extension	Specify any PRO-specific eligibility criteria (e.g., language/reading requirements or pre-randomisation 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.	
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 (e.g., 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 (e.g., 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 (e.g., systolic blood pressure), analysis metric (e.g., change from baseline, final value, time to event), method of aggregation (e.g., median, proportion), and time point 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 (e.g., overall health-related quality of life, specific domain, specific symptom) and, for each one, the analysis metric (e.g., change from baseline, final value, time to event) and the principal time point or period of interest.	
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 et al)[1 2]	SPIRIT13-PRO Extension	Include a schedule of PRO assessments, providing a rationale for the time points, and justifying if the initial assessment is not pre-randomisation. Specify time windows, whether PRO collection is prior to clinical assessments, and, if using multiple questionnaires, whether order of administration will be standardized.	
Sample size	14	Estimated number of participants needed to achieve study objectives and how it was	SPIRIT14-PRO Elaboration	When a PRO is the primary end point, state the required sample size (and how it was	

		determined, including clinical and statistical assumptions supporting any sample size calculations		determined) and recruitment target (accounting for expected loss to follow-up). If sample size is not established based on the PRO end point, then discuss the power of the principal PRO analyses.	
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 (e.g., computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (e.g., 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 (e.g., 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 (e.g., 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					
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (e.g., duplicate measurements, training of assessors) and description of study instruments (e.g., 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 (e.g., 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.	
			SPIRIT-18a (ii)-PRO Extension	Include a data collection plan outlining the permitted mode(s) of administration (e.g., paper, telephone, electronic, other) and setting (e.g., clinic, home, other).	
			SPIRIT-18a (iii)-PRO Extension	Specify whether more than 1 language version will be used and state whether translated versions have been developed using currently recommended methods.	
			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.	
	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 minimizing 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.	

Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (e.g., double data entry; range checks for data values); reference to where details of data management procedures can be found, if not in the protocol			
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.	
	20b	Methods for any additional analyses (e.g., subgroup and adjusted analyses)			
	20c	Definition of analysis population relating to protocol nonadherence (e.g., as randomized analysis) and any statistical methods to handle missing data (e.g., multiple imputation)	SPIRIT20c-PRO Elaboration	State how missing data will be described and outline the methods for handling missing items or entire assessments (e.g., approach to imputation and sensitivity analyses).	
Methods: Monitoring					
	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			
Harms	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 standardized way. Describe how this process will be explained to participants; eg, in the participant information sheet and consent form.	
Auditing	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			
Protocol amendments	25	Plans for communicating important protocol modifications (e.g., changes to eligibility criteria,			

		outcomes, analyses) to relevant parties (e.g., investigators, research ethics committees/institutional review boards, trial participants, trial registries, journals, regulators)			
Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorized 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, health care 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 authorized 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			

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.	**To be added in proof**
6a	SPIRIT6a-PRO Extension	Describe the PRO-specific research question and rationale for PRO assessment and summarize PRO findings in relevant studies.	**To be added in proof**
7	SPIRIT7-PRO Extension	State specific PRO objectives or hypotheses (including relevant PRO concepts/domains).	**To be added in proof**
10	SPIRIT10-PRO Extension	Specify any PRO-specific eligibility criteria (e.g., language/reading requirements or pre-randomisation 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.	**To be added in proof**
12	SPIRIT12-PRO Extension	Specify the PRO concepts/domains used to evaluate the intervention (e.g., overall health-related quality of life, specific domain, specific symptom) and, for each one, the analysis metric (e.g., change from baseline, final value, time to event) and the principal time point or period of interest.	**To be added in proof**
13	SPIRIT13-	Include a schedule of	**To be added in**

Table 2 Table of contents and resources

	PRO Extension	PRO assessments, providing a rationale for the time points, and justifying if the initial assessment is not pre-randomisation. Specify time windows, whether PRO collection is prior to clinical assessments, and, if using multiple questionnaires, whether order of administration will be standardized.	**proof**
14	SPIRIT14-PRO Elaboration	When a PRO is the primary end point, 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 end point, then discuss the power of the principal PRO analyses.	**To be added in proof**
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 (e.g., 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	**To be added in proof**

		accordance with any user manual and specify and justify deviations if planned.	
	SPIRIT-18a (ii)-PRO Extension	Include a data collection plan outlining the permitted mode(s) of administration (e.g., paper, telephone, electronic, other) and setting (e.g., clinic, home, other).	**To be added in proof**
	SPIRIT-18a (iii)-PRO Extension	Specify whether more than 1 language version will be used and state whether translated versions have been developed using currently recommended methods.	**To be added in proof**
	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.	**To be added in proof**
18b	SPIRIT-18b (i)-PRO Extension	Specify PRO data collection and management strategies for minimizing avoidable missing data.	**To be added in proof**
	SPIRIT-18b (ii)-PRO Elaboration	Describe the process of PRO assessment for participants who discontinue or deviate from the assigned intervention protocol.	**To be added in proof**
20a	SPIRIT20a-PRO Elaboration	State PRO analysis methods, including any plans for addressing multiplicity/type I (α) error.	**To be added in proof**
20c	SPIRIT20c-PRO Elaboration	State how missing data will be described and outline the methods for	**To be added in proof**

		handling missing items or entire assessments (e.g., approach to imputation and sensitivity analyses).	
22	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 standardized way. Describe how this process will be explained to participants; eg, in the participant information sheet and consent form.	**To be added in proof**
Additional Resources			
Table 1. SPIRIT 2013 and SPIRIT-PRO Extension checklist: Recommended Items to Address in a Clinical Trial Protocol			**To be added in proof**
Glossary of terms			**To be added in proof**
Protocol template			**Add link in proof**

The development of the SPIRIT-PRO Extension followed the Enhancing Quality and Transparency of Health Research (EQUATOR) Network's methodological framework for guideline development,[13] and has been published elsewhere.[10] Ethical approval was provided by the University of Birmingham Ethical Review Board (ERN_16-0819). 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 three-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 of 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.

Glossary

Concept: "The specific measurement goal (i.e., 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."[15]

Domain: "A sub concept represented by a score of an instrument that measures a larger concept comprised of multiple domains. For example, psychological function is a larger concept with multiple domains (emotional and cognitive function) that are measured by relevant items."[15]

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 combined[16] (e.g. change from baseline at 6 weeks in mean fatigue score).[17]

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".[18]

Important or key secondary PROs/end points: Some PRO measures (particularly health-related quality-of-life measures) are multidimensional, producing several domain-specific outcome scales; e.g., 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 (CONSORT PRO Extension 2b),[18] 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 (e.g. using the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE™)).[19]

Instrument: "A means to capture data (e.g., 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."[15]

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.[20]

Item: "an individual question, statement, or task (and its standardized response options) that is evaluated by the patient to address a particular concept".[15]

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 health care professional".[21]

Outcome*: the variable to be measured. It is the measurable characteristic that is influenced or affected by an individuals' baseline state or an intervention as in a clinical trial or other exposure[16] (e.g. a fatigue score).

Patient-reported outcome (PRO): 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.[15 18] PROs are assessed by self-reported questionnaires, referred to as PRO measures (PROMs) or instruments.[16]

Primary outcome: the most important outcome in a trial, pre-specified 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".[15]

Secondary outcomes: outcomes pre-specified in the protocol to assess additional effects of the intervention; some PROs may be identified as important or key secondary outcomes.

SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials.[1 2]

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 time point whereby the result would still be deemed to be clinically relevant.[22]

* 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 (i.e. the outcome) by fully specifying four components: measurement variable (e.g. fatigue "in the past week" as measured by the QLQ-C30), analysis metric (e.g., change in fatigue from baseline, final fatigue value, time to clinically important increase in fatigue (and "event"), method of aggregation (e.g., median fatigue, proportion of patients with severe fatigue, proportion of patients with clinically important change in fatigue), and time point. 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 Explanation and Elaboration 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 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, whilst others may wish to add PRO content to existing sections of the protocol.

Modelled after other reporting guidelines,[2 23 24] 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 25] 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: Multicenter 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]"[26]

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.[27] 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 United Kingdom (UK) National Institute for Health Research (NIHR) Health Technology Assessment programme explicitly described who was responsible for the PRO component (Supplement 1).[3]

SPIRIT-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 3, Randomized, Double-Blind, Placebo-Controlled Study Evaluating the Efficacy and Safety of Idelalisib (GS-1101) in Combination with Rituximab for Previously Treated Chronic Lymphocytic Leukemia

PRO endpoints: 2°

"3.5. Endpoint Selection Rationale

3.5.2.2. Health-Related Quality of Life (HRQL)
Direct patient reporting of outcomes using standardized methods has become an increasingly important component of therapeutic assessment. Evaluation of patient-reported outcomes (PROs) is particularly relevant in patients who cannot be cured of disease [Reference].
PRO questionnaires have been previously used in CLL [chronic lymphocytic leukaemia] 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 vs 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 anemia, 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 co-morbid 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 3 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."[28]

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 (e.g. 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 (e.g. 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.[29] This information is also relevant to research ethics committees/IRBs and funders responsible for reviewing the scientific integrity and ethical aspects of the trial.

SPIRIT-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 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, high dose palliative radiotherapy (HDPRT) versus concurrent chemotherapy and HDPRT (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 six 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 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.
- 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"[30]

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 (RATE-AF)

PRO endpoints: 1°, 2°, exploratory

"3.1 Hypothesis

Null Hypothesis for primary outcome: No difference in patient-reported quality of life (measured using the physical functioning domain of the SF36 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 SF36 questionnaire).

3.2 Primary objective

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

3.3 Secondary objectives

- Generic and AF-specific patient-reported HRQL using the SF-36 global and domain-specific scores, the AFEQT overall score and the EQ-5D-5L summary index and visual analogue scale at six and twelve months.
- Echocardiographic left-ventricular ejection fraction (LVEF) and diastolic function (E/e' and composite of diastolic indices) at twelve months.
- Functional assessment, including 6-minute walking distance achieved, change in European Heart Rhythm Association (EHRA) class and cognitive function at six and twelve months.
- Change in B-type natriuretic peptide (BNP) levels as a surrogate for total cardiac strain at six months.
- Change in heart rate from baseline and group comparison using 24-hour ambulatory ECG."[31]

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 time-point (PICOT) and the estimand framework.[32] 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 SISAQOL Consortium,[33] 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.[33] Although the SISAQOL recommendations were published for oncology trials, the principles apply more broadly. Pre-specification of objectives and hypotheses encourages identification of key PRO domains and time-points. 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 health-related quality of life (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 time points during a trial, such as baseline, end of treatment, then a series of longer-term follow-ups. Pre-specification of objectives and

hypotheses – focussing on the most important PRO domains and time points – 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.[18] 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 (e.g. "to compare the effect of treatment A versus treatment B on fatigue") rather than in terms of a particular direction of effect.[2 34] In contrast, the PRO hypothesis states the predicted effect of the interventions on the trial outcomes (e.g. "patients allocated to treatment A will have less fatigue than those allocated to treatment B"). [35-37]

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 (e.g. language/reading requirements or pre-randomisation completion of PRO).[29] 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 (CARES-SF) and Cover Sheet must be completed prior to registration and randomization on SWOG-9133.[38]

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.[29] 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, e.g. anxiety/emotional well-being.[39] 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 randomization should be defined and should not be too long.

It may not always be possible to collect PROs from all study participants, e.g. 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,[40] and data on use of other language or culturally appropriate PRO instruments is often missing from the protocol.[13 41]

Where the needs of specific groups have been identified (e.g. not fluent in English) but not accommodated in the study protocol (e.g. 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.[18]

SPIRIT-12-PRO Extension: Specify the PRO concepts/domains used to evaluate the intervention (e.g., overall health-related quality of life, specific domain, specific symptom) and, for each one, the analysis metric (e.g., change from baseline, final value, time to event) and the principal time point 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 initial rate control therapy (RATE-AF)

PRO endpoints: 1°, 2°, exploratory

"12.1.1 Primary Outcome
Patient-reported quality of life (HRQL) - SF-36 physical component summary score at six months

12.1.2 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 six and twelve months
- AFEQT overall score at six and twelve months"[31]

Trial name: A Phase 3, Randomized, Double-Blind, Placebo-Controlled Study Evaluating the Efficacy and Safety of Idelalisib (GS-1101) in Combination with Rituximab for Previously Treated Chronic Lymphocytic Leukemia

PRO endpoints: 2°

"Change in HRQL domain and symptom scores based on the Functional Assessment of Cancer Therapy: Leukemia (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 randomization 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 randomization 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."[28]

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 (e.g., 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 (e.g., change from baseline, final value, time to event); the method of aggregation, which refers to the summary measure format for each study group (e.g., mean, proportion with score > 2); and the specific measurement time point 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 time-points (i.e. where greatest impact of interventions are expected), and develops that concept further by encouraging protocol contributors to think about how these PRO domains and time-points will be analysed, i.e. the analysis metric.[33] To ensure transparency and credibility of the analysis, it is recommended that there is pre-specification of the PRO concepts/domains, analysis metric(s) and time-point(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.[33 42] The selected key domains, time-points, 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, organizing 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, i.e. direct impact of the disease and treatment (e.g. symptoms such as pain, fatigue, nausea, rash and anxiety) or more distal, "knock-on" effects, (e.g. functional status and global quality of life), as illustrated for ovarian cancer in (Figure 1, inspired by the Wilson and Cleary model[43]). 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.[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 pre-specified alongside the levels of statistical and clinical significance for the study and any responder definition in use.[15] Time-points for analysis should be chosen to best address the research question, whilst 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.[15 45]

The example, Idelalisib and Rituximab Improve Progression-Free Survival Over Rituximab Alone in Unfit Patients with Relapsed Chronic Lymphocytic Leukemia: A Phase 3 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 time-point (6 months) as the primary outcome, with other SF-36 domains, questionnaires and time-points specified as secondary outcomes.

Figure 1 Proximal and distal effects of therapy on patient symptoms and quality of life adapted from Wilson and Cleary, 1995.[43]

SPIRIT-13-PRO Extension: Include a schedule of PRO assessments, providing a rationale for the time points, and justifying if the initial assessment is not pre-randomisation. 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 high dose palliative radiotherapy (HDPRT) versus concurrent chemotherapy + HDPRT (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).[30]

Figure 2 Example schedule of PRO assessments in the TROG 11.03 P-Lung GP Trial

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))

Figure 3 Example schedule of PRO assessments in the MRC SUPREMO TRIAL (BIG 2-04)[46]

Figure 4 Flow diagram schedule of PRO assessments in the MRC SUPREMO TRIAL (BIG 2-04)[46]

Explanation

A clear and concise schedule of PRO assessments (Figures 2, 3 & 4) can: assist trial staff to be organised and prepared for participant visits, inform study participants about the methods and expectations of trial participation, and facilitate review of participant burden by research ethics committees/IRBs.[29] The scheduled PRO assessments should provide the data required to address the study's PRO objectives. When selecting appropriate time-points 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 (e.g. 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 pre-randomisation 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 post-baseline, or on Day 30 +/- 3 days post-surgery. Time-windows for each scheduled PRO assessment require an unambiguous reference point, to ensure that PRO data collection captures clinically relevant time points 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, e.g. 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 (i.e. no net effect on estimates of interest, such as the impact of interventions on PROs) or systematic (i.e. causing bias).

SPIRIT-14-PRO Elaboration: When a PRO is the primary end point, 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 end point, 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 four points on the 0 to 100 ThyPRO composite scale, corresponding to a small to moderate effect. In previously obtained data, the standard deviation of ThyPRO-scores (sigma level) was 20 points. With a correlation between observations on the 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 multicenter 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 cosmesis and body image score (CBIS) is defined as an improvement of 20% of the cosmesis score (8 points). Given the reported standard deviation of the CBIS between 4-6 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 (alpha) of 0.05 and a power of 0.90. Estimating a 10% dropout rate, which is common in randomized controlled trials, 55 patients will be randomized per arm."[48]

Trial group: TROG
Trial name: A randomised phase III study of radiation doses and fractionation schedules for ductal carcinoma in situ (DCIS) 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 standard deviations of a continuous scale such as fatigue or physical symptoms, with 80% power at a two-sided alpha 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.[45] Ideally, the criteria for clinical significance (e.g. 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.[15]

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 (e.g., 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 (e.g. SF-36, Stroke Impact Scale, Quality of Life Index) and the tool chosen was a compromise between psychometric properties and adequacy of content for mild stroke. The 32 item questionnaire Quality of Life Index (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, socio-economic, psychological/spiritual and family), is evaluated in terms of satisfaction and importance on a six-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: a = 0.90) [Reference] and thus should be responsive to therapy-induced change [References]. A one point difference was observed in the first six months post-stroke descriptive follow-up (n = 63) for an effect size of 0.33 [References]. A two-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 multi-centre 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 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 wellbeing.[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.[15 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 (COSMIN) 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 (e.g. 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 standardized 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.[18]

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 (e.g., 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.[29]

SPIRIT-18a(ii)-PRO Extension: Include a data collection plan outlining the permitted mode(s) of administration (e.g., paper, telephone, electronic, other) and setting (e.g., clinic, home, other).

Examples

Trial name: Early surgery versus optimal current step-up practice for chronic pancreatitis: a multi-centre randomised control trial (ESCAPE)

PRO endpoints: 1°, 2°

"The Izbicki pain score will be assessed every two 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 Izbiki pain score forms and return envelops. Patients will be contacted by telephone every two 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 less than 3 minutes. 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 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 (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™ (FACT-G and 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 two-weekly questionnaires which participants complete at home from the 24-week time-point, where the participant consents to this, reminders for completion are sent by email or text message to the participant by the research team at 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 pre-planned 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.[29] Of note, the FDA has previously recommended that there is a back-up plan for electronic PRO data collection (e.g., web-, phone-, 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 (ISPOR) 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 standardized 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).

Further, 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 3, Randomized, Open Label Trial of Lenalidomide/dexamethasone With or Without Elotuzumab in Relapsed or Refractory Multiple Myeloma (ELOQUENT – 2)

PRO endpoints: 2°, exploratory

"5.8 Outcomes Research Assessments
5.8.1 HRQL Assessments
To assess the impact of treatment, subject's quality of life will be measured using 3 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 to 10 on any single item is generally considered to be clinically significant.[*Reference*] Pre-testing was carried out in the UK, Norway, Sweden, Denmark, and Germany. Field testing of the module has been conducted in a range of Phase 3 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- 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 generalizability 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 National Institute for Health Research 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, multi-national 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.[18 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, pre-specification 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 Pediatric Acquired Brain Injury - a Randomized 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 EF 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 (rs=0.88 for teachers, .82 for parents), internal consistency (Cronbach's αs=0.80– 0.98), and moderate correlations have been detected between teacher and parent ratings (rs=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 (i.e. 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."[15 42] The US FDA also discourages the use of proxy-reported outcomes to inform labelling claims, recommending observer reports for observable phenomenon only (e.g. 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, dependant 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 (e.g. age, relationship to the patient, gender, proxy literacy, relationship and exposure to the patient,[92]) 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.

SPIRIT-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 randomized multicentre trial of Carboplatin + Liposomal Doxorubicin vs Carboplatin + Paclitaxel in patients with ovarian cancer (MITO-2 (Multicentre Italian Trials in Ovarian Cancer-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 randomization to Letrozole or placebo for women previously diagnosed with primary breast cancer completing five 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°

"4.9 Quality of Life
Mandatory for NCIC CTG centres and optional for centers within other cooperative groups:
Patient is able (i.e. 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 randomization.

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 sources[95] 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; or when the PRO assessment is overly burdensome on the patient (e.g. 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.[29] 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,[29] 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 (e.g. 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.[29] 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.[29]

The MITO-2 and MA.17R examples above illustrate some of these strategies. MA.17R comes from the Canadian Cancer Trials Group (CCTG), a multi-centre 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 pre-randomisation 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 maximize compliance.

Our prior work suggests that few trials actively specify such procedures in their protocols; 46.7% HTA protocols[3] and 38.5% of international ovarian cancer trials[4] 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 multi-centre, open-label, randomised, two-arm Phase III trial on the effect on progression free survival 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 multi-centre, 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 progression-free survival (PFS); quality of life was a secondary endpoint. The protocol stated that:

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

The protocol also specified post-progression 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 post-progression, which may explain some inconsistency among sites, with some sites collecting PRO data post-progression, and other not.

As stated in the AURELIA PRO paper (*ref, 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 minimize bias,[102] ensures that staff collect all required PRO data in a standardized 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 post-discontinuation/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 3, randomized double-blind, placebo-controlled study of the efficacy and safety of 2 doses of Tofacitinib (Cp-690,550) in subjects with active psoriatic arthritis and an inadequate response to at least one TNF inhibitor (OPAL BEYOND)

PRO endpoints: 1°, 2°, exploratory

"9.2.1. Analysis of Primary Endpoint

There are two primary endpoints in A3921125 and two doses of tofacitinib each of which will be compared to placebo for each endpoint. In order to control for Type I error a step-wise 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 10mg ACR20 response rate at 3 months, tofacitinib 5mg ACR20 response rate at 3 months, tofacitinib 10mg HAQ-DI at 3 months, tofacitinib 5mg HAQ-DI at 3 months. This gate-keeping or step-down approach strongly protects the Type I error rate at the 0.05 (2-sided) level…

9.2.1.1. 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 confidence intervals for the differences. The primary analysis will be ACR20 response rate at Month 3. The ACR20 response rate will also be analysed at other time points as a secondary analysis.
Missing values due to a subject dropping from the study for any reason (e.g., lack of efficacy or adverse event) will be handled by setting the ACR20 value to nonresponsive.

9.2.1.2. 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 nonresponsive. 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.

9.2.2. 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 10mg will be tested vs placebo first, followed by tofacitinib 5mg. 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 (e.g., remission rates) will follow the analyses described above for binary variables (e.g., ACR20) and continuous endpoints will follow the same type of analyses described above for continuous endpoints (e.g., 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).[45] This can be contained by pre-specifying the key PRO domain(s) or overall score of interest and the principal time-point(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 end points are tested in a defined sequence that contains the overall type I error to the desired level, or conventional non-hierarchical methods (e.g. Bonferroni correction), should be specified *a* priori.[15] 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 pre-specified 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 (e.g., 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°

"10.3.2 Compliance
Missing data may hamper assessment of HRQL in clinical trials. This may occur because centers 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 less than 2% on average and should not be a problem. The former problem will be minimized by ensuring that participating centers are properly informed and motivated towards HRQL assessment. During the study, compliance with completing HRQL questionnaires will be investigated at each time point. 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 time point using a chi-square 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 bi-annual compliance reviews, then the Protocol Writing Committee would review this to either improve compliance or consider terminating the HRQL assessment in the trial…

10.4.3 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 nonresponse (the whole questionnaire is missing for a patient). As item non-response occurs less than 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 modeling techniques (i.e. 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 (i.e. there are no PRO data available for analysis from the participant at that assessment time point). 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.[33 109]

2) Handling missed assessments: missing assessments present a major problem for PRO analyses; lead to a loss of power and wider confidence intervals 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 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,[29] 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 (i.e., 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.[33] Additionally, SISAQOL is developing a set of macros to describe patterns of missing data, and to evaluate imputation methods for use in sensitivity analysis.[33 35 37]

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; e.g., in the participant information sheet 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 wellbeing 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 wellbeing 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 wellbeing 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]

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 (for example, 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 i.e., bias caused by "any intervention other than the experimental maneuver 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 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 is 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.

Box Regulatory/HTA perspective

Regulatory agencies, such as the MHRA, EMA and FDA are placing more focus on capturing the patient experience when developing drugs.[15 42 55 119] 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.

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 wellbeing 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 health-care 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 is 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 health-care research: the patient.’[120]

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 didn’t know before, and how that will be used to help people. Because that’s why we participate in research; it’s to help people like us.

Protocol template

The PRO protocol template [journal to link to template] 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

Supplement 2 outlines additional items recommended for inclusion in other trial documentation, such as the SAP, participant information sheet, and training and guidance documents for staff.[10 121] 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 (CRO) 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.[1 2 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-centered 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 results[3] and use of PRO data to inform patient-centered care. Several SPIRIT-PRO items correspond to items on the CONSORT (Consolidated Standards of Reporting Trials) -PRO checklist.[10 18] 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 9 122-125] Worryingly, a recent review of cancer trials suggested that 49,568 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.[122] Non-reporting of PRO findings is widespread,[7 9 122-130] 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 health-care research resources.[131-133] 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.[134] 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,[135] 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.[119 136]

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 institutional review boards 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),[33] reporting (CONSORT-PRO),[13] and presentation of results (Figure 5).[137-139] These tools are currently being disseminated by the PROTEUS 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 methodologic 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.[10 140 141]

Figure 5 Resources to promote high quality PRO trial design, analysis and dissemination [9]

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.

Article Information

Corresponding Author: Melanie Calvert, PhD, Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, Edgbaston, Birmingham B15 2TT, England (m.calvert@bham.ac.uk).

In memory of Doug Altman (1948-2018), statistician, pioneer, luminary. “To maximise the benefit to society, you need to not just do research, but do it well”.

Contributorship: Drs MJ Calvert and MT King 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. Drs MJ Calvert and MT King are cochairs of the SPIRIT-PRO Group.

Concept and design: MJ Calvert, MT King.

Acquisition, analysis, or interpretation of data: MJ Calvert, MT King, R Mercieca-Bebber, OL Aiyegbusi, D Kyte, A Slade, A-W Chan, E Basch, JA Bell, AV Bennett, V Bhatnagar, J Blazeby, A Bottomley, J Brown, M Brundage, L Campbell, JC Cappelleri, H Draper, A Dueck, C Ells, L Frank, RM Golub, I Griebsch, K Haywood, BL King-Kallimanis, L Martin, S Mitchell, T Morel, L Nelson, JM Norquist, D O’Connor, MJ Palmer, D Patrick, G Price, A Regnault, A Retzer, D Revicki, J Scott, R Stephens, G Turner, A Valakas, G Velikova, M von Hildebrand, A Walker, L Wenzel.

Drafting of the manuscript: MJ Calvert, D Kyte, R Mercieca-Bebber, A Slade, A-W Chan, MT King.+ section writers G Velikova, A Regnault, D Revicki, AV Bennett, S Mitchell, L Wenzel, MJ Palmer, J Brown, G Turner, A Retzer, A Walker

Critical revision of the manuscript for important intellectual content: MJ Calvert, MT King, R Mercieca-Bebber, OL Aiyegbusi, D Kyte, A Slade, A-W Chan, E Basch, JA Bell, AV Bennett, V Bhatnagar, J Blazeby, A Bottomley, J Brown, M Brundage, L Campbell, JC Cappelleri, H Draper, A Dueck, C Ells, L Frank, RM Golub, I Griebsch, K Haywood, BL King-Kallimanis, L Martin, S Mitchell, T Morel, L Nelson, JM Norquist, D O’Connor, MJ Palmer, D Patrick, G Price, A Regnault, A Retzer, D Revicki, J Scott, R Stephens, G Turner, A Valakas, G Velikova, M von Hildebrand, A Walker, L Wenzel.

Obtained funding: MJ Calvert, D Kyte, MT King, D Altman, J Blazeby, J Brown, M Brundage, J Coast, H Draper, M von Hildebrand, J Ives, R Mercieca-Bebber, G Price, L Roberts, A Slade.

Supervision: MJ Calvert

Conflict of Interest Disclosures: 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. Calvert 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. Calvert has received personal fees from Astellas, Takeda, Merck, Daiichi Sankyo,

Glaukos, GSK and the Patient-Centered Outcomes Research Institute (PCORI) outside the submitted work. Mercieca-Bebber reports non-financial support from University of Birmingham. Aiyegbusi, Kyte, and Retzer, reports grants from National Institute of Health Research (NIHR). Aiyegbusi and Kyte report grants from Birmingham Biomedical Research Centre (BRC). Aiyegbusi reports grants from UCB Pharma and also receives funding from the Health Foundation and declares personal fees from Gilead Sciences Ltd. Kyte and Retzer reports grants from Innovate UK and Macmillan Cancer Support. Kyte 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. Basch 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. Cappelleri reports other from Pfizer Inc., and is an employee and a stockholder of Pfizer Inc. Griebsch is a fully paid employee of Boehringer Ingelheim International GmbH. Ells is Chair of the Government of Canada Interagency Advisory Panel on Research Ethics. Martin reports non-financial support from Daiichi-Sankyo and Cell & Gene Therapy Catapult. Morel reports other from UCB. Nelson reports other from GlaxoSmithKline, including employment and ownership of stock in GSK. Stephens 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. Walker 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 three years, no other relationships or activities that could appear to have influenced the submitted work.

Funding/Support: 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

Calvert 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. King is supported by the Australian government through Cancer Australia. Mercieca-Bebber is supported by the Australian Government by a National Health and Medical Research Council (NHMRC) research fellowship. Blazeby is supported by the NIHR Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust. She is also an NIHR Senior Investigator.

Role of the Funder/Sponsor: 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. Work was funded by an unrestricted grant from UCB

Pharma, and a participant (Thomas Morel) contributed as a co-author and member of the industry advisory group.

Disclaimer: The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, the NIHR, the Department of Health, or any other employer or institution. Views of authors, Delphi participants, and stakeholder participants are individual views and may not represent the views of the broader stakeholder group or host institution. This work reflects the views of the authors and should not be construed in any way to represent the views or policies of the US Food and Drug Administration, Medicines and Healthcare products Regulatory Agency (MHRA), or any other employer or institution.

Additional Contributions and Acknowledgements: With thanks to 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] eAppendix in Supplement 1 made by the SPIRIT-PRO Executive, the ISOQOL Best Practices for PROs in Randomized Clinical Trials Protocol Checklist Taskforce, the international stakeholders responsible for stakeholder survey distribution and stakeholders who completed the stakeholder survey, the Delphi panellists andthe SPIRIT-PRO International Consensus Meeting Participants.

Data sharing

References to protocols available in the public domain are provided. Permissions to access unpublished protocols may be sought via request to Melanie Calvert: m.calvert@bham.ac.uk. Requests will be forwarded to the relevant research team for consideration.

We would like to thank the trial groups, the Trans Tasman Radiation Oncology Group, the Australasian Leukaemia and Lymphoma Group, Breast International Group (BIG), Canadian Cancer Trials Group (CCTG), Cancer Trials Ireland, European Organisation for Research and Treatment of Cancer (EORTC), International Breast Cancer Study Group (IBCSG) and Scottish Cancer Trials Breast Group (SCTBG); and the individual investigators who granted permission for us to publish excerpts from their trial protocols; and those trial teams that made their protocols publicly available through the published domain.

References

1. Chan A-W, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. *Annals of Internal Medicine* 2013;158(3):200-07. doi: 10.7326/0003-4819-158-3-201302050-00583
2. Chan A-W, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;8(346) doi: 10.1136/bmj.e7586
3. 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(10):e110229. doi: 10.1371/journal.pone.0110229 [published Online First: 2014/10/22]
4. Mercieca-Bebber R, Friedlander M, Kok PS, et al. The patient-reported outcome content of international ovarian cancer randomised controlled trial protocols. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2016;25(10):2457-65. doi: 10.1007/s11136-016-1339-x [published Online First: 2016/06/14]
5. Kyte D, Retzer A, Ahmed K, et al. Systematic Evaluation of Patient-Reported Outcome Protocol Content and Reporting in Cancer Trials. *JNCI: Journal of the National Cancer Institute* 2019 doi: 10.1093/jnci/djz038
6. 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(10):e012281. doi: 10.1136/bmjopen-2016-012281 [published Online First: 2016/10/05]
7. 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(1):5. doi: 10.1186/s41687-017-0008-3 [published Online First: 2017/01/01]
8. 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 Out* 2019;17(1):156-56. doi: 10.1186/s12955-019-1220-z

9. Kyte DG, Retzer A, Keeley T, et al. Systematic evaluation of Patient-Reported Outcome (PRO) protocol content and reporting in cancer clinical trials: the EPiC study. *Quality of Life Research* 2017;26(1):55-56.
10. Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for Inclusion of Patient-Reported Outcomes in Clinical Trial Protocols: The SPIRIT-PRO Extension. *JAMA* 2018;319(5):483-94. doi: 10.1001/jama.2017.21903 [published Online First: 2018/02/08]
11. 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(10):e76625. doi: 10.1371/journal.pone.0076625 [published Online First: 2013/10/15]
12. 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. *Contemporary Clinical Trials Communications* 2018;9:23-32. doi: 10.1016/j.conctc.2017.11.009
13. Moher D, Schulz KF, Simera I, et al. Guidance for Developers of Health Research Reporting Guidelines. *PLOS Medicine* 2010;7(2):e1000217. doi: 10.1371/journal.pmed.1000217
14. 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(10):e110216. doi: 10.1371/journal.pone.0110216
15. FDA. Guidance for Industry, Patient-Reported Outcome Measures: Use in Medical Product Development to Suport Labeling Claims 2009 [Available from: https://www.fda.gov/media/77832/download accessed 25/11/2019.
16. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource 2016 [Available from: https://www.ncbi.nlm.nih.gov/books/NBK326791/.
17. Curtis L HA, Weinfurt K. Choosing and Specifying Endpoints and Outcomes: Introduction. In: Rethinking Clinical Trials: A Living Textbook of Pragmatic Clinical Trials. 2019 [Available from: DOI: 10.28929/009 accessed 12/03/2020.

18. Calvert M, Blazeby J, Altman DG, et al. Reporting of Patient-Reported Outcomes in Randomized Trials: The CONSORT PRO Extension. *JAMA* 2013;309(8):814-22. doi: 10.1001/jama.2013.879
19. Kluetz PG, Chingos DT, Basch EM, et al. 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. doi: 10.14694/edbk_159514 [published Online First: 2016/06/02]
20. ClinicalTrials.gov. Glossary of Common Site Terms [Available from: https://clinicaltrials.gov/ct2/about-studies/glossary accessed 12/03/2020 2020.
21. FDA. Drug Development Tool Glossary [updated 08/15/2017. Available from: https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/ddt-glossary accessed 03/07/2020 2020.
22. 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. *European Journal of Cancer* 1997;33(1):20-28. doi: https://doi.org/10.1016/S0959-8049(96)00412-1
23. 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. doi: 10.1136/bmj.c869
24. 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;363 doi: 10.1136/bmj.k1614
25. SPIRIT-Statement [Available from: https://www.spirit-statement.org/ accessed 27/04/2020 2020.
26. 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. *Journal of Clinical Oncology* 2014;32(17):1804-11. doi: 10.1200/jco.2013.54.3694
27. 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. doi: 10.1136/bmj.j3453 [published Online First: 2017/08/05]

28. Furman RR, Sharman JP, Coutre SE, et al. Idelalisib and Rituximab in Relapsed Chronic Lymphocytic Leukemia. *New England Journal of Medicine* 2014;370(11):997-1007. doi: 10.1056/NEJMoa1315226
29. 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(6):e010938. doi: 10.1136/bmjopen-2015-010938 [published Online First: 2016/06/18]
30. 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 from: https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=363229&showOriginal=true&isReview=true accessed 05/12/2019.
31. RATE-AF Trial Protocol [Available from: https://www.birmingham.ac.uk/research/activity/mds/trials/bctu/trials/portfolio-v/RateAF/index.aspx accessed 21/11/2019.
32. EMA. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials 2019 [Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf.
33. 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. *The Lancet Oncology* 2020;21(2):e83-e96. doi: 10.1016/S1470-2045(19)30790-9
34. Fleming TR. Clinical trials: discerning hype from substance. *Ann Intern Med* 2010;153(6):400-6. doi: 10.7326/0003-4819-153-6-201009210-00008 [published Online First: 2010/09/22]
35. 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. *The Lancet Oncology* 2018;19(9):e459-e69. doi: https://doi.org/10.1016/S1470-2045(18)30418-2

36. 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(10):e582-e89. doi: 10.1016/s1470-2045(19)30335-3 [published Online First: 2019/10/04]
37. 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. *The Lancet Oncology* 2016;17(11):e510-e14. doi: https://doi.org/10.1016/S1470-2045(16)30510-1
38. SWOG-9208 Quality of Life and Health Status in Patients With Stage I or Stage II Hodgkin's Disease [Available from: https://www.clinicaltrials.gov/ct2/show/NCT00002563 accessed 25/11/2019.
39. Brooks MM, Jenkins LS, Schron EB, et al. Quality of Life at Baseline: Is Assessment After Randomization Valid? *Medical Care* 1998;36(10):1515-19.
40. Blümle A, Meerpohl JJ, Rücker G, et al. Reporting of eligibility criteria of randomised trials: cohort study comparing trial protocols with subsequent articles. *BMJ* 2011;342:d1828. doi: 10.1136/bmj.d1828
41. Calvert M, Kyte D, von Hildebrand M, et al. Putting patients at the heart of health-care research. *The Lancet* 2015;385(9973):1073-74. doi: https://doi.org/10.1016/S0140-6736(15)60599-2
42. 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 from: https://www.ema.europa.eu/en/documents/other/appendix-2-guideline-evaluation-anticancer-medicinal-products-man_en.pdf.
43. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *Jama* 1995;273(1):59-65. [published Online First: 1995/01/04]
44. 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(7):1553-8. doi: 10.1158/1078-0432.CCR-15-2035 [published Online First: 2016/01/14]

45. 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. doi: 10.1111/j.1524-4733.2007.00273.x [published Online First: 2007/11/13]
46. Supremo useful links [Available from: https://www.supremo-trial.com/useful-links.asp accessed 19/03/20 2020.
47. Winther KH, Watt T, Bjørner JB, et al. The chronic autoimmune thyroiditis quality of life selenium trial (CATALYST): study protocol for a randomized controlled trial. *Trials* 2014;15(1):115. doi: 10.1186/1745-6215-15-115
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 Surgery* 2011;11(1):24. doi: 10.1186/1471-2482-11-24
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. *The Lancet Oncology* 2020 doi: https://doi.org/10.1016/S1470-2045(20)30085-1
50. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology* 2008;61(2):102-09. doi: https://doi.org/10.1016/j.jclinepi.2007.03.012
51. King MT. A point of minimal important difference (MID): a critique of terminology and methods. *Expert Review of Pharmacoeconomics & Outcomes Research* 2011;11(2):171-84. doi: 10.1586/erp.11.9
52. Regnault A, Hamel JF, Patrick DL. Pooling of cross-cultural PRO data in multinational clinical trials: how much can poor measurement affect statistical power? *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2015;24(2):273-7. doi: 10.1007/s11136-014-0765-x [published Online First: 2014/09/06]

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 Neurology* 2010;10(1):3. doi: 10.1186/1471-2377-10-3
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. doi: 10.1186/1471-2407-11-466 [published Online First: 2011/11/01]
55. Kluetz PG, O'Connor DJ, Soltys K. Incorporating the patient experience into regulatory decision making in the USA, Europe, and Canada. *The Lancet Oncology* 2018;19(5):e267-e74. doi: https://doi.org/10.1016/S1470-2045(18)30097-4
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. doi: 10.1136/bmj.k5267 [published Online First: 2019/01/27]
57. COMET Initiative [Available from: http://www.comet-initiative.org/ accessed 25/11/19.
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. doi: 10.1111/j.1524-4733.2007.00275.x [published Online First: 2007/11/13]
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(8):967-77. doi: 10.1016/j.jval.2011.06.014 [published Online First: 2011/12/14]
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(8):978-88. doi: 10.1016/j.jval.2011.06.013 [published Online First: 2011/12/14]

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. *Quality of Life Research* 2013;22(8):1889-905. doi: 10.1007/s11136-012-0344-y
62. Prinsen CA, 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(1):449. doi: 10.1186/s13063-016-1555-2 [published Online First: 2016/09/14]
63. Valderas JM, Ferrer M, Mendívil J, et al. Development of EMPRO: A Tool for the Standardized Assessment of Patient-Reported Outcome Measures. *Value in Health* 2008;11(4):700-08. doi: 10.1111/j.1524-4733.2007.00309.x
64. Rothrock NE, Kaiser KA, Cella D. Developing a Valid Patient-Reported Outcome Measure. *Clinical Pharmacology & Therapeutics* 2011;90(5):737-42. doi: 10.1038/clpt.2011.195
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 related outcome measures* 2018;9:353-67. doi: 10.2147/PROM.S156279
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 Gastroenterology* 2013;13(1):49. doi: 10.1186/1471-230X-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(1):598. doi: 10.1186/1471-2407-12-598
68. Standard vs Modified Drug Therapy (STAR) 2015 [Available from: https://njl-admin.nihr.ac.uk/document/download/2024304 accessed 03/05/20 2020.
69. FDA. Methods to Identify What is Important to Patients & Select, Develop or Modify Fit-for-Purpose Clinical Outcomes Assessments 2018 [Available from: 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. *Quality of Life Research* 2016;25(3):559-74. doi: 10.1007/s11136-015-1110-8

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 in Health* 2009;12(4):419-29. doi: https://doi.org/10.1111/j.1524-4733.2008.00470.x

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(5):501-16. doi: 10.1016/j.jval.2014.06.005 [published Online First: 2014/08/17]

73. Lonial S, Dimopoulos M, Palumbo A, et al. Elotuzumab Therapy for Relapsed or Refractory Multiple Myeloma. *New England Journal of Medicine* 2015;373(7):621-31. doi: 10.1056/NEJMoa1505654

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 in Health* 2005;8(2):94-104. doi: 10.1111/j.1524-4733.2005.04054.x

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 in Health* 2009;12(4):430-40. doi: http://dx.doi.org/10.1111/j.1524-4733.2008.00471.x

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. *The Lancet Oncology* 2011;12(13):1240-48. doi: https://doi.org/10.1016/S1470-2045(11)70212-1

77. Slade A, Retzer A, Ahmed K, et al. Systematic Review of the Use of Translated Patient Reported Outcome Measures in Cancer Trials. Submitted, under review.

78. EORTC Translation Procedure Manual [Available from: https://www.eortc.org/app/uploads/sites/2/2018/02/translation_manual_2017.pdf accessed 25/11/2019.

79. Herdman M, Fox-Rushby J, Badia X. 'Equivalence' and the translation and adaptation of health-related quality of life questionnaires. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 1997;6(3):237-47. doi: 10.1023/a:1026410721664 [published Online First: 1997/04/01]

80. Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 1998;7(4):323-35. doi: 10.1023/a:1024985930536 [published Online First: 1998/06/04]

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(4):1175-77. doi: https://doi.org/10.1378/chest.09-1684

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 - The PRO-ESOR project. *Medical Care* 2004;42(1):37-48. doi: 10.1097/01.mlr.0000103529.63132.77

83. Regnault A, Herdman M. Using quantitative methods within the Universalist model framework to explore the cross-cultural equivalence of patient-reported outcome instruments. *Quality of Life Research* 2015;24(1):115-24. doi: 10.1007/s11136-014-0722-8

84. Hypher R, Brandt A, 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. doi: 10.1136/bmjopen-2019-029273

85. Surr CA, Walwyn REA, Lilley-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(1):300. doi: 10.1186/s13063-016-1416-z

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. *Palliative medicine* 2016;30(1):6-22. doi: 10.1177/0269216315589898 [published Online First: 2015/06/13]
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. *Journal of pain and symptom management* 2013;46(6):925-37. doi: https://doi.org/10.1016/j.jpainsymman.2013.01.010
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. *Journal of pain and symptom management* 2011;42(3):354-65. doi: 10.1016/j.jpainsymman.2010.11.018 [published Online First: 2011/04/02]
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 services research* 2008;43(3):833-48. doi: 10.1111/j.1475-6773.2007.00820.x [published Online First: 2008/05/06]
90. Howland M, Allan KC, Carlton CE, et al. Patient-rated versus proxy-rated cognitive and functional measures in older adults. *Patient related outcome measures* 2017;8:33-42. doi: 10.2147/prom.s126919 [published Online First: 2017/03/30]
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. *Journal of pain and symptom management* 2013;46(6):925-37. doi: 10.1016/j.jpainsymman.2013.01.010 [published Online First: 2013/05/01]
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 services research 2018;53(5):3809-24. doi: 10.1111/1475-6773.12980 [published Online First: 2018/05/29]

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. *J Clin Oncol* 2011;29(27):3628-35. doi: 10.1200/jco.2010.33.8566 [published Online First: 2011/08/17]
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. *Journal of Clinical Oncology* 2018;36(6):563-71. doi: 10.1200/JCO.2017.75.7500
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. *Clin Trials* 2018;15(1):95-106. doi: 10.1177/1740774517741113 [published Online First: 2017/11/11]
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. *Statistics in medicine* 1998;17(5-7):781-96. doi: 10.1002/(sici)1097-0258(19980315/15)17:5/7<781::aid-sim821>3.0.co;2-o [published Online First: 1998/04/29]
97. Fairclough DL, Peterson HF, Chang V. Why are missing quality of life data a problem in clinical trials of cancer therapy? *Statistics in medicine* 1998;17(5-7):667-77. doi: 10.1002/(sici)1097-0258(19980315/15)17:5/7<667::aid-sim813>3.0.co;2-6 [published Online First: 1998/04/29]
98. Bell ML, Fiero M, Horton NJ, et al. Handling missing data in RCTs; a review of the top medical journals. *BMC Medical Research Methodology* 2014;14(1):118. doi: 10.1186/1471-2288-14-118
99. FDA. FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency 2020 [updated 21/09/2020. Available from: https://www.fda.gov/media/136238/download accessed 22/10/2020 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. *Journal of Clinical Oncology* 2014;32(13):1302-08. doi: 10.1200/jco.2013.51.4489

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. *Journal of Clinical Oncology* 2014;32(13):1309-16. doi: 10.1200/jco.2013.51.4240
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. *New England Journal of Medicine* 2017;377(16):1525-36. doi: 10.1056/NEJMoa1615977
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. *JNCI: Journal of the National Cancer Institute* 1993;85(5):365-76. doi: 10.1093/jnci/85.5.365
105. FDA. Multiple Endpoints in Clinical Trials. Guidance for Industry 2017 [Available from: https://www.fda.gov/media/102657/download accessed 19/03 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. *Journal of Clinical Oncology* 2013;32(4):320-26. doi: 10.1200/JCO.2013.50.5669
107. Fairclough DL, Cella DF. Functional Assessment of Cancer Therapy (FACT-G): Non-response to individual questions. 1996;5(3):321-29. doi: 10.1007/bf00433916
108. Fayers P, Aaronson N, Bjordal K, et al. Bottomley A, on behalf of the EORTC Quality of Life Group: The EORTC QLQ–C30 Scoring Manual 3rd edition. 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 Medical Research Methodology* 2011;11(1):105. doi: 10.1186/1471-2288-11-105

110. Little RJ, Cohen ML, Dickersin K, et al. The design and conduct of clinical trials to limit missing data. 2012;31(28):3433-43. doi: 10.1002/sim.5519
111. McIsaac MA, Cook RJ, Poulin-Costello M. Incomplete Data in Randomized Dermatology Trials: Consequences and Statistical Methodology. *Dermatology* 2013;226(1):19-27. doi: 10.1159/000346247
112. Gamble C, Krishan A, Stocken D, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *JAMA* 2017;318(23):2337-43. doi: 10.1001/jama.2017.18556
113. Mazza G CC, Madeline Pe, Dorme L, Oliver K, Bottomley A, Sloan J, Dueck A, . 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.
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(10):e026080. doi: 10.1136/bmjopen-2018-026080
115. Kyte D, Draper H, Calvert M. Patient-Reported Outcome Alerts: Ethical and Logistical Considerations in Clinical Trials. *JAMA* 2013;310(12):1229-30. doi: 10.1001/jama.2013.277222
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(1):e0144658. doi: 10.1371/journal.pone.0144658
117. Sackett DL. Clinician-trialist rounds: 5. Cointervention bias – how to diagnose it in their trial and prevent it in yours. *Clinical Trials* 2011;8(4):440-42. doi: 10.1177/1740774511410995
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. FDA. CDER Patient-Focused Drug Development 2019 [Available from: https://www.fda.gov/drugs/development-approval-process-drugs/cder-patient-focused-drug-development accessed 23/04 2020.
120. 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. doi: 10.1111/j.1524-4733.2007.00267.x [published Online First: 2007/11/13]
121. Calvert M, Kyte D, Mercieca-Bebber R, et al. Supplement 3 - Summary Recommendations for Inclusion of PRO-Specific Information in the Trial Protocol, Supplementary Documents for Training for Staff: 2018; [Available from: https://jamanetwork.com/journals/jama/fullarticle/2671472.
122. Schandelmaier S, Conen K, von Elm E, et al. Planning and reporting of quality-of-life outcomes in cancer trials. *Annals of Oncology* 2015;26(9):1966-73. doi: 10.1093/annonc/mdv283
123. 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(18):3335-42. doi: 10.1002/cncr.29489
124. Mercieca-Bebber R, Rouette J, Calvert M, et al. Preliminary evidence on the uptake, use and benefits of the CONSORT-PRO extension. *Quality of Life Research* 2017;26(6):1427-37. doi: 10.1007/s11136-017-1508-6
125. 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. *Annals of Oncology* 2018;29(12):2288-95. doi: https://doi.org/10.1093/annonc/mdy449
126. 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. *Critical Reviews in Oncology/Hematology* 2020;146:102877. doi: https://doi.org/10.1016/j.critrevonc.2020.102877
127. 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. *Clinical Genitourinary Cancer* 2019;17(5):332-47.e2. doi: https://doi.org/10.1016/j.clgc.2019.07.007

128. Mathoulin-Pelissier S, Gourgou-Bourgade S, Bonnetain F, et al. Survival End Point Reporting in Randomized Cancer Clinical Trials: A Review of Major Journals. *Journal of Clinical Oncology* 2008;26(22):3721-26. doi: 10.1200/JCO.2007.14.1192
129. 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. doi: https://doi.org/10.1016/j.lungcan.2019.10.022
130. 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. *Journal of Clinical Oncology* 2008;26(13):2205-11. doi: 10.1200/JCO.2007.14.6795
131. Ioannidis JPA, Greenland S, Hlatky MA, et al. Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet* 2014;383(9912):166-75. doi: 10.1016/S0140-6736(13)62227-8
132. Glasziou P, Altman DG, Bossuyt P, et al. Reducing waste from incomplete or unusable reports of biomedical research. *The Lancet* 2014;383(9913):267-76. doi: 10.1016/S0140-6736(13)62228-X
133. European Commission. Clinical Trials Regulations (EU) No 536/2014 2019 [Available from: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/regulation5362014_qa_en.pdf accessed 19/03/2020.
134. 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 from: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/2017_01_26_summaries_of_ct_results_for_laypersons.pdf accessed 03/05/20 2020.
135. Rivera S, Stephens R, Mercieca-Bebber R, et al. “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

136. Turner G, Aiyegbusi OL, Price G, et al. Moving beyond project-specific patient and public involvement in research. *Journal of the Royal Society of Medicine* 2020;113(1):16-23. doi: 10.1177/0141076819890551

137. Snyder C, Smith K, Holzner B, et al. Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data. *Quality of Life Research* 2019;28(2):345-56. doi: 10.1007/s11136-018-2020-3

138. 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(10):1848-59. doi: 10.1002/cncr.30530

139. Brundage M, Blackford A, Tolbert E, et al. Presenting comparative study PRO results to clinicians and researchers: beyond the eye of the beholder. *Quality of Life Research* 2018;27(1):75-90. doi: 10.1007/s11136-017-1710-6

140. 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. *Quality of Life Research* 2017;26(6):1393-404. doi: 10.1007/s11136-016-1465-5

141. Frank L, Basch E, Selby JV, et al. The PCORI Perspective on Patient-Centered Outcomes Research. *JAMA* 2014;312(15):1513-14. doi: 10.1001/jama.2014.11100

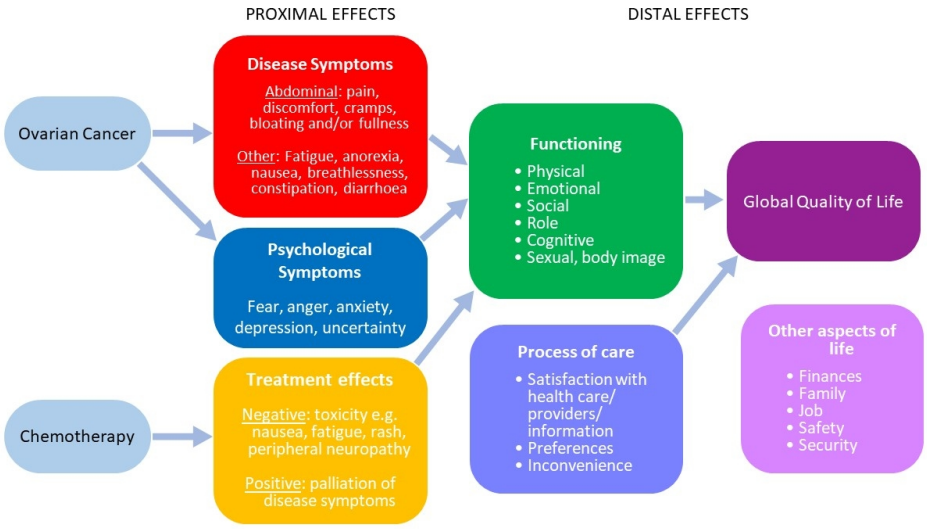


Figure 1 Proximal and distal effects of therapy on patient symptoms and quality of life adapted from Wilson and Cleary, 1995.

338x190mm (96 x 96 DPI)

PRO endpoints: 1°, 2°

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 PRO assessments in the TROG 11.03 P-Lung GP Trial

Visits[a] / Investigations	Patients involved	Screening	Post (+/-) chemo pre (+/-) RT	Post (+/-) RT	1 yr	2 yr	3 yr	4 yr	5 yr	6 yr	7 yr	8 yr	9 yr	10 yr	Recurrence[c]
		Baseline 1	2	3	4	5	6	7	8	9	10	11	12	13	
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[1]	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 (d)	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.
(i) **The only exception are patients in the cardiac substudy who receive chemotherapy. This is the only group of patients who must attend a post chemotherapy visit.**
(ii) **For patients receiving chemotherapy, follow up will be on completion of radiotherapy or at 3 months after chemotherapy in non-irradiated patients.**
For patients not receiving chemotherapy follow up will be on completion of radiotherapy or at 3 months after surgery in non-irradiated patients.
(b) Questioning for symptoms of recurrent breast cancer, examination of loco-regional area and other relevant clinical areas for evidence of recurrence depending on clinicalfeatures.
(c) 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.
(d) 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 PRO assessments in the MRC SUPREMO TRIAL (BIG 2-04)

141x96mm (144 x 144 DPI)

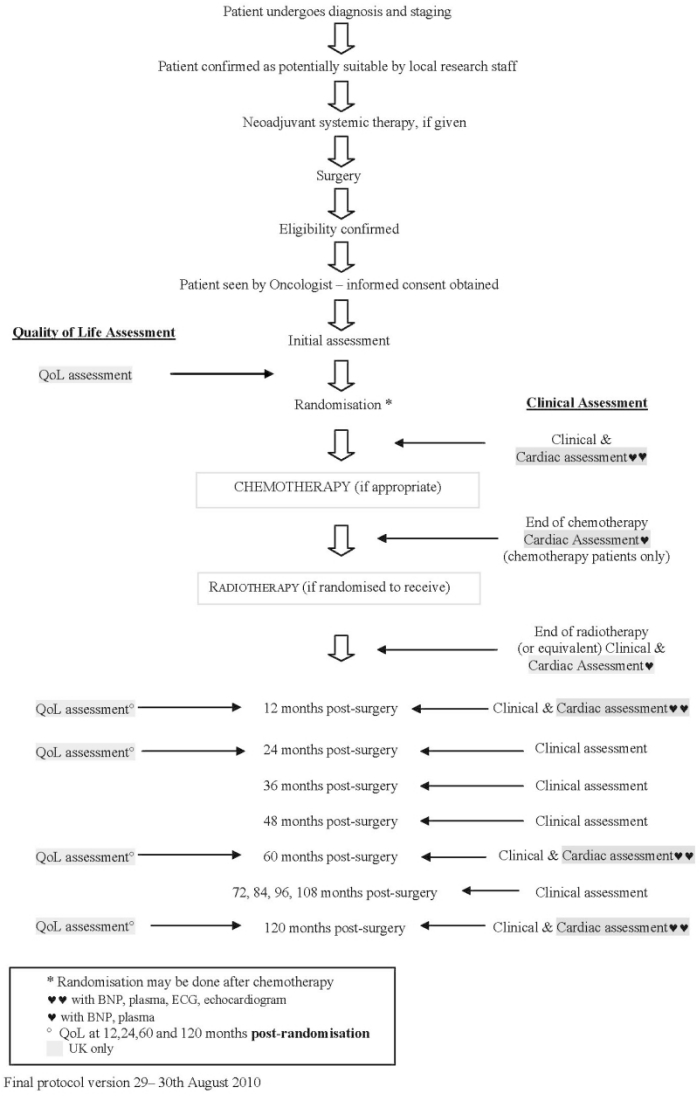


Figure 4 Flow diagram schedule of PRO assessments in the MRC SUPREMO TRIAL (BIG 2-04)

210x297mm (200 x 200 DPI)

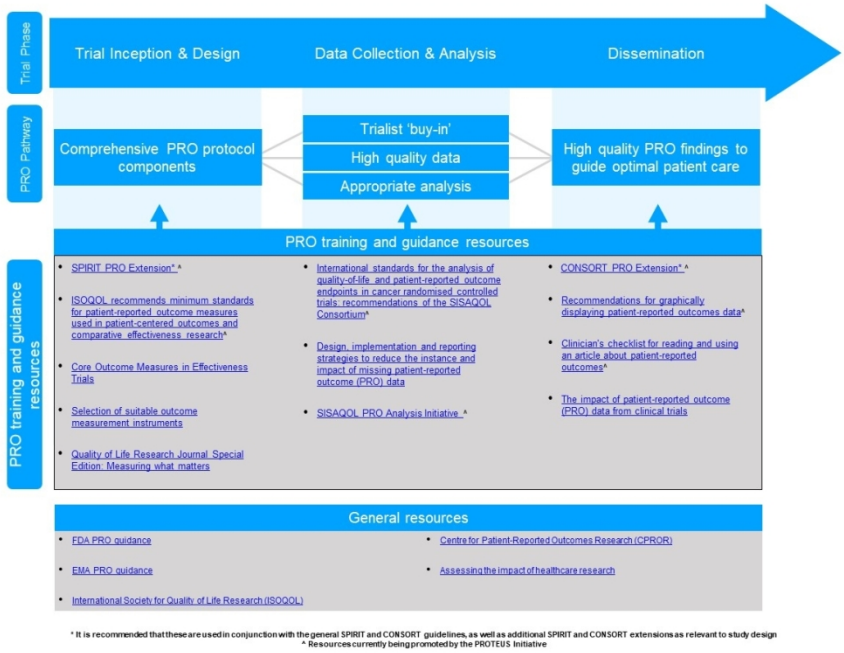


Figure 5 Resources to promote high quality PRO trial design, analysis and dissemination

361x270mm (96 x 96 DPI)

SPIRIT-PRO

PROtocol Template

About this PROtocol Template

Patient-reported outcome (PRO) data from clinical trials can provide valuable evidence to inform shared decision making, labeling claims, clinical guidelines, and health policy; however, the PRO content of clinical trial protocols is often suboptimal.

To address this issue an international, consensus-based, PRO-specific protocol guidance (the SPIRIT-PRO Extension) was developed and published in 2018:

> Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for Inclusion of Patient-Reported Outcomes in Clinical Trial Protocols: The SPIRIT-PRO Extension. *JAMA.* 2018;319(5):483–494. doi:https://doi.org/10.1001/jama.2017.21903

This PROtocol template aims to promote implementation and use of the SPIRIT-PRO Extension for trials where PROs are a primary or key secondary outcome; however content is also recommended for use where PROs are exploratory outcomes. This PROtocol template was designed as a stand-alone document which may be incorporated into any clinical trial protocol. When using this addendum please cite as:

Journal to add E&E paper reference please

Nothing in this template should be construed to represent or warrant that persons using this template have complied with all applicable laws and regulations. All individuals and organizations using this template have the responsibility for complying with the applicable laws and regulations or regulatory requirements for the relevant jurisdiction.

We recommend integration of key SPIRIT-PRO information within relevant sections of the protocol (e.g., rationale, schedule of assessments, objectives, endpoints, and statistical analysis). Additional notes from the SPIRIT-PRO group and the industry advisory group (IAG) have also been provided where necessary.

In addition we recommend a separate PRO specific section of the protocol which provides further background information, justification for selection of measures, details on psychometric properties of measures and data collection procedures. The protocol template aims to serve as a guide and sections can be moved to best fit with existing trial templates, however we recommend the use of the SPIRIT-PRO checklist to ensure all content has been covered (Page 18). Efforts should be made by protocol writers to avoid unnecessary repetition of content.

Protocol writers can confirm that they have successfully adhered to the SPIRIT-PRO guideline using the checklist available here:
https://jamanetwork.com/journals/jama/article-abstract/2671472

Protocol writers are encouraged to read and consider other relevant resources which are beyond the scope of the SPIRIT-PRO Extension as detailed below:

References and useful resources:

Evidence-based recommendations for the minimum content of a clinical trial protocol

Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, Hróbjartsson A, Mann H, Dickersin K, Berlin J, Doré C, Parulekar W, Summerskill W, Groves T, Schulz K, Sox H, Rockhold FW, Rennie D, Moher D. SPIRIT 2013 Statement: Defining standard protocol items for clinical trials. *Ann Intern Med* 2013;158:200-207.

Chan A-W, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin J, Dickersin K, Hróbjartsson A, Schulz KF, Parulekar WR, Krleža-Jerić K, Laupacis A, Moher D. SPIRIT 2013 Explanation and Elaboration: Guidance for protocols of clinical trials. *BMJ* 2013;346:e7586.

Enhancing the Quality and Transparency Of health Research (EQUATOR) Network: https://www.equator-network.org/

Estimands

FDA-ASCO Public Workshop: 2019 Clinical Outcome Assessments in Cancer Clinical Trials Fourth Annual Workshop. King-Kallimanis, B. Systematically defining research objectives and framing questions using the estimand framework. https://www.fda.gov/drugs/news-events-human-drugs/fda-asco-public-workshop-2019-clinical-outcome-assessments-cancer-clinical-trials-fourth-annual [accessed 19/12/19]

E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). October 2017. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM582738.pdf

Akacha M, Bretz F, Ruberg S. Estimands in clinical trials - broadening the perspective. Stat Med. 2017;36(1):5-19.

Akacha M, Bretz F, Ohlssen D, Rosenkranz G, Schmidli H. Estimands and Their Role in Clinical Trials. Statistics in Biopharmaceutical Research. 2017;9(3):268-71.

Permutt T. A taxonomy of estimands for regulatory clinical trials with discontinuations. Statistics in Medicine. 2016;35(17):2865-75.

Bell ML, Floden L, Rabe BA, Hudgens S, Dhillon HM, Bray VJ, Vardy JL. Analytical approaches and estimands to take account of missing patient-reported data in longitudinal studies. *Patient Relat Outcome Meas*. 2019;10:129–140. Published 2019 Apr 16. doi:10.2147/PROM.S178963

Patient Reported adverse events

If the trial is assessing patient reported adverse event and symptom monitoring consider how this links to adverse reporting in the protocol.

Kluetz PG, Chingos DT, Basch EM, Mitchell SA. 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. doi: 10.14694/EDBK_159514.

Tolerability

If PROs are being used to assess tolerability this should be described.

Kluetz PG, Kanapuru B, Lemery S, Johnson LL, Fiero MH, Arscott K, Barbachano Y, Basch E, Campbell M, Cappelleri JC, Cella D, Cleeland C, Coens C, Daniels S, Denlinger CS, Fairclough DL, Hillard JR, Minasian L, Mitchell SA, O'Connor D, Patel S, Rubin EH, Ryden A, Soltys K, Sridhara R, Thanarajasingam G, Velikova G, Coons SJ. Informing the Tolerability of Cancer Treatments Using Patient-Reported Outcome Measures: Summary of an FDA and Critical Path Institute Workshop. Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2018;21(6):742-7.

Analysis

The Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data (SISAQOL) Consortium aims to develop recommendations to standardize the analysis of PRO data in cancer randomized controlled trials. https://event.eortc.org/sisaqol/

Coens C, Pe M, Dueck AC, Sloan J, Basch E, Calvert M, Campbell A, Cleeland C, Cocks K, Collette L, Devlin N, Dorme L, Flechtner HH, Gotay C, Griebsch I, Groenvold M, King M, Kluetz PG, Koller M, Malone DC, Martinelli F, Mitchell SA, Musoro J, O'Connor D, Oliver K, Piault-Louis E, Piccart M, Quinten C, Reijneveld JC, Schürmann C, Smith AW, Soltys KM, Taphoorn M, Velikova G, Bottomley A. International standards for the analysis of quality-of-life and patient-reported outcome endpoints in cancer randomised controlled trials: recommendations of the SISAQOL Consortium. The Lancet Oncology 2020;21(2):e83-e96.

Bottomley A, Pe M, Sloan J, Basch E, Bonnetain F, Calvert M, Campbell A, Cleeland C, Cocks K, Collette L, Dueck AC, Devlin N, Flechtner HH, Gotay C, Greimel E, Griebsch I, Groenvold M, Hamel JF, King M, Kluetz PG, Koller M, Malone DC, Martinelli F, Mitchell SA, Moinpour CM, Musoro J, O'Connor D, Oliver K, Piault-Louis E, Piccart M, Pimentel FL, Quinten C, Reijneveld JC, Schürmann C, Smith AW, Soltys KM, Sridhara R, Taphoorn M, Velikova G, Coens C. Moving forward toward standardizing analysis of quality of life data in randomized cancer clinical trials. Clinical trials (London, England). 2018;15(6):624-30.

Pe M, Dorme L, Coens C, Basch E, Calvert M, Campbell A, Cleeland C, Cocks K, Collette L, Dirven L, Dueck AC, Devlin N, Flechtner HH, Gotay C, Griebsch I, Groenvold M, King M, Koller M, Malone DC, Martinelli F, Mitchell SA, Musoro J, O'Connor D, Oliver K, Piault-Louis E, Piccart M, Pimentel FL, Quinten C, Reijneveld JC, Schürmann C, Sloan J, Velikova G, Bottomley. Statistical analysis of patient-reported outcome data in randomised controlled trials of locally advanced and metastatic breast cancer: a systematic review. The Lancet Oncology. 2018;19(9):e459-e69.

Bottomley A, Pe M, Sloan J, Basch E, Bonnetain F, Calvert M, Campbell A, Cleeland C, Cocks K, Collette L, Dueck AC, Devlin N, Flechtner HH, Gotay C, Greimel E, Griebsch I, Groenvold M, Hamel JF, King M, Kluetz PG, Koller M, Malone DC, Martinelli F, Mitchell SA, Moinpour CM, Musoro J, O'Connor D, Oliver K, Piault-Louis E, Piccart M, Pimentel FL, Quinten C, Reijneveld JC, Schürmann C, Smith AW, Soltys KM, Taphoorn M, Velikova G,

Coens C. Analysing data from patient-reported outcome and quality of life endpoints for cancer clinical trials: a start in setting international standards. The Lancet Oncology. 2016;17(11):e510-e4.

Other resources

Cappelleri JC, Zou KH, Bushmakin AG, Alvir JMJ, Alemayehu D, Symonds T. Patient-Reported Outcomes: Measurement, Implementation and Interpretation. Boca Raton, Florida: Chapman & Hall/CRC; 2013.

de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine. Cambridge, UK: Cambridge University Press; 2011.

Fairclough DL. Design and Analysis of Quality of Life Studies in Clinical Trials. 2nd ed. Boca Raton, Florida: Chapman & Hall/CRC; 2010.

Fayers FM, Machin D. Quality of Life: The Assessment, Analysis and Interpretation of Patient-reported Outcomes. 3rd ed. Chichester, England: John Wiley & Sons Ltd.; 2016.

Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A Practical Guide to Their Development and Use. 5th ed. New York, NY: Oxford University Press; 2015.

Missing Data

Mercieca-Bebber R, Palmer MJ, Brundage M, Calvert M, Stockler MR, King MT. 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(6):e010938.

Little RJ, D'Agostino R, Cohen ML, et al. The Prevention and Treatment of Missing Data in Clinical Trials. New England Journal of Medicine 2012;367(14):1355-60.

Little RJ, Cohen ML, Dickersin K, et al. The design and conduct of clinical trials to limit missing data. Stat Med 2012;31(28):3433-43. doi: 10.1002/sim.5519

Patient Public Involvement (PPI) and Patient Experience

Kluetz PG, O'Connor DJ, Soltys K. Incorporating the patient experience into regulatory decision making in the USA, Europe, and Canada. The Lancet Oncology. 2018;19(5):e267-e74.

Table of Contents

Title page

PRO content author:

Affiliation(s), telephone:

Specify the individual(s) responsible for the PRO content of the trial protocol. (SPIRIT-5a-PRO Elaboration)

Explanation: Providing information (e.g. name, affiliation, contact details) on who wrote the PRO-specific aspects of the trial protocol promotes transparency and accountability and identifies the appropriate point of contact for resolution of any PRO specific queries. When patients have actively contributed to this process, this should be documented as per recent guidance for the reporting of patient and public involvement.[7]
Additional notes: *The PRO author should be part of the protocol writing committee. The trial study coordinator should ensure that PROs are harmonized with all the other clinical endpoints.*

1. Protocol Summary

The protocol synopsis is a short (1 to 2 pages) summary of the key points of the protocol (including PRO-specific information). This section of the protocol should be completed after the main text to ensure consistency with the main text.

1.3. Schedule of Activities (SoA)

Include a schedule of PRO assessments, providing a rationale for the time points, and justifying if the initial assessment is not pre-randomization. Specify time windows, whether PRO collection is prior to clinical assessments, and, if using multiple questionnaires, whether order of administration will be standardized (SPIRIT-13-PRO Extension).
Additional notes: *The PRO schedule of assessments should be summarised in a table alongside all other trial assessments to provide trial staff with a single point of reference. Ideally PRO assessments should appear in the table in the order which assessments will take place. If more than one PRO measure is being used, each may be specified on a separate line if they are administered at different time points.*

Procedure	Screening (up to X days before Day 1)	Intervention Period [Days or Weeks, etc.]									E/D *	Follow-up (X days after last dose)	Notes E/D = Early Discontinuation
		–1 (time window)	1 (time window)	2 (time window)	3 (time window)	4 (time window)	5 (time window)	6 (time window)	7 (time window)	8 (time window)			
List PRO assessments consistent with desired order of completion in study visits (more than one line may be necessary). Further details should be provided in Section 4.2. (SPIRIT 13 PRO Extension)													

2. Introduction

2.1. Study Rationale

2.1.1. Summary of PRO-specific rationale

Describe the PRO-specific research question and rationale for PRO assessment and summarize PRO findings in relevant studies. (SPIRIT-6a-PRO Extension)
Additional note: *As with other clinical information, the PRO-specific rationale should be summarized here and detailed information including summary of PRO findings in relevant studies should be provided in Section 4.2. Scientific Rationale for Study Design or in the PRO-specific section 8.1.*

Explanation: Inclusion of PROs in a trial requires careful consideration and planning. A clearly defined question helps with selection of measures and specification of hypotheses and analyses. When the PRO is a secondary or exploratory outcome, a brief rationale may be adequate.[6]

3. Objectives and Endpoints

State specific PRO objectives (including relevant PRO concepts/domains). (SPIRIT-7-PRO Extension)
Additional notes: *The level of detail on the PRO endpoints should also be similar to the other clinical endpoints to ensure harmonization. State whether specific PRO domains will be used for confirmatory or descriptive or exploratory purposes (SISAQOL recommendation statement (RS)1).[4] If primary or secondary objectives are confirmatory objectives (draw conclusions about treatment efficacy), there is a need to specify whether the objective is to show superiority, equivalence or non-inferiority (SISAQOL RS2).[4]*

Explanation: Pre-specification of objectives encourages identification of key PRO domains and time points, reducing the risk of multiple statistical testing and selective reporting of PROs based on statistically significant results.[8]

Specify the PRO *endpoint: the* concepts/domains used to evaluate the intervention (e.g., overall health-related quality of life, specific domain, specific symptom) and, for each one, the analysis metric (e.g., change from baseline, final value, time to event) and the principal time point or period of interest. (SPIRIT-12-PRO Extension)
Additional note: *Utility measures (e.g. EQ-5D) that inform cost-utility evaluations should also be included here.*

Explanation: The PRO concepts/domains and time points for assessment should closely align with the trial objectives and hypotheses. Because of the risk of multiple statistical testing, the domain(s) and principal time point(s) for analyses should be specified *a priori*.[8 9]

Objectives (including PRO objectives)	[Endpoints]
Primary	
1.	
Secondary	
2.	
Tertiary/Exploratory	
3.	

4. Study Design

4.1. Scientific Rationale for Study Design (PRO-specific) [Note this section could also be combined into the PRO-specific section 8.1]

Describe the PRO-specific research question and rationale for PRO assessment and summarize PRO findings in relevant studies. (SPIRIT-6a-PRO Extension)
Additional notes: *Detailed information should be provided here. Information not provided in Section 2.1.1. Summary of PRO-specific rationale should be presented here. Indicate how PRO evaluation aligns with the overall trial design, so it addresses the research objectives.*

Explanation: Inclusion of PROs in a trial requires careful consideration and planning. A clearly defined question helps with selection of measures and specification of hypotheses and analyses. When the PRO is a secondary outcome, a brief rationale may be adequate.[6]

5. Study Population

Specify any PRO-specific eligibility criteria (e.g., language/reading requirements or pre-randomization[1] completion of PRO). (SPIRIT-10-PRO Extension)
If PROs will not be collected from the entire study sample, provide a rationale and describe the method for obtaining the PRO subsample. (SPIRIT-10-PRO Extension)
Additional notes: *Efforts should be made to provide translated PRO versions where needed to promote inclusivity in PRO completion. Specify reasons if translated versions are not provided and certain populations are excluded. PRO-specific eligibility criteria should be included alongside other eligibility criteria to ensure research personnel have a single point of information.*

Explanation: Any PRO-specific eligibility criteria should be considered at the design stage of the trial and clearly specified in the protocol. In large trials, sufficient power may be achieved by collecting PROs from a representative subset of participants, while in some trials it may not be possible to collect PROs in the entire population (e.g., because of non-availability of validated questionnaires in all languages)[10]; in such instances, the rationale for the sampling method should be described.

6. Study Intervention

6.1. PRO capture method involving a medical device

Notes: *Some ePRO systems may meet the requirement for a medical device (for example by providing actionable alerts). If the ePRO fulfils the definition of a medical device, then the protocol should refer to the relevant regulatory directive for that jurisdiction.*

7. Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal from the Study

7.1. Discontinuation of Study Intervention

Describe the process of PRO assessment for participants who discontinue or deviate from the assigned intervention protocol. (SPIRIT-18b (ii)-PRO Elaboration)
Additional note: *Participants who withdraw from the study should be provided pre-paid packaging to return PRO assessments and ePRO devices if applicable.*

Explanation: A clear plan for collection of PROs for trial participants who withdraw early from a study or who discontinue the intervention helps minimize bias,[11] ensures that staff collect all required PRO data in a standardized and timely way, and may assist ethical appraisal of the study.[6]

8. Study Assessments and Procedures

8.1. Efficacy Assessments

8.1.1. Patient Reported Outcomes

Specify the PRO Hypotheses (SPIRIT-7-PRO Extension)
Additional note: *Also specify whether the objective is to support superiority, equivalence or non-inferiority.*

Explanation: Pre-specification of objectives and hypotheses encourages identification of key PRO domains and time points, reducing the risk of multiple statistical testing and selective reporting of PROs based on statistically significant results.[8]

Justify the PRO instrument to be used and describe domains, number of items, recall period, instrument scaling and scoring (e.g., 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. (SPIRIT-18a (i)-PRO Extension).
Additional note: *Interpretation guidelines should include reference to pre-defined clinical relevance thresholds, preferably specific to the PRO population.*

Explanation: The selection of PROs to be used in a clinical trial requires careful consideration. Ideally, the measure should be validated in the target population.[12] Consideration should be given to the number of questionnaires to be used, acceptability of the questions, and the likely patient burden (e.g., time taken for completion, cognitive burden, emotional burden). Justification for the measures selected will help trial personnel understand why specific measures are being used.[13] Questionnaires should be used in accordance with any existing user manuals to promote data quality and ensure standardized scoring, and any deviations should be described.[6]

Include a schedule of PRO assessments, providing a rationale for the time points, and justifying if the initial assessment is not pre-randomization. Specify time windows, whether PRO collection is prior to clinical assessments, and, if using multiple questionnaires, whether order of administration will be standardized. (SPIRIT-13-PRO Extension)

Additional notes: *State the mode of administration (e.g., paper and/or electronic capture). Baseline PROs should be completed prior to randomization allocation. The windowing for the first on treatment assessment should ensure the recall period does not overlap with day 0 of the trial (e.g., if recall is 7 days the minimum window should be trial day 7). The order of PRO administration should be the same for the entire duration of the study. Specify when patients are still expected to complete a PRO assessment (e.g., whether PROs will be collected after disease progression or not). This will allow calculation of completion rates (SISAQOL recommendation statement (RS) 19-20).[4] This information can be summarised in section 1.3 with further detail provided in this PRO-specific section.*

Explanation: Provision of an easy-to-follow schedule will assist staff and may help reduce missing data.[1] Collecting PRO data prior to randomization helps ensure an unbiased baseline assessment, and if specified as an eligibility criterion, ensures data completeness. This is important because baseline PRO data are often used as a covariate in analyses and are essential to calculating change from baseline. Completion of PROs prior to clinical assessments (as these may influence patient responses) and standardization of the order of questionnaire administration are advised to help reduce measurement error.[14] Allowable time windows for each scheduled PRO assessment should be specified to ensure that PRO data collection captures the effect of the clinical event(s) of interest.[6]

Include a data collection plan outlining the permitted mode(s) of administration (e.g., paper, telephone, electronic, other) and setting (e.g., clinic, home, other). (SPIRIT-18a (ii)-PRO Extension)

Additional notes: *Specify whether participants would be completing assessments at the site. If so, they should be completed prior to any other assessments, procedures, or discussions with their care team. Participants should also be given time and a quiet space to complete the assessments. If electronic PRO collection tools are unavailable, damaged or lost, describe procedure to follow for the collection of PROs according to the SoA.*

Explanation: It is important that both research personnel and trial participants understand how, when, and where PRO data will be collected in the study. Increasingly, electronic PRO assessment is undertaken in trials, so evidence of equivalence between different modes of administration should be considered.[15] If electronic PRO measures contain only minor modifications with respect to the paper-based versions, usability testing and cognitive debriefing may provide sufficient evidence of equivalence.[15 16] The setting for PRO data collection should be described and standardized across trial intervention groups and sites.[6]

Specify PRO data collection and management strategies for minimizing avoidable missing data. (SPIRIT-18b (i)-PRO Extension)
Additional notes: *Signpost to statistical analysis plan (SAP) as required. Collect reasons for missed PRO assessments during the trial. These could be documented in the case report forms (CRF). SISAQOL, CONSORT-PRO Extension and PROTEUS may be consulted for further guidance on missing data.*[2 4 5]

Explanation: Missing data are a particular problem for PROs because participants with the poorest outcomes in a trial often are those who do not complete planned PRO assessments, and data cannot be obtained retrospectively beyond the time frame of interest or from medical records. This is a potentially significant source of bias and may reduce trial power.[17] It is important to note that not all missing PRO data are avoidable: participants have the right to decide not to complete questionnaires. Common reasons for avoidable missing PRO data are administrative errors, lack of explanation of the importance of PRO data, and overly burdensome questionnaires. Addressing these in the protocol should help minimize avoidable missing data. Examples of protocol content include ensuring that PRO end points and hypotheses are clearly defined and scientifically compelling, providing a rationale for PRO assessment, clearly specifying the PRO assessment time points, defining acceptable PRO assessment time windows, aligning PRO assessment time points to clinic visits (if clinically informative), minimizing participant burden, and specifying the importance of complete PRO data.[1]

Specify whether more than one language version will be used and state whether translated versions have been developed using currently recommended methods. (SPIRIT-18a (iii)-PRO Extension)

Explanation: Multinational trials, or national trials involving participants with different languages, require measures that have been translated and culturally adapted where needed using appropriate methodology.[18 19] This may influence the selection of measure to be used because inclusion of a wide range of participants can help ensure the generalizability of trial results. Plans to use translated versions should be specified in the protocol, citing references when available.[6]

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. (SPIRIT-18a (iv)-PRO Extension)
Additional note: *If a proxy reporting is used, this should ideally be the same person throughout the trial.*

Explanation: In some contexts, such as trials involving young children or cognitively impaired participants, it may be necessary for someone other than a trial participant to respond on that participant's behalf. Note that proxy is different from someone assisting a person to respond to questionnaires (e.g., a nurse reads questions to a patient and writes

down their actual answers). Clear justification and specification of proxy reporting in the protocol allows external reviewers to assess potential bias and facilitates trial reporting in accordance with CONSORT-PRO.[2]

8.2. Safety Assessments

8.2.1 Monitoring of PRO data

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 standardized way. Describe how this process will be explained to participants; e.g., in the participant information sheet and consent form. (SPIRIT-22-PRO Extension)
Additional note: *If PROs are used in place of or as an add-on to standard solicited AE recording (at visits/phone calls the investigators asks if there have been any problems since last visit) they should be recognised as part of the safety monitoring procedures of the trial.*

Explanation: Evidence suggests that monitoring and management of PRO alerts (psychological distress or physical symptoms evident from PRO responses that may require an immediate response) vary across and within trials.[13 18 20] To protect the interests of trial participants and minimize potential bias, it is important to specify plans for monitoring.[21] If monitoring is not planned (e.g., in a low-risk study in which alerts are not anticipated), this should also be briefly stated in the protocol, the participant information sheet, and the consent form. Alternative support mechanisms for participants should be outlined.[6]

9. Statistical considerations

When a PRO is the primary end point, 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 end point, then discuss the power of the principal PRO analyses. (SPIRIT-14-PRO Elaboration)
Additional note: *Enough information on secondary, tertiary and even exploratory PRO endpoints should be provided to justify their inclusion.*

Explanation: In studies in which PROs are the primary outcome or end point, the target sample size will generally be based on an a priori sample size calculation for that end point.[9] Ideally, the criteria for clinical significance (e.g., minimal important difference, responder definition) should be specified when known.[22 23] If PROs are a secondary end point, researchers should specify whether the sample size provides sufficient power to test the principal PRO hypotheses.[9]

9.1. Statistical analyses

State PRO analysis methods, including any plans for addressing multiplicity/type I (α) error. (SPIRIT-20a-PRO Elaboration)
Additional notes: *Also state if there are no plans to address multiplicity/type I (α) error. Pre-define clinical relevance thresholds for the domains (SISAQOL recommendation statement (RS) 3-5).[4] Methods should align with specified endpoints and reflect superiority, equivalence or non-inferiority objectives (SISAQOL RS 2).[4] A power analysis can be conducted to assess whether a clinical relevant difference as specified in the objective can be detected reliably with the given sample size for the PRO population (SISAQOL RS 2).[4] If sensitivity analysis is planned, this should be pre-specified (SISAQOL RS 32).[4]*

Explanation: Many questionnaires, such as health-related quality-of-life measures, are multidimensional and therefore may yield several summary scores (e.g., multiple domains and an overall score). Furthermore, PROs are usually assessed at multiple time points. Statistical analysis of all domains and time points implies multiple hypothesis testing, which inflates the probability of false-positive results (type I error).[9] This can be contained by pre-specifying the key PRO domain(s) or overall score of interest and the principal time point(s). Any plans to address multiplicity, such as stepwise or sequential analyses, whereby multiple end points are tested in a defined sequence that contains the overall type I error to the desired level, or conventional non-hierarchical methods (e.g., Bonferroni correction), should be specified a priori.[8] The protocol should either fully address these issues or provide a summary with reference to where full details can be found (e.g., in the statistical analysis plan).[6]

9.1.1. Missing PRO data

State how missing data will be described and outline the methods for handling missing items or entire assessments (e.g., approach to imputation and sensitivity analyses). (SPIRIT-20c-PRO Elaboration)
Additional notes: *State whether PRO data is missing at the item level or at entire PRO assessment level and whether intermittent or lost to follow-up. In addition, state how missing PRO data is recorded and categorised. SISAQOL recommendations may be consulted for further guidance on missing data.[4]*

Explanation: There are 2 levels of missing PRO data: (1) patient completion of some but not all items within an instrument and (2) absence of the entire PRO assessment. Whether and how missing items should be imputed is usually specified in an instrument's scoring algorithm. When entire PRO assessments are missed, analysis requires assumptions about why those data were missing (i.e., the missing data mechanism). There are a range of statistical approaches, each with specific assumptions. Inappropriate method selection may lead to potentially biased and misleading results.[1 11] The protocol should acknowledge and summarize these issues, with full details provided in the statistical analysis plan.[6]

9.1.2. Other Analyse(s)

Note: *If psychometric analyses of the PRO measure(s) or any additional research questions are planned with the PRO data collected, this should be acknowledged in the protocol, with reference to statistical analysis plan (SAP), Calvert et al 2018 (JAMA Supplement 3).*[6]

10. Supporting Documentation and Operational Considerations

10.1. Appendix 1: Regulatory, Ethical, and Study Oversight Considerations

10.1.1. Publication Policy

Notes: *It is important to publish PRO data. There is evidence that this is not done enough or well. Publication of PRO data should be done according to CONSORT and CONSORT-PRO*

Note: *Whilst not a formal recommendation in the SPIRIT-PRO guidance, it is regarded as good practice to include, depending on the mode of administration, paper copies of PRO or screenshots of ePRO instruments.*
All copyrights and version information should be clearly showing in the appended PRO instruments with due acknowledgment for PROs requiring permission for use.

10.2. Appendix Copies of PROs

Glossary

Concept: "The specific measurement goal (i.e., 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."[24]
Domain: "A sub concept represented by a score of an instrument that measures a larger concept comprised of multiple domains. For example, psychological function is a larger concept with multiple domains (emotional and cognitive function) that are measured by relevant items."[24]
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 combined[25] (e.g. change from baseline at 6 weeks in mean fatigue score).[26]
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"[27]
Important or key secondary PROs/end points: Some PRO measures (particularly health-related quality-of-life measures) are multidimensional, producing several domain-specific outcome scales; e.g., 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 (CONSORT PRO Extension 2b),[27] 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 (e.g. using the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE™))[28]
Instrument: "A means to capture data (e.g., 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."[24]
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.[29]
Item: "an individual question, statement, or task (and its standardized response options) that is evaluated by the patient to address a particular concept"[24]
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 health care professional".[30]
Outcome*: the variable to be measured. It is the measurable characteristic that is influenced or affected by an individuals' baseline state or an intervention as in a clinical trial or other exposure[25] (e.g. a fatigue score).
Patient-reported outcome (PRO): 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.[24 27] PROs are assessed by self-reported questionnaires, referred to as PRO measures (PROMs) or instruments.[25]
Primary outcome: the most important outcome in a trial, pre-specified 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"[24]
Secondary outcomes: outcomes pre-specified in the protocol to assess additional effects of the intervention; some PROs may be identified as important or key secondary outcomes
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials.[31 32]
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 time point whereby the result would still be deemed to be clinically relevant.[33]

* 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 (i.e. the outcome) by fully specifying four components: measurement variable (e.g. fatigue "in the past week" as measured by the QLQ-C30), analysis metric (e.g., change in fatigue from baseline, final fatigue value, time to clinically important increase in fatigue (and "event"), method of aggregation (e.g., median fatigue, proportion of patients with severe fatigue, proportion of patients with clinically important change in fatigue), and time point. Note that using theses definitions, several endpoints can be defined from the same outcome source data, revealing the distinction and relationship between "outcome" and "endpoint" for PROs.

List of Abbreviations

CONSORT – Consolidated Standards of Reporting Trials

CRF – case report form

ePRO – electronic patient-reported outcome

PPI – Patient Public Involvement

PRO – patient-reported outcome

PRO-CTCAE – Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events

PROTEUS – Patient-Reported Outcomes Tools: Engaging Users & Stakeholders

RS – recommendation statement

SAP – statistical analysis plan

SISAQOL – Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data

SoA – Schedule of Activities

SPIRIT – Standard Protocol Items: Recommendations for Interventional Trials

References

1. 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(6):e010938. doi: 10.1136/bmjopen-2015-010938 [published Online First: 2016/06/18]
2. Calvert M, Blazeby J, Altman DG, et al. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *Jama* 2013;309(8):814-22. doi: 10.1001/jama.2013.879 [published Online First: 2013/02/28]
3. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Annals of internal medicine* 2013;158(3):200-7. doi: 10.7326/0003-4819-158-3-201302050-00583 [published Online First: 2013/01/09]
4. 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. *The Lancet Oncology* 2020;21(2):e83-e96. doi: 10.1016/S1470-2045(19)30790-9
5. Snyder C, Smith K, Holzner B, et al. Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data. *Quality of Life Research* 2019;28(2):345-56. doi: 10.1007/s11136-018-2020-3
6. Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for Inclusion of Patient-Reported Outcomes in Clinical Trial Protocols: The SPIRIT-PRO Extension. *Jama* 2018;319(5):483-94. doi: 10.1001/jama.2017.21903
7. 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. doi: 10.1136/bmj.j3453 [published Online First: 2017/08/05]
8. FDA. Guidance for industry. Patient-reported outcome measures: use in medicinal product development to support labeling claims. Silver Spring, MD: US Department of Health and Human Services Food and Drug Administration, 2009.
9. 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. doi: 10.1111/j.1524-4733.2007.00273.x [published Online First: 2007/11/13]
10. Kyte D, Retzer A, Ahmed K, et al. Systematic Evaluation of Patient-Reported Outcome Protocol Content and Reporting in Cancer Trials. *JNCI: Journal of the National Cancer Institute* 2019;111(11):1170-78. doi: 10.1093/jnci/djz038
11. Fairclough DL. Design and analysis of quality of life studies in clinical trials. New York: Chapman & Hall/CRC Press 2002.
12. Rothrock NE, Kaiser KA, Cella D. Developing a valid patient-reported outcome measure. *Clin Pharmacol Ther* 2011;90(5):737-42. doi: 10.1038/clpt.2011.195 [published Online First: 2011/10/05]
13. 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(10):e76625. doi: 10.1371/journal.pone.0076625 [published Online First: 2013/10/15]
14. Leidy NK, Revicki DA, Geneste B. Recommendations for evaluating the validity of quality of life claims for labeling and promotion. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 1999;2(2):113-27. doi: 10.1046/j.1524-4733.1999.02210.x [published Online First: 2006/05/06]
15. 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 in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 2009;12(4):419-29. doi: 10.1111/j.1524-4733.2008.00470.x [published Online First: 2009/11/11]

16. Rutherford C, Costa D, Mercieca-Bebber R, et al. Mode of administration does not cause bias in patient-reported outcome results: a meta-analysis. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2016;25(3):559-74. doi: 10.1007/s11136-015-1110-8 [published Online First: 2015/09/04]
17. Bernhard J, Cella DF, Coates AS, et al. Missing quality of life data in cancer clinical trials: serious problems and challenges. *Statistics in Medicine* 1998;17(5-7):517-32. doi: 10.1002/(sici)1097-0258(19980315/15)17:5/7<517::Aid-sim799>3.0.Co;2-s
18. 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. doi: 10.1016/j.conctc.2017.11.009 [published Online First: 2018/04/27]
19. 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 in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 2005;8(2):94-104. doi: 10.1111/j.1524-4733.2005.04054.x [published Online First: 2005/04/05]
20. 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(1):e0144658. doi: 10.1371/journal.pone.0144658 [published Online First: 2016/01/20]
21. Kyte D, Draper H, Calvert M. Patient-reported outcome alerts: ethical and logistical considerations in clinical trials. *Jama* 2013;310(12):1229-30. doi: 10.1001/jama.2013.277222 [published Online First: 2013/09/26]
22. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of clinical epidemiology* 2008;61(2):102-9. doi: 10.1016/j.jclinepi.2007.03.012 [published Online First: 2008/01/08]
23. King MT. A point of minimal important difference (MID): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res* 2011;11(2):171-84. doi: 10.1586/erp.11.9 [published Online First: 2011/04/12]
24. FDA. Guidance for Industry, Patient-Reported Outcome Measures: Use in Medical Product Development to Suport Labeling Claims 2009 [Available from: https://www.fda.gov/media/77832/download accessed 25/11/2019.
25. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource 2016 [Available from: https://www.ncbi.nlm.nih.gov/books/NBK326791/.
26. Curtis L HA, Weinfurt K. Choosing and Specifying Endpoints and Outcomes: Introduction. In: Rethinking Clinical Trials: A Living Textbook of Pragmatic Clinical Trials. 2019 [Available from: DOI: 10.28929/009 accessed 12/03/2020.
27. Calvert M, Blazeby J, Altman DG, et al. Reporting of Patient-Reported Outcomes in Randomized Trials: The CONSORT PRO Extension. *JAMA* 2013;309(8):814-22. doi: 10.1001/jama.2013.879
28. Kluetz PG, Chingos DT, Basch EM, et al. 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. doi: 10.14694/edbk_159514 [published Online First: 2016/06/02]
29. ClinicalTrials.gov. Glossary of Common Site Terms [Available from: https://clinicaltrials.gov/ct2/about-studies/glossary accessed 12/03/2020 2020.
30. FDA. Drug Development Tool Glossary [updated 08/15/2017. Available from: https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/ddt-glossary accessed 03/07/2020 2020.
31. Chan A-W, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. *Annals of Internal Medicine* 2013;158(3):200-07. doi: 10.7326/0003-4819-158-3-201302050-00583

32. Chan A-W, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;8(346) doi: 10.1136/bmj.e7586
33. 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. *European Journal of Cancer* 1997;33(1):20-28. doi: https://doi.org/10.1016/S0959-8049(96)00412-1

Supplement 1

Protocol evidence per final SPIRIT-PRO item.

SPIRIT PRO Item number and description	SPIRIT-PRO wording prior to finalisation	% HTA protocols including item	% EPiC protocols including item	% international ovarian cancer protocols including item
(5a) Specify the individual(s) responsible for the PRO content of the trial protocol		6.67%	22.78%	23.1%
(6a) Describe the PRO-specific research question and rationale for PRO assessment and summarise PRO findings in relevant studies	Describe what is currently known about PROs in this area and explain the gaps in the literature	49.33%	32.91%	42.3%
	Provide a rationale for the inclusion of PROs as appropriate to the study population, intervention, context, objectives and setting	8.00%	33.54%	57.7%
(7) State specific PRO objectives or hypotheses (including relevant PRO concepts/domains	State the PRO study objective in relation to PRO domain/s, patient population and timeframe	77.33%	73.42% (17.09% in relation to dimension, population or timeframe)	30.8%
	State the PRO hypothesis and corresponding null hypothesis and to which outcome(s) the hypothesis relates	18.67%	-	PRO hypothesis provided 19.2%
(10) Specify any PRO-specific eligibility criteria (e.g. 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 methods for obtaining the PRO subsample	If PROs will be collected in a subset of the study population or in specific centres, include a description/rationale for the sampling method	0.00%	10.76%	11.5%
	State the inclusion/exclusion criteria for PRO endpoint(s) (e.g., language/reading requirements)	45.33%	50.00%	7.7%

	Specify if PRO completion is pre-randomisation eligibility requirement	-	-	7.7%
(12) Specify the PRO concepts/domains used to evaluate the intervention (e.g. overall health-related quality of life, specific domain, specific symptom) and for each one, the analysis metric (e.g. change from baseline, final value, time to event) and the principal time point or period of interest	Describe the PRO constructs used to evaluate the intervention e.g. overall QOL, specific domain, specific symptom	-	-	73.1%
(13) Include a schedule of PRO assessments, providing a rationale for the time points, 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	Specify the timepoint(s) for PRO analysis (including the principle timepoint of interest) and provide the rationale for these	Timing specified 97.33%	Timing specified 83.54%	42.3%
	Include PRO assessments in the main protocol schedule of assessments, specifying which PRO measures (PROMs) will be used at each assessment	-	-	96.2%
	Specify if baseline PRO assessment should be completed before randomisation	-	-	53.8%
	Specify the targeted time and acceptable time windows for each PRO assessment	-	-	26.9%
	If PROs are to be completed in the clinic: specify timing of PROM delivery in relation to clinical assessments (e.g. before/whilst/after seeing clinician and/or clinical assessments)	-	-	54%
	Justify the timing of PRO assessments. Scheduled	Timing justified 6.67%	Timing justified 12.03%	23.1%

	PRO assessments should link to research questions, hypotheses, length of recall, disease/treatment natural history, planned analysis and time of comparison must be comparable for both arms			
(14) When a PRO is the primary end point, 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 end point, then discuss the power of the principal PRO analyses	If PRO is the primary endpoint, state the required PRO sample size, otherwise discuss the power of the PRO analysis	50.67%	25.95%	30.8%
(18a i) Justify the PRO instrument to be used and describe the domains, number of items, recall period, and instrument scaling and scoring (e.g. 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	Describe the PROMs including, number of items/domains, instrument scaling/scoring, reliability, content and construct validity, responsiveness, sensitivity, acceptability, recall period. Provide references as appropriate	PROM identified 100%; Justification in relation to study hypotheses 41.33%; Justified in relation to measurement properties 37.33%; Justified in relation to acceptability/patient burden 14.67%	PROM identified 63.29%; Justification in relation to study hypotheses 36.71%; Justified in relation to measurement properties 46.84%; Justified in relation to acceptability/patient burden 29.11%	Justification for measure used 84.6%
(18a ii) Include a data collection plan outlining the permitted mode(s) of administration (e.g. paper, telephone, electronic, other)	Include a pre-specified data collection plan	84.00% included brief details of PRO data collection procedures but often omitted information surrounding mode of administration,	57.59%	46.2%

and setting (e.g. clinic, home, other)		setting and proxy reporting: 8.00% included PRO data collection guidelines/training information for trial personnel.		
(18a iii) Specify whether more than one language version will be used and state whether translated versions have been developed using currently recommended methods	Provide evidence of measurement equivalence across modes (i.e., when mixing modes of PRO data collection) and/or of cross cultural validity where different language versions of questionnaires are used	-	-	7.7%
(18a iv) 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		-	-	0.0%
(18b i) Specify PRO data collection and management strategies for minimising avoidable missing data	Specify procedures for data collection and management methods to minimise missing data. E.g. checking completed PROMs (including who will check forms and how will they deal with missing PROMs or missing items).	-	-	38.5%
(18b ii) Describe the process of PRO assessment for participants who discontinue or deviate from the assigned intervention protocol	Establish process for PRO assessment at (and beyond) withdrawal for patients who withdraw early from a study or who go 'off-study'/'off treatment'	-	-	34.6%
(20a) State PRO analysis methods, including any plans for addressing multiplicity/type I (a) error	Include an a priori description of all planned PRO analyses pertaining to the study hypotheses	PRO statistical analysis plan provided? 96.00%	PRO statistical analysis plan provided? 53.16%	61.5%

	Pre-specify sequence of testing/exploratory analyses to control for multiplicity or pre-specify domains (e.g. in a regulatory trial/labelling claim) (Common in pharma trials. Involves pre-specifying domains that alpha would be spent on, or ordering the domains in priority & alpha would be spent down the list)	Plans to address multiplicity of PRO data provided? 1.33%	10.13%	7.7%
(20c) State how missing data will be described and outline the methods for handling missing items or entire assessments (e.g. approach to imputation and sensitivity analyses)	State how missing data will be described	45.33%	30.38%	-
	Describe method for handling missing assessments (e.g. approach to imputation and sensitivity analyses)	45.33%	30.38%	-
	Describe methods for scoring endpoints. Where possible, reference scoring manuals for summated scales from PROM (domain-specific &/or total) & methods for handling missing items, and methodological papers for composite endpoints (e.g. QTWiST)	-	-	53.8%
(22) 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; e.g. in the participant information sheet and consent form	Include an a priori plan for consistent/standardised management of PRO alerts (symptoms reported by patients that exceed a pre-defined level of severity) to be clearly communicated to all appropriate trial staff	10.67%	0.63%	0.0%
	Specify whether PRO forms will be used to	4.00%	3.80%	7.7%

	influence therapy or patient management (i.e. will the clinician use PRO responses to inform the patient's care?)			

Supplement 2

Summary recommendations for inclusion* of PRO specific information in the trial protocol, supplementary documents or training for staff.

SPIRIT PRO Item Description	SPIRIT 2013 / SPIRIT- PRO Item Number	Guidance/training for trial staff (eg, site initiation/face to face or online training/operations manual)	Information/ guidance for participants (eg, participant information sheet)	Statistical Analysis Plan
Administrative information				
Specify the individual(s) responsible for the PRO content of the trial protocol	✓ SPIRIT-5a-PRO Elaboration	+		
Introduction				
Describe the PRO specific research question and rationale for PRO assessment, and summarize PRO findings in relevant studies.	✓ SPIRIT-6a-PRO Extension	✓	✓[1]	
State specific PRO objectives or hypotheses (including relevant PRO concepts/domains).	✓ SPIRIT-7-PRO Extension			✓
Methods: Participants, interventions, and outcomes				
Specify any PRO-specific eligibility criteria (eg, language/reading requirements or pre-randomization completion of PRO). If PROs will not be collected in the entire study sample, provide a rationale and describe the method for obtaining the PRO subsample.	✓ SPIRIT-10-PRO Extension	✓		✓
Identify the PRO endpoint as the primary, secondary (and if so - whether a key/important secondary), or an exploratory endpoint.	✓ SPIRIT-12			✓
Specify the PRO concepts/domains used to evaluate the intervention (eg, overall HRQOL, specific domain, specific symptom) and, for each one, the analysis metric (eg, change from baseline, final value, time to event) and the principal time point or period of interest.	✓ SPIRIT-12-PRO Extension			
Include a schedule of PRO assessments, providing a rationale for the time points, and justifying if the initial assessment is not pre-randomization. Specify: time windows; whether PRO collection is prior to clinical assessments; and if using multiple questionnaires, whether order of administration will be standardized.	✓ SPIRIT-13- PRO Extension	✓		✓
Where 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 PRO endpoint, then discuss the power of the principal PRO analyses.	✓ SPIRIT-14-PRO Elaboration			✓
Methods: Data collection, management, and analysis				

Justify the PRO instrument to be used, and describe domains, number of items, recall period, instrument scaling/scoring (eg, range and direction of scores indicating a good/poor outcome). Evidence of PRO instrument measurement properties, interpretation guidelines, and patient acceptability/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.	✓ SPIRIT-18a(i)-PRO Extension			✓[2]
Provide evidence of measurement equivalence across modes (i.e. when mixing modes of PRO data collection) and/or of cross cultural validity where different language versions of questionnaires are used.				✓[3]
Outline plans for evaluation of measurement properties, if appropriate (eg, if not previously validated in the population of interest).				✓[4]
Specify the estimated time to complete each assessment, and discuss feasibility of assessment for the population.		✓	✓[5]	
Include a data collection plan outlining the permitted mode(s) of administration (eg, paper, telephone, electronic, other) and setting (eg, clinic, home, other).	✓ SPIRIT-18a(ii)-PRO Extension	✓	✓	
Specify whether more than one language version will be used, and state whether translated versions have been developed using currently recommended methods.	✓ SPIRIT-18a(iii)-PRO Extension	✓		
Where the trial context requires someone other than the trial participant to answer on their behalf (a proxy reported outcome), state and justify this. Provide/cite evidence of the validity of proxy assessment if available.	✓ SPIRIT-18a(iv)-PRO Extension	✓		
Specify who will administer the PROM (eg, a physician, nurse etc).		✓		
If it is permissible for another person to help the study participant complete the PROM, describe what type and level of assistance is acceptable.		✓	✓	
Include a plan for systematically training and contacting local site personnel to ensure that they understand the content and importance of collecting PRO data. Ideally coordinated by a lead data manager who monitors PRO completion rates in real time and communicates with sites if completion rates are suboptimal.		✓		
Specify PRO data collection and management strategies for minimising avoidable missing data.	✓ SPIRIT-18b(i)-PRO Extension	✓		
Include guidance on discussing importance of PROs with patient		✓		
Describe the process of PRO assessment for participants who discontinue or deviate from their assigned intervention protocol.	✓ SPIRIT-18b(ii)-PRO Elaboration	✓		
Specify that a named person/position at each centre (and/or centrally) be nominated to take responsibility for administration, collection and checking of PROM, specify whether this is the treating clinician or not.		✓[6]		
Specify how an electronic PRO system/database will be maintained and how the investigator will meet regulatory requirements and ensure data integrity and security.	✓ SPIRIT-19	✓	✓[7]	
Specify plan to monitor PRO compliance, including adherence to time windows.	✓ SPIRIT-19	✓	✓[8]	
Include an overview of PRO administration (data collection), and data handling/transmission and storage procedures	✓ SPIRIT-19	✓	✓[7]	

Item 39: Include an a priori description of all planned PRO analyses pertaining to the study hypotheses. Item 44: Include a priori identified summary statistics (as appropriate).	✓ SPIRIT-20a			✓
State the assumptions of PRO analyses.				✓
Include an a priori estimation of PRO effect size.	✓ SPIRIT-14			✓
Specify intention-to-treat or per-protocol PRO analyses.	✓ SPIRIT-20c			✓
Specify the minimum PRO response rate and acceptable degree of timing deviation (i.e. acceptable time windows for each PRO assessment time point) before the PRO objective is compromised. Specify the minimum PRO response rate and acceptable degree of timing deviation (i.e. acceptable time windows for each PRO assessment time point) before the PRO objective is compromised.	✓ SPIRIT-14			✓
Describe methods for scoring endpoints. Where possible, reference scoring manuals for summated scales from PROM (domain-specific and/or total) and methods for handling missing items, and methodological papers for composite endpoints (eg, QTWiST).	✓ SPIRIT-20a			✓
State PRO analysis methods including any plans for addressing multiplicity/type 1 (α) error.	✓ SPIRIT-20a-PRO Elaboration			✓
Specify the criteria for clinical significance (eg, state minimal [clinical] important difference and/or responder definition (size and duration of benefit)).	✓ SPIRIT-14			✓
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).	✓ SPIRIT-20c PRO Elaboration			✓
Monitoring				
Describe the role of the Data Monitoring Committee and Quality Assurance for PROs.	✓ SPIRIT-21a			
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 standardized way. Describe how this process will be explained to participants, eg, in the participant information sheet and consent form.	✓ SPIRIT-22 PRO Extension	✓	✓[7]	
Ethics and dissemination				
Describe informed consent procedure for PRO assessment.		✓		
Include detailed plans for regular feedback to participants via letter/newsletter on PRO aspect of study.		✓	✓	

Footnotes:*"Recommended for inclusion" indicates that >50% of the Delphi and Stakeholder Survey participants endorsed the item (or at least one item if two items were merged for the consensus meeting) for inclusion and/or the item was considered important for inclusion at the consensus meeting.
+Recommended by the Delphi panel but excluded following discussion at the consensus meeting.
[1] <50% of Delphi panellists voted to include this item, but item was included following discussion at the consensus meeting. It was felt important to include a brief rationale for PRO assessment in the PIS.
[2] Use of the PRO instrument in accordance with the user manual should be specified in the SAP.
[3] Plans for dealing with different modes of administration should be specified in the SAP. This was not supported by the Delphi but was identified as important by the consensus panel.
[4] If validation of the PRO instrument is planned as part of the trial then this should be pre-specified in the main trial SAP. If validation is planned as a separate sub-study, this should be specified a separate study protocol and SAP.
[5] Estimation of time to complete each PRO assessment should be included in guidance and training for staff and in information for trial participants.
[6]This should also be recorded in the delegation of duties log.
[7] The participant information sheet should contain information regarding the storage and security of PRO data and provide information on who will access their PRO data and for what purpose.
[8] If plans to monitor adherence to time windows, include reminders for participants eg. Via text or SMS, the relevant details should be specified in the information to participants.