SISAQOL-IMI consensus-based guidelines to design, analyse, interpret and present patient-reported outcomes in cancer clinical trials

Authors

Cecilie D. Amdal[1], Ragnhild S. Falk[1], Ahu Alanya[2], Michael Schlichting[3], Satrajit Roychoudhury[4], Vishal Bhatnagar[5], Lisa M. Wintner[6], Antoine Regnault[7], Anders Ingelgård[8], Corneel Coens[2], Saskia le Cessie[9], Bernhard Holzner[6], Jane Chang[4], Martin Taphoorn[9], Paul Cislo[4], Johannes M. Giesinger[6], Joseph C. Cappelleri[4], Vivek Pawar[10], Silene ten Seldam[11], Elektra J. Papadopoulos[12], Melanie J. Calvert[13], Kenth L. Joseph[1], Andrew Bottomley[2], Ingolf Griebsch[8], Juan I Arraras[14], Guro L Astrup[1], Ethan Basch[15], Andrej Belančić[16], Michael Brundage[17], Alicyn Campbell[18], Khadija Rerhou Rantell[19], Kim Cocks[20], Nathan Cherny[21], Sonya Eremenco[22], Montse Ferrer[23], Mallorie H. Fiero[5], Christoph Gerlinger[24], Els Goetghebeur[25], Ulrich Grouven[26], Alexandra Lauer[8], Olalekan Lee Aiyegbusi[13], Abigirl Machingura[2], Junki Mizusawa[27], Geert Molenberghs[28], Morten Aa. Petersen[29], Jaap C. Reijneveld[30], Jolie Ringash[31], Gerhard Rumpold[32], Claudia Rutherford[33], Chantal Quinten[34], Kavita Sail[12], Maxime Sasseville[35], Willi Sauerbrei[36], Anja Schiel[37], Ashley W. Smith[38], Claire Snyder[39], Galina Velikova[40], Xin S. Wang[41], Kristin Bjordal*[1] and Madeline Pe*[2]on behalf of the SISAQOL-IMI Consortium

*Shared last authorship

Affiliations

[1] Oslo University Hospital, Oslo, Norway (C D Amdal MD PhD, R S Falk PhD, Kenth L. Joseph PhD, GL Astrup PhD, Prof K Bjordal MD PhD)

[2] European Organisation for Research and Treatment of Cancer (EORTC) Headquarters, Brussels, Belgium (A. Alanya PhD, C Coens MSc, *A Bottomley PhD (former)*, A Machingura MSc, M. Pe PhD)

[3] Merck, Healthcare KGaA, Darmstadt, Germany (M Schlichting MSc)

[4] Pfizer, New York, NY, USA (S Roychoudhury PhD, J Chang MPH, P Cislo PhD, J C Cappelleri PhD)

[5] US Food and Drug Administration, Silver Spring, MD, USA, USA (V Bhatnagar MD, M Fiero PhD)

[6] University Hospital of Psychiatry II, Medical University of Innsbruck, Innsbruck, (L M Wintner PhD, Prof. B Holzner PhD, Prof. J M Giesinger PhD)

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

[8] Boehringer Ingelheim, Ingelheim am Rhein, Germany (A Ingelgård PhD, *I Griebsch (former),* A Lauer PhD)

[9] Department of Clinical Epidemiology (Prof S Le Cessie PhD), Department of Neurology (Prof M J B Taphoorn MD PhD), Leiden University Medical Center, Leiden, the Netherlands

[10] EMD Serono Inc., Billerica, MA, USA (V Pawar PhD)

[11]Myeloma Patients Europe, Brussels, Belgium (S ten Seldam MSc)

[12] AbbVie, Mettawa, IL, USA (EJ Papadopoulos MD MPH, K Sail PhD)

[13] Centre for Patient Reported Outcomes Research and NIHR Biomedical Research Centre, University of Birmingham, Birmingham, UK (Prof M Calvert PhD, Prof O L Aiyegbusi PhD)

[14] Hospital Universitario de Navarra, Pamplona, Spain (J Arraras PhD)

[15] American Society for Clinical Oncology, USA (E Basch)

[16] Clinical Hospital Center Rijeka, Rijeka, Croatia (A Belančić MD)

[17] Queen's University at Kingston, Ontario, Canada (M Brundage MD MSc)

[18] Patient Relevant Evidence, San Francisco, CA, USA (A Campbell MPH)

[19] Medicines and Healthcare products Regulatory Agency, London, UK (K Rerhou Rantell PhD)

[20] Adelphi Values, Bollington, UK (K Cocks PhD)

[21] European Society for Medical Oncology, Lugano, Switzerland (N Cherny MD)

[22] Critical Path Institute (C-Path), AZ, USA (S Eremenco MA)

[23] Hospital del Mar Research Institute, CIBERESP, ISCIII, Spain (M Ferrer PhD)

[24] Bayer AG, Berlin, Germany,and Gynecology, Obstetrics and Reproductive Medicine, University Medical School of Saarland, Homburg, Germany (C Gerlinger PhD)

[25] Ghent University, Ghent, Belgium (Prof E Goetghebeur PhD)

[26] Institute for Quality and Efficiency in Health Care (IQWiG), Cologne, Germany (U Grouven PhD)

[27] Japan Clinical Oncology Group (JCOG), National Cancer Center Hospital, Tokyo, Japan (J Mizusawa PhD)

[28] KU Leuven, Leuven, Belgium, and Hasselt University, Hasselt, Belgium (Prof G Molenberghs PhD)

[29] Region Hovedstaden, Copenhagen, Denmark (M Aa Petersen MSc)

[30] Amsterdam University Medical Center, Amsterdam, The Netherlands (J Reijneveld MD, PhD)

[31] University Health Network, Toronto, Canada (J Ringash MD MSc)

[32] Evaluation Software Development, Austria (G Rumpold PhD)

[33] University of Sydney, Faculty of Medicine and Health, Susan Wakil School of Nursing and Midwifery, Sydney Quality of Life Office, Sydney Australia (C Rutherford PhD)

[34] European Medicines Agency, Amsterdam, The Netherlands (C Quinten PhD)

[35] Health Canada, Ottawa, Canada (M Sasseville PhD)

[36] Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany (Prof W Sauerbrei PhD)

[37] Norwegian Medical Products Agency, Oslo, Norway (A Schiel PhD)

[38] National Cancer Institute, Bethesda, MD, USA (AW Smith PhD)

[39] Johns Hopkins University School of Medicine, Baltimore, MD, USA (Prof C Snyder PhD)

[40] Leeds Institute of Medical Research, University of Leeds, Leeds, UK and Leeds Cancer Centre, St James's Hospital, Leeds, UK (Prof G Velikova MD PhD)

[41] University of Texas MD Anderson Cancer Center, Houston, TX, USA (Prof XS Wang MD, MPH)

Corresponding author
Dr Madeline Pe, Quality of Life Department, European Organisation for Research and Treatment of Cancer, Brussels 1200, Belgium
Email **madeline.pe@eortc.org**

Summary

Standardising the implementation of patient-reported outcomes (PROs) in clinical trials is crucial for evaluating the benefits and risks of cancer treatments. The Setting International Standards in Analysing Patient-Reported Outcomes and Quality of Life Endpoints in Cancer Clinical Trials-Innovative Medicines Initiative (SISAQOL-IMI) has developed 146 consensus-based recommendations for designing, analysing, interpreting and presenting PROs in cancer clinical trials. This initiative, undertaken from 2021 to 2025, involved experts, including statisticians, PRO measurement experts, clinicians, and patient representatives from 41 organisations representing regulatory agencies, academia, the pharmaceutical industry, health-technology assessment bodies and patient advocates. SISAQOL-IMI provides guidance on implementation of PRO in randomised controlled trials and single arm trials, terminology, definitions and selection of PRO score interpretation thresholds, and visualising PRO results for different audiences. To facilitate the implementation of these standards, in addition to this publication, four key outputs are available: an interactive table, a guidebook, plain language materials and a glossary (www.sisaqol-imi.org).

Search strategy and selection criteria

References for this Review were identified through searches of PubMed with the search terms *("patient reported outcome analysis") OR ("quality of life analysis") AND "cancer" AND "clinical trials"*. No date restrictions were included. Articles were also identified through searches of the authors' own files and recommendations by the SISAQOL-IMI Consortium. Only papers published in English were reviewed. The final reference list was generated based on originality and relevance to the broad scope of this Review.

Introduction

Health-related quality of life and other patient-reported outcomes (PROs) are recognised as important for evaluating the benefits and risks of cancer treatments. These outcomes are important to a broad range of stakeholders: clinicians, academics [1-4], patient advocates [5], drug developers, international regulatory agencies and health technology assessment (HTA) bodies[6-8]. Although guidelines exist on how to include PROs in protocols [9], how to report them in trial publications [10], and how to create graphical presentations [11], they do not provide information on agreed methodological standard for the design, analysis, interpretation, and reporting of PRO data that would be acceptable for various decision-makers. Previous reviews have consistently demonstrated that vague PRO research objectives, poorly defined PRO endpoints, including unclear definitions of clinically meaningful change or difference, and lack of transparency in the reporting of PRO findings have raised concerns about their reliability[12]. This gap may impact the comparative evaluation of cancer clinical trials and hinder the optimal use of PRO data in the decision-making of various stakeholders. A common framework is essential across different clinical trial environments (e.g., academic and industry-sponsored) to ensure that trials generate high-quality PRO data that meet the needs of various stakeholders and for less experienced researchers to access a best-case methodology. To establish consensus recommendations, it was necessary to bring together different stakeholders to develop guidelines that accomplish these needs.

Expanding the SISAQOL "Setting International Standards in Analysing Patient-Reported Outcomes and Quality of Life Endpoints in Cancer Clinical Trials" work [12] into the SISAQOL-Innovative Medicines Initiative (SISAQOL-IMI) was a logical step forward. SISAQOL-IMI aimed to consider various stakeholder needs, adopt recent developments in the methodological literature (e.g., the estimands framework from ICH E9 (R1), which offers guidance on statistical principles for clinical trials aiming for pharmaceutical product registration) [13, 14]. SISAQOL-IMI involved expert statisticians and other PRO measurement experts to develop practical tools that support the implementation of these the recommendations [15]. The goal of the SISAQOL-IMI was to establish consensus-based guidelines for designing, analysing, interpreting and presenting PROs in cancer clinical trials.

Four key scientific priority areas were identified:

a) Randomised controlled trials (RCTs), where PROs can be used to evaluate the clinical benefit of an intervention and/or describe the patient perspective, for instance, to complement clinician-reported adverse event data.
b) Single arm trials (SATs), where PROs often are used to describe the patient perspective, for example, to support generation of future PRO related hypotheses in an RCT setting or to complement clinician-reported adverse events. In settings where an RCT is not feasible to evaluate clinical outcomes [16] results from SATs are sometimes accepted by regulatory authorities [17]. Recommendations are needed on how to analyse and interpret PROs in such studies [18] to leverage their potential for decision-making.
c) Presenting and visualising PRO results in trials, where graphic displays are commonly used to visualise results. Recommendations are required for optimally presenting data for different audiences.
d) Defining clinically relevant thresholds for differences and changes in PRO scores. Interpretation of the clinical relevance or meaningfulness of differences and changes in PRO scores is necessary, but challenging due to heterogeneity in the definition of these concepts, and inconsistencies in both terminology and the methodology on which they are based [19-21].

Despite many design and analysis considerations being similar between SATs and RCTs, the absence of randomisation in SATs implies that more care is needed to reduce bias and avoid misleading interpretations.

To support the dissemination and implementation of the recommendations among various expert and stakeholder groups, SISAQOL-IMI has generated scientific and plain language versions of the recommendations, supported by an online, interactive glossary. This manuscript provides an overview of the consensus process, the methods used and the project outcomes, including key recommendations and the final outputs: the interactive table (a webtool that allows users to easily navigate through the recommendations, and get a tailored set of recommendations based on their PRO research objective and variable of interest), the guidebook, the plain language recommendations, and the glossary. It concludes with lessons learned and outlines the plan for implementation and sustainability.

Methodology

Implementing the Vision: Organising Workstreams to Harmonise PRO Standards

Individual researchers and organisations often follow their own procedures and standards for the design, analysis and interpretation of PRO data. This lack of consistency had led to varying analytical approaches and, at times, confusing or non-comparable findings, making it difficult for stakeholders to use PRO data to effectively inform decision-making. The SISAQOL-IMI Consortium was established to address these gaps by bringing together relevant stakeholders who use PROs in the evaluation of cancer treatments, including international regulatory bodies, health technology assessment bodies, industry, academic and professional societies alongside experts in statistics, PRO measurement, clinical oncology, and patient advocacy. This collaboration ensured that the resulting recommendations are both methodologically robust and accessible to both technical and non-technical audiences. More information on the Consortium has been reported in a previous publication[15].

The work was organised in eight different work packages (WPs), as illustrated in the Appendix (page 1), with international, multidisciplinary participation, including patient representatives, in all WPs. The Steering Committee, composed of WP leaders and a management team, met bimonthly to address issues, adjust work plans, initiate actions, and ensure alignment across WPs. The General Assembly (GA), with the 41 participating organisations, acted as the decision-making body. The Consortium included over 180 members from 15 countries, representing 33 funded organisations and eight with other agreements. Details on the SISAQOL-IMI Consortium's organisation are available on the Innovative Health Initiatives website: (Innovative Health Initiative | IHI Innovative Health Initiative)[22].

Design of the consensus process

The consensus-building process used in this project employed a modified Delphi method. While the traditional Delphi approach relies on anonymous individual expert surveys conducted over multiple rounds with the possibility of individual experts to review the results and reconsider their votes based on the additional information provided, our method used anonymous surveys at the organizational level, combined with in-person or hybrid meetings and active engagement with diverse stakeholders, including patient partners. This approach

enabled real-time discussion, iterative refinement of recommendation statements, and a more inclusive and transparent decision-making process.

There were five yearly meetings (see Table 1). The first consensus process focused on prioritising concepts. Statements were not developed at this point. Consensus process two and three focused on statements related to RCTs (WP2), SATs (WP3) and PRO score interpretation thresholds (WP6). The third process also included statements on how to present PRO results (WP4). The fourth consensus process included one statement related to patient involvement and the final updates of statements for RCTs, SATs, and PRO score interpretation thresholds. Consensus meeting five focused on the final ratification of the consensus recommendations developed in the previous years and sustainability plan.
The overall framework of the consensus process for the development of the recommendations is described in more detail below (see also Figure 1). Each process was initiated by the ongoing work within each WP, followed by the consensus process in the SISAQOL-IMI consortium with voting rounds, discussions, revisions, harmonisation, validation, and the final ratification of the recommendations by the GA. SISAQOL-IMI defined "statements" as the specific formulation of each new advice developed and ratified within the consortium, while "recommendations" were defined as the final output which included statements with the corresponding examples and explanations.

Development of "statements" for vote

To develop the statements for each of the four priority areas (RCTs, SATs, visualisation and presentation of PRO results, and interpretation of PRO results), the work was divided among four scientific WPs. These WPs used a multi-step process to gather the information needed to develop the statements (see Figure 1).
The initial step involved conducting comprehensive literature reviews. These reviews collated relevant information, evaluated current standards and identified gaps in the literature, while highlighting areas of similarities and divergence. The methods used to select and extract relevant data were described in research protocols, with details provided in Appendix 5 of the online Guidebook (www.sisaqol-imi.org).
Each literature review informed the statements within its respective priority area.
Harmonisation across WPs took place after the draft statements were developed.

The literature review on RCTs focused on current practices of PRO analysis, existing stakeholder guidelines and key methodological recommendations for PRO analysis in RCTs[12]. The literature review on SATs focused on current practices and methodological recommendations on design, analysis, reporting and interpretation of SATs[16]. The literature review on visualisation focused on evidence on the graphical representation of PRO data[11] and more general information for the design of PRO visualisations. The literature review on PRO score interpretation thresholds focused on publications on clinically meaningful change thresholds between 2009 and 2021. Studies establishing PRO score interpretation thresholds for the most frequently used PRO measures in oncology and methodological articles discussing application of these thresholds were included. Using the information from the literature reviews, expert discussions were conducted within the WPs. Results from these discussions fed into the formulation of the initial set of statements for RCTs, SATs, visualisation and presentation of PRO data, and PRO score interpretation thresholds that were included in the consensus survey.

Consensus voting and review process

Proposed statements were evaluated by the full SISAQOL-IMI consortium –41 organizations—with each organization casting one vote. Initial votes were collected via pre-meeting online surveys using a 5-point Likert scale (from Strongly Agree to Strongly Disagree), with options for "Don't Know" and "Not Applicable" along with qualitative comments. Work package leads reviewed survey feedback, addressed comments, and revised statements in collaboration with the experts within their WP. Statements achieving consensus by a two-thirds majority across all stakeholder groups, with no concerns raised, did not have to be revoted (i.e., second vote) during the in-person consensus meeting but were made available online via SharePoint for additional comments from the Consortium. Statements that did not reach the two-thirds majority but received at least half of the votes; and/or statements that reached two-thirds majority but raised concerns in some stakeholder groups were discussed, revised if necessary and re-voted on during the second round of voting at the consensus meeting. The second round of voting was an important step since some votes may have been a result of misinterpretation of a concept or statement and needed further discussions or elaboration with the rest of the Consortium. Unresolved statements (that received less than one-half of the votes or did not reach sufficient consensus) were either withdrawn or revised for inclusion in the following year's consensus process.

The Consortium agreed upon these voting rules, including the two-step voting procedure (Appendix 2 on page 2). Recognising that unanimous agreement was not always achievable, the Consortium developed a "diverging views document" to capture differences in perspectives[15].

Cross-cutting Workstreams

Participation and Patient Engagement

WP7 coordinated the consensus-building process, actively encouraging timely engagement from all participating organizations. All SISAQOL-IMI organisations participated in all consensus surveys, except for one organisation that opted out of voting in consensus survey 4 due to time constraints. Each GA/consensus meeting was attended by approximately 80 attendees representing all stakeholder groups. Between seven and 12 patient representatives participated in pre-survey discussions and surveys, and between four and six attended the consensus meetings. To encourage input from patient representatives, four workshops were arranged prior to the consensus meetings to address key issues and clarify complex concepts.

Independent validation

Different initiatives ensured transparency of the consensus process and evaluated whether the statements were accurate, easy to understand and feasible to be implemented in clinical research. The Independent Scientific Advisory Board provided continuous critical review of statements developed by the scientific WPs, clarifying any concerns through discussions with the responsible WP leaders. In addition, one WP (WP 5) performed a two-step independent validation of the preliminary statements, involving interviews with experts and pilot testing of the statements. First, they conducted interviews with 17 individuals with various expertise within oncology: statisticians, clinicians, PRO methodologists. These experts represented academia, industry, regulatory/HTA bodies, and non-profit cancer organisations. Their aim is to evaluate the clarity of the statements. While most statements were interpreted as intended, some confusion arose from unfamiliar terminology or concepts. The WPs used the feedback to revise their statements as needed. Thereafter, 12 experts tested the preliminary statements by applying them to a study protocol with a defined PRO objective, setting up a statistical analysis plan, and outlining how the PRO results would be presented.

Harmonization across WPs, language review, and development of the glossary

To ensure harmonisation of recommendations between RCTs and SATs, the two WPs reviewed each other's recommendations. Recommendations developed for RCTs or SATs that could also be applicable for the other WP were either adopted by the other WP without changes, or adapted with minor changes to the statements, explanations, or examples. To ensure a common understanding and consistent terminology both in scientific and plain language, a glossary was created[15]. A dedicated team developed scientific and plain language versions using a hierarchy of recognised dictionaries (see Appendix 3, page 3). The Consortium reviewed and agreed upon all the proposed terms. The glossary enabled multiple stakeholders with diverse backgrounds and training, including patient representatives, to actively participate in meaningful discussion and decisions across WPs. In addition, this resource supported terminology harmonization across WP statements and was important for external participants during the independent validation process. Once all recommendations were available, a professional language editor reviewed all the recommendations in close collaboration with WP leaders, statisticians and PRO methodologists, clinicians and patient representatives to ensure they were unambiguous, easy to understand and consistent across WPs. The language editor then reviewed all final output documents to harmonise the language.

Presentation of Final Recommendations

Given the breadth of over 140 recommendations, the guidance was designed to be intuitive and easy to navigate. The table format allows users to directly access the cell that aligns with their specific PRO objective and endpoint, streamlining the process of identifying relevant guidance. This structure also encourages users to approach their study planning with clearly defined PRO endpoints.

The recommendations were structured as concise statements accompanied by explanations and examples, presented by study design (RCTs or SATs). To facilitate navigation, the recommendations are arranged in a tabular format, with columns specifying the study objective (confirmatory—superiority and equivalence/non-inferiority—or descriptive), and rows specifying the PRO variable of interest (the PRO endpoint). This structure creates 30 individual 'cells' (18 for RCTs, and 12 for SATs), each representing a unique combination of study design and PRO variable of interest (e.g., an RCT with a confirmatory superiority objective [column] and time to PRO improvement [row] endpoint). Figure 2 demonstrates the structure for RCTs as displayed in the interactive table. Within each cell, recommendations

are structured based on an analytical framework, consisting of the estimands framework of ICH E9 (R1)[14] and five additional attributes (PRO score interpretation thresholds, study design considerations, external comparison (for single-arm studies only), analysis considerations and results visualisation and presentation).

Key results and outputs of SISAQOL-IMI

SISAQOL-IMI ratified 146 out of the 149 proposed statements related to RCTs (WP2, n = 50), SATs (WP3, n = 43), visualisation and presentation of PRO results (WP4, n = 25), PRO score interpretation thresholds (umbrella term replacing clinically meaningful change) (WP6, n = 27), and patient involvement (WP8, n = 1). Figure 2 describes the development and evolution of the statements throughout consensus processes 2, 3 and 4. During the process, nine statements required re-voting after revision. Due to the substantial overlap between statements for RCTs and SATs, 25 shared recommendations were further harmonised across the WPs (19 from RCTs and 6 from SATs) and adopted either identically or adapted with minor wording changes.

Among the 146 accepted statements, the level of agreement was high, ranging from 70% to 100%, and 82% of the statements had agreement above 85% (see Appendix, pages 4-7). Forty-two statements reached consensus in the first round of voting, while 22 were adapted with minor wording changes without discussion. Eighty-two statements required discussion during consensus meetings. For five statements, divergent views among the stakeholder groups (see Appendix, pages 8-9) were included as "considerations" along with the final recommendation.

The key results for each of the four scientific priority areas from the WPs are presented in Table 2.

a) For RCTs, the recommendations emphasise the need to align estimands with research objectives, an area identified as lacking in current practice[23]. The recommendations provide guidance on both general issues, such as strategies for intercurrent events, handling missing data, and overall PRO analysis strategy, including the need for supplementary and sensitivity analyses. They also address specific issues dependent on

the type of analytic metrics used (e.g., time to event, responder analyses) including how to account for repeated measures.

b) For SATs, gaps in the current practice for addressing research questions related to PROs were identified [16]. The recommendations focus on what to consider when including PROs in SATs, providing guidance on formulating research questions taking into account the attributes of the estimands framework, and addressing challenges such as the absence of a randomised control group, handling of intercurrent events and missing data, which are quite distinct from absence of data following death. These recommendations are illustrated in a SAT case study with and without an external control group[24].

c) With regards to visualisation and presentation of PRO results, the recommendations consider both scientific audiences and non-specialist readers. In addition to recommendations on graph types based on previous evidence[11], advice is provided on the information to include in visualisations, tailored to specific contexts. For example, graphs presenting PRO data should be consistent with the pre-specified domains and time frames of the trial, with exploratory or descriptive results clearly labelled. They should also include details on sample size, intercurrent events, and missing data to clarify the basis for estimates. Statistical significance should be reported mainly for confirmatory objectives or labelled as exploratory when applicable. Scales in figures should reflect the full range of PRO scores, and the directionality of results should be clearly stated. For non-specialist readers, statistical significance is best indicated by symbols (e.g., asterisks) rather than p-values.

   In addition, based on an informal process, general guidelines on creating graphical representations are included. These focus on effectively using colour, highlights, and figure captions, improving readability, maintaining consistency, and avoiding visual clutter. This general advice was not subject to a formal consensus voting process.

d) To harmonise terminology for PRO thresholds used for interpretation of clinically meaningful change, the umbrella term “PRO score interpretation threshold” was introduced that refers to both patient-level and group-level data. This umbrella term is complemented with specific terms for patient- and group-level settings. Recommendations link different types of PRO thresholds to specific statistical analysis methods to ensure the correct interpretation of results. In addition, key criteria are provided for selecting an appropriate PRO score interpretation threshold. For example, the threshold should

preferably be anchor-based rather than distribution-based, the threshold should be established in a suitable patient population, and the anchors should be patient-centred. Finally, the recommendations encourage reporting how thresholds were chosen and applied in the analysis and interpretation of PRO data.

To support and encourage the implementation of the final recommendations, the Consortium has produced five key outputs, all of which will be available on the SISAQOL-IMI website [https://www.sisaqol-imi.org/]. These include the current SISAQOL-IMI publication; an interactive table that facilitates access to recommendations tailored to specific PRO objectives and endpoints; a guidebook offering detailed background on the development of the recommendations and instructions for using the SISAQOL-IMI outputs; plain language materials such as checklists and a plain language glossary; and a scientific glossary (see Appendix, page 12-13).

Discussion and lessons learned

The success of the SISAQOL-IMI project stems from consensus among diverse stakeholders and this was achieved through constructive, open, and results-driven collaboration that led to broadly accepted solutions. High agreement on most statements is expected to promote recognition and implementation of the recommendations within the scientific community and other stakeholder groups. The involvement of a wide range of stakeholders including statisticians, clinicians and patient representatives ensures the relevance of the SISAQOL-IMI content, which in turn supports and facilitates its uptake and application. The clinicians, patient representatives and patient advocates involved ensure the relevance of this work to patient care.

For SISAQOL-IMI, achieving broad consensus was prioritised over a simple majority vote. Differing perspectives were carefully considered, leading to statements adjusted during re-voting. Instead of including multiple disclaimers, the diverging views document was created. As a compromise, statements where the ideal situation was perceived as potentially unfeasible for a specific trial setting included the clause “any deviation should be justified”.

This extensive consensus process produced an agreed set of recommendations reflecting a shared understanding of good practice for PRO endpoints in cancer clinical trials. This comprehensive approach helps prevent the proliferation of multiple smaller guidelines. It

bridges a gap between PRO-specific guidelines with no analytical focus [9, 10], and analytical guidelines [13, 14], which do not specifically address key topics related to PRO endpoints.

Rigorous standards for PROs are needed, similar to those existing for other scientific research areas. SISAQOL-IMI agreed that all consensus recommendations should be methodologically robust and acceptable to the stakeholders. The process was transparent, thorough, and comprehensive, involving parallel and overlapping processes such as feedback on statements, recommendations, reports, and glossary development. The WP leaders' and Consortium's dedication and tight timelines were key to success.
Unlike other consensus processes, this effort involved organisations as institutional members rather than individuals, requiring internal discussions before reaching consensus. Face-to-face interactions at yearly General Assembly/consensus meetings proved invaluable, as informal discussions during breaks helped clarify issues and resolve disagreements, fostering stronger consensus. Patient involvement was crucial. Over time, special meetings and dialogues were arranged in order to strengthen their involvement. Providing adequate training and support enabled patient representatives to understand the discussions and contribute actively.

The SISAQOL-IMI project has already had a significant impact on the field of PROs and their application in cancer clinical trials. It has been cited in the Food and Drug Administration Guidance "Submitting Patient-Reported Outcome Data in Cancer Clinical Trials"[25] and received the American Statistical Association's Statistical Partnership Among Academia, Industry and Government (SPAIG) Award for collaboration among academia, industry, and government[26]. The recommendations have been recognised as an important solution to advance the PRO field at the European Medicines Agency/EORTC workshop[7] on PRO data in regulatory decision-making. Additionally, a collaboration with ESMO-Magnitude of Clinical Benefit Scale (ESMO/MCBS) has begun to provide methodological support for addressing relevant clinical questions[2].

Although the SISAQOL-IMI recommendations are tailored to cancer clinical research, the methods and concepts are broadly applicable. The Consortium expects they will be relevant for research in disease groups other than cancer, but further studies are needed to validate their relevance and effectiveness across disease populations.

Implementation and sustainability plan

The SISAQOL-IMI final outputs (interactive table, guidebook, plain language material, and glossary) will be freely available on the website [https://www.sisaqol-imi.org/], along with an instructional video on how to use the tools effectively.
To ensure ongoing progress after the project concludes, SISAQOL-IMI has established an updated steering committee and secretariat, along with a sustainability plan for regular updates and revisions. Digital and in-person courses will be developed, and the recommendations will be presented at international conferences. By raising awareness among professionals, the goal is to promote the use of the guidelines to improve PRO design, analyses, interpretation and visualisation of results of cancer clinical trials. SISAQOL-IMI will develop clinical trial protocol and statistical analysis plan templates to demonstrate how the SISAQOL-IMI recommendations can be integrated into these key trial documents. These templates will systematically incorporate PRO elements in a logical sequence, with example text aligned with the SISAQOL-IMI recommendations. The templates will be freely accessible on the project website.

Conclusion

Aligning statistical approaches, terminology, interpretation and visualisation of PRO results from cancer clinical trials is crucial to optimise the use of these data in decision-making within and across relevant stakeholders in the production and use of clinical trial evidence. To facilitate effective implementation in updated guidelines, publications and future studies, trial results should be presented in a clear and accessible format. The SISAQOL-IMI recommendations will help to achieve this goal and thus standardise inclusion of PROs in clinical trials in the future and make the results more transferable to clinical care and individual patient well-being. The training activities and sustainability plan are essential to ensure the long-term impact of our efforts.

Figures legends

Figure 1 Framework of each consensus process

Figure 2 Screenshot of the interactive table using the analytical framework for organisation of statements for one example

Figure 3 Development and evolution of statements within each work package

Contributors

All authors are members of the SISAQOL-IMI Consortium and were involved in the conceptualization of this Policy Review. CDA led in the drafting of the manuscript with support from AA, RSF, KB and MP. MS, SR, VB, LMV, AR were all part of the core writing committee. MP, AI, AA CC, MS, SR, SLC, BH, LMW, JC, MT, PC, JMG, JCC, KB, VP, StS, EJP are all work package leaders and contributed to the collection and analysis of the SISAQOL-IMI recommendations. All authors representing the 41 organizations in SISAQOL-IMI collected and provided the data in the development of the SISAQOL-IMI recommendations. All authors interpreted and reviewed the manuscript and the final outputs of the SISAQOL-IMI recommendations. All 41 organisations involved in SISAQOL-IMI reviewed and approved the final version of this Policy Review.

All authors had access to data at the stakeholder level. Organisation-level data were processed by MP, AA, and RSF. All authors had final responsibility for the decision to submit the manuscript for publication.

Declaration of interests

- KBJ, JG, BH, GR, reports receiving funding paid to their institution from the SISAQOL-IMI grant to conduct this work.
- ABe, CQ, NC, MF, received travel grants from SISAQOL-IMI to participate in the annual meeting.
- EB reports receiving consulting fees as a scientific advisor for AZ, Navigating Cancer, Verily, N-Power Medicine, Savor, Thyme Care and Resilience Health
- AC reports being an employee of Genentech from 2011 to 2019 and then AZ from year 2020 to 2024. However the current work was done as part of PRE LLC and on external time. AC reports patent unrelated to this work, AZ ownership. AC reports stocks, leadership roles, receipt of equipment and other services, and other financial / non-financial interest from AZ.
- PC, JC, SR reports being an employee of Pfizer and receives stocks and stock options as part of his employment.
- MC reports receiving grants paid to her institution from NIHR Birmingham Biomedical Research center, Health data research UK, Innovate UK, Macmillan Cancer Support, GSK, Research England, European Commission and EFPIA, Brain Tumour Charity, Gilead, Janssen, NIHR, UKRI, Merck. She also reports receiving revenue from commercial license of the symptom burden questionnaire – long COVID; She receives consulting fees from Aparito Ltd, Boehringer Ingelheim, CIS Oncology, Merck, Genentech, GSK, Pfizer, Vertex Pharmaceuticals, ICON, Halfloop,

University of Michigan, Northwestern University, EuroQoL, Shionogi, Cell and Gene Therapy Catapult. She also received payment or honoraria from University of Maastricht, Sount-Eastern Norway Regional Health Authority, Cochrane Portugal, Singapore National Medical Research Council. She reports leadership role in PROTEUS Consortium which she receives a consultancy fee. She reports a family member has stocks in GSK.

- OLA reports receiving grants and contracts paid to University of Birmingham from NIHR BRC, ARC, NIHR Blood and Transplant Research Unit in Precision Transplant and Cellular Therapeutics, Anthony Nolan, Sarcoma UK, Merck. OLA also reports personal consulting fees paid through University of Birmingham from Merck, GSK, Gilead, BI, and Innovate UK
- MB reports receiving an unrestricted research grant paid to his institution from Pfizer
- KC received travel grants from SISAQOL-IMI to attend the annual meeting. She reports consultancy fees from pharmaceutical and medical devices companies for statistical services.
- SE reports membership fees paid by PRO Consortium member firms to CPATH to support staff time to participate in the Consortium. She reports CPATH receiving grants and broad agency announcement contract payments from the US FDA. She reports receiving travel funds to participate in the American College of Neuropsychopharmacology 2022 Annual meeting. She reports receiving travel grants to attend the SISAQOL-IMI meeting.
- CG reports being an employee of Bayer and holds stock options as part of his employment.
- MS reports being an employee of Merck KGAa and receiving stocks/stock options as part of his employment.
- EP reports being an employee of AbbVie and this work was done as part of her employment at AbbVie
- JR is the Chair of the EORTC quality of life group (unpaid) and is the member of scientific advisory board of EpilepsieNL (unpaid). He received travel grants to attend the annual SISAQOL-IMI meeting.
- StS reports that her employer Myeloma Patients Europe receives grant and sponsorship from various pharmaceutical companies, which are listed here: https://www.mpeurope.org/about-mpe/our-funding/

- CS reports receiving research funding from Pfizer and Genentech to her institution. She received personal consulting fees from Shionogi and Movember, and travel costs were covered by Shionogi to present on a panel at the CPATH COA Consortium meeting
- GV reports receiving grants from NIHR, Pfizer and Yorkshire Cancer Research – all paid to her institution. She reports receiving consulting fees paid to her from Pfizer, Roche, Seagen. She also reports payments and honorarias paid to her from Pfizer, Roche, Novartis, Eisai and Sanofi. She reports receiving support from Pfizer and Roche to attend meetings and travel. She reports receiving payment as part of her participation on a data safety monitoring board or advisory board from Roche, Seagen and AZ. She holds leadership roles as part of the EORTC board of directors and NCRI Chair of Living with and beyond cancer group.

Acknowledgements

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

This publication reflects the views of the individual authors and should not be construed to represent official views or policies of the European Medicines Agency (EMA), the US Food and Drug Administration (FDA), US National Cancer Institute (NCI), Medicines and Healthcare products Regulatory Agency (MHRA), Institute for Quality and Efficiency in Health Care (IQWiG), Health Canada, the Norwegian Medical Products Agency (NOMA), the American Society of Clinical Oncology (ASCO), the European Society for Medical Oncology (ESMO) or any other institution, organisation, or entity. This project received no funding from the US National Institutes of Health (NIH). Neither IMI nor the EU, nor EFPIA are responsible for any use that may be made of the information contained therein.

We would like to express our sincere appreciation to former WP leaders Carla Mamolo, Jinma Ren, Jayne Galinsky, Linda Dirven as well as former SISAQOL-IMI Steering Committee members Daniel O'Connor and Kathy Oliver, for their invaluable efforts and dedication. We are also grateful to Rosa Abruzzese for her extensive review and harmonisation of the language in the final deliverables of the project. Finally, we appreciate Monika Turek's help for her organisational and administrative support during the entire SISAQOL-IMI project, including the submission of this manuscript.

References

1. Booth CM, Sengar M, Goodman A, Wilson B, Aggarwal A, Berry S, et al. Common Sense Oncology: outcomes that matter. *Lancet Oncol.* 2023;24(8):833-5.
2. Oosting SF, Barriuso J, Bottomley A, Galotti M, Gyawali B, Kiesewetter B, et al. Methodological and reporting standards for quality-of-life data eligible for European Society for Medical Oncology-Magnitude of Clinical Benefit Scale (ESMO-MCBS) credit. *Ann Oncol.* 2023;34(4):431-9.
3. Lacombe D, Bean J, Bogaerts J, Bottomley A, Collette L, Gilotay C, et al. 50 years of the EORTC and the central role of the EORTC Headquarters. *European Journal of Cancer Supplements.* 2012;10(1):3-12.
4. Friends of Cancer Research. Supporting patient-centric approach to dose optimization in oncology: patient-reported outcomes (PROs). 2022. Available from: https://friendsofcancerresearch.org/wp-content/uploads/Supporting_Patient-entric_Approach_Dose_Optimization_Oncology-PROs.pdf; 2022.
5. WECAN (Work group of European Cancer Patient Advocacy network). Patient-reported outcomes, and other patient-relevant measures and endpoints. Accessed Dec 11, 2024. Available from: https://wecanadvocate.eu/academy/patient-reported-outcomes-and-other-patient-relevant-measures-and-endpoints/
6. file:///C:/Users/mpe/OneDrive - EORTC/Desktop/SISAQOL-IMI final manuscript/www.fda.gov/regulatoryU.S. Food and Drug Administration. Core patient-reported outcomes in cancer clinical trials. October 2024. Available from: www.fda.gov/regulatory
7. EMA and European Organisation for Research and Treatment of Cancer (EORTC) workshop: How can patient-reported outcomes (PRO) and health-related quality of life (HRQoL) data inform regulatory decisions? | European Medicines Agency (EMA). 29.02.2024. Available from: https://www.ema.europa.eu
8. Directorate-General for Health and Food Safety, EU. Guidance on reporting requirements for multiplicity issues and subgroup, sensitivity and post hoc analyses in joint clinical assessments. 2024. Available from: https://health.ec.europa.eu/
9. Calvert M, Kyte D, Mercieca-Bebber R, Slade A, Chan AW, King MT, et al. Guidelines for Inclusion of Patient-Reported Outcomes in Clinical Trial Protocols: The SPIRIT-PRO Extension. *JAMA* 2018;319(5):483-94.
10. Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *JAMA* 2013;309(8):814-22.
11. Snyder C, Smith K, Holzner B, Rivera YM, Bantug E, Brundage M. Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data. *Qual Life Res* 2019;28(2):345-56.
12. Coens C, Pe M, Dueck AC, Sloan J, Basch E, Calvert M, et al. International standards for the analysis of quality-of-life and patient-reported outcome endpoints in cancer randomised controlled trials: recommendations of the SISAQOL Consortium. *Lancet Oncol* 2020;21(2):e83-e96.

13. ICH Expert Working Group. ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conferenceon Harmonisation E9 Expert Working Group. Stat Med. 1999;18:1905-42.

14. ICH Expert Working Group. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. European Medicines Agency; 2017. Published Feb 17, 2020. Available from: www.ema.europa.eu

15. Pe M, Alanya A, Falk RS, Amdal CD, Bjordal K, Chang J, et al. Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints in Cancer Clinical Trials-Innovative Medicines Initiative (SISAQOL-IMI): stakeholder views, objectives, and procedures. *Lancet Oncol* 2023;24(6): e270-83.

16. Liu L, Choi J, Musoro JZ, Sauerbrei W, Amdal CD, Alanya A, et al. Single-arm studies involving patient-reported outcome data in oncology: a literature review on current practice. *Lancet Oncol* 2023;24(5): e197-206.

17. Agrawal S, Arora S, Amiri-Kordestani L, de Claro RA, Fashoyin-Aje L, Gormley N, et al. Use of Single-Arm Trials for US Food and Drug Administration Drug Approval in Oncology, 2002-2021. *JAMA Oncol* 2023;9(2): e266-72.

18. Di Maio M. The Value of Patient-Reported Outcomes in Single-Arm Cancer Trials. *Cancer Invest* 2023;41(5): e491-4.

19. Cappelleri JC, Bushmakin AG. Interpretation of patient-reported outcomes. Stat *Methods Med Res* 2014;23(5): e460-83.

20. Coon CD, Cappelleri JC. Interpreting Change in Scores on Patient-Reported Outcome Instruments. *Ther Innov Regul Sci* 2016;50(1):22-9.

21. Bushmakin AG, Cappelleri JC. A Practical Approach to Quantitative Validation of Patient-Reported Outcomes: A Simulation-Based Guide Using SAS. Hoboken, NJ: John Wiley & Sons; 2022.

22. Initiative IH. SISAQOL-IMI Establishing international standards in the analysis of patient reported outcomes and health-related quality of life data in cancer clinical trials. 2021.Available from: https://www.ihi.europa.eu/.

23 Machingura A, Coens C, Pe M, Alanya A, Regnault A, Mazerolla F. Methodological work and use of Patient-Reported Outcomes data in randomised controlled trials (RCTs) in cancer: Literature reviews on current practices and guidelines. 2024.

24. Thomassen D, Roychoudhury S, Amdal CD, Reynders D, Musoro JZ, Sauerbrei W, et al. The role of the estimand framework in the analysis of patient-reported outcomes in single-arm trials: a case study in oncology. *BMC Med Res Methodol* 2024;24(1): e290.

25. US. Food and Drug Adminstration. Submitting Patient Reported Outcome Data in Cancer Clinical Trials, Guidance for Industry, Technical Specifications Document. 2023. Contract No: FDA-2018-D-1216.

26. American Statistical Association. Statistical Partnerships Among Academe, Industry, and Government (SPAIG) award. Available from: https://www.amstat.org/your-career/awards/statistical-partnerships-among-academe-industry-and-government-(spaig)-award.

Figure 1

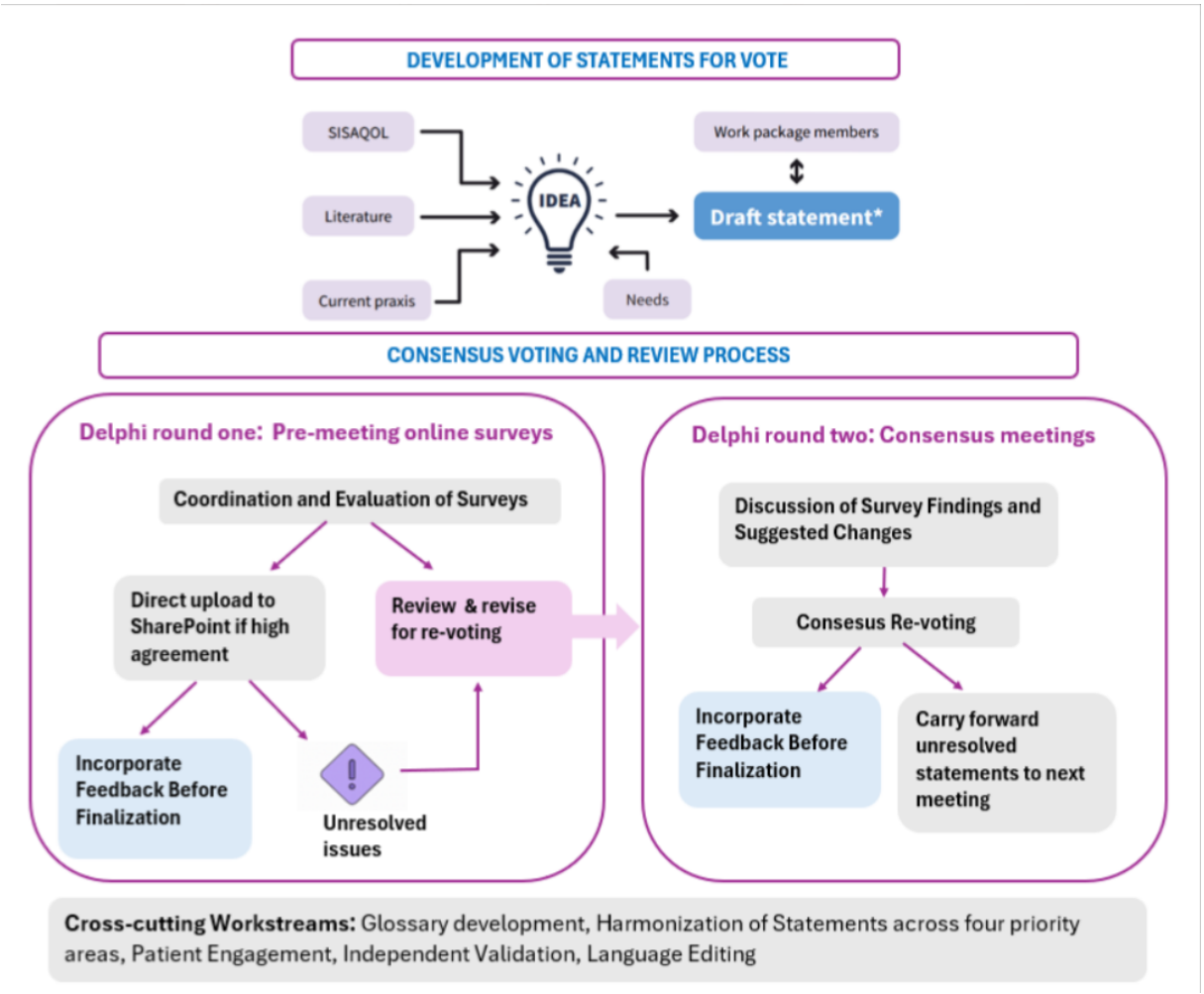


Figure 2

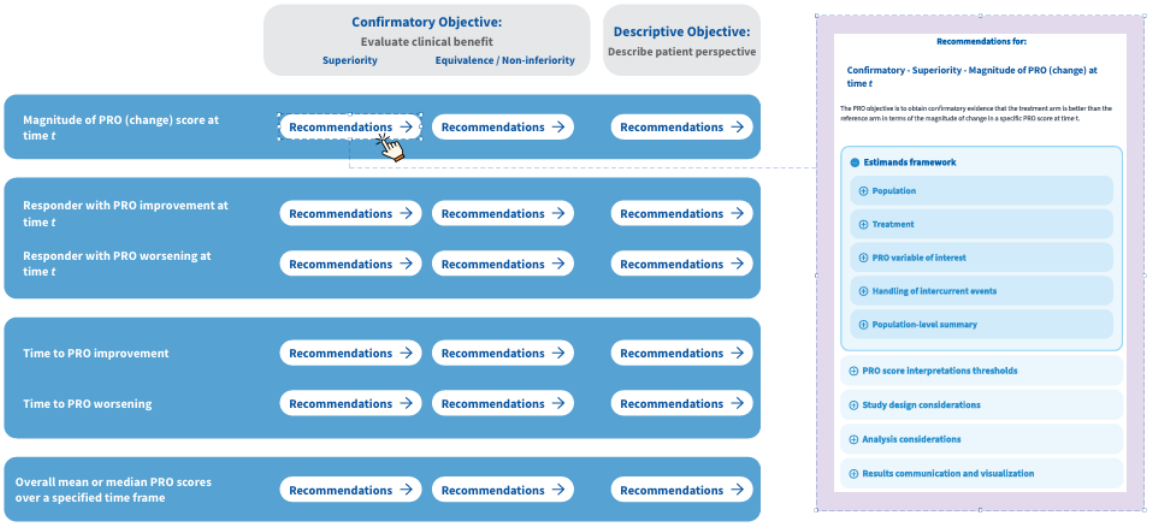


Figure 3

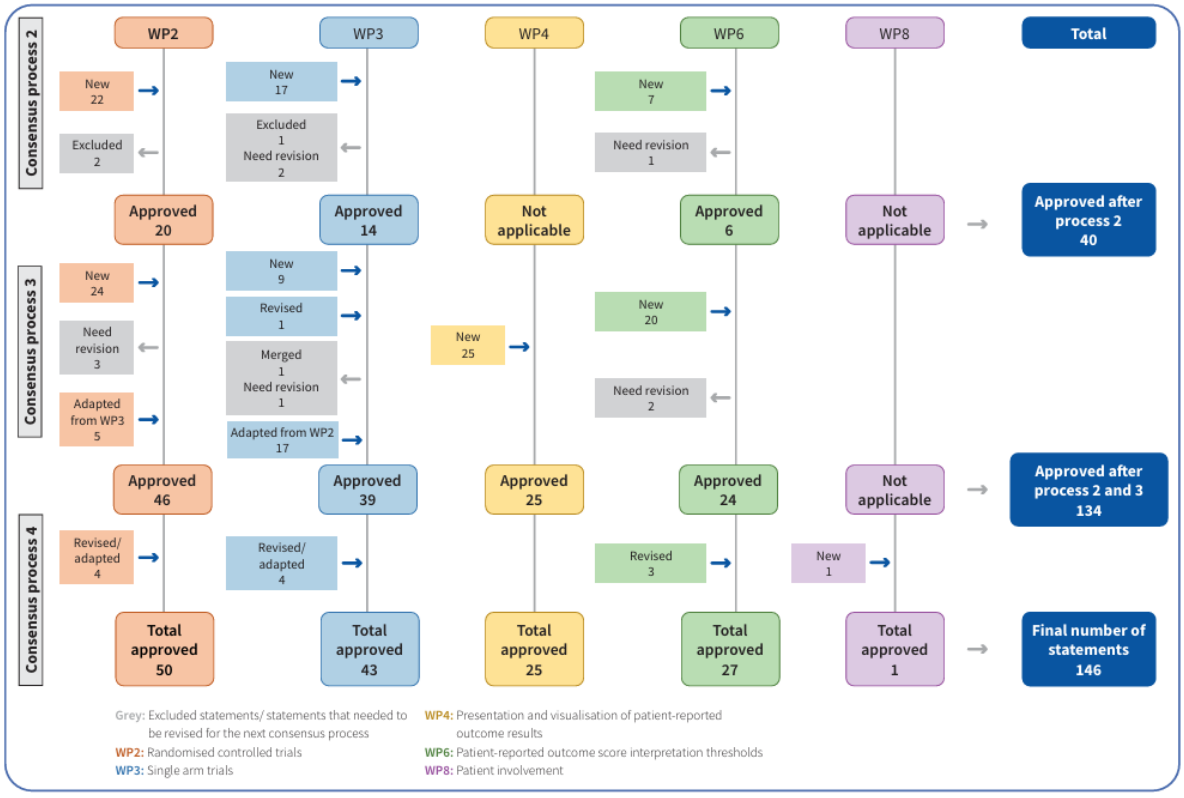


Table 1 Milestones for the general assemblies (GAs)

General Assembly meetings	Year	Milestones
1	2021	Defined the goals, priority of patient-reported (PRO) objectives and identified expectations
2	2022	Ratification of the first set of recommendations for cancer randomised controlled trials, single arm trials and clinical meaningful change/ PRO score interpretation thresholds
3	2023	Ratification of the updated and expanded version of recommendations for cancer randomised controlled trials, single arm trials, visualisation and presentation of PRO results and for clinical meaningful change/ PRO score interpretation thresholds
4	2024	Ratifications of the final version of recommendations for cancer randomised controlled trials, single arm trials, visualisation and presentation of PRO results and for clinical meaningful change/ PRO score interpretation thresholds
5	2025	Ratification of the final output and sustainability plan

Table 2. Scientific key results for each work package

Randomised controlled trials (WP2)	Current standards	SISAQOL-IMI standards	Rationale for change
Addressing death in patient-reported outcome (PRO) analysis	There is no standard method for addressing death in PRO analysis for RCTs. Hypothetical strategies are often used without specifying the underlying assumptions. According to the established guideline on statistical principles for clinical trials, ICH E9 (R1), a hypothetical strategy i s a method used to estimate the treatment effect by assuming a hypothetical scenario: what if the intercurrent event (e.g., death) did not occur. Commonly used methods often do not reflect the intended estimand/objective.	There are different strategies to address death, as an intercurrent event in RCTs (hypothetical, composite, while alive, principal stratum). The choice of strategy will have an impact on the treatment effect estimate and its interpretation. Protocols should define and justify a clear strategy in line with the assumptions based on the pre-defined PRO objective and discussed with relevant stakeholder groups.	Four main strategies for addressing death as an intercurrent event can be considered. For each strategy, the underlying assumptions and resulting interpretation need to be considered in the selection of the best-suited strategy that fits the context of PRO objectives, disease setting and study constraints.
Missing data versus Intercurrent event	Many statistical analyses assume the same number of observations per patient. Unobserved data are often considered as missing and addressed without consideration of underlying	The overall PRO analysis strategy should include a main PRO analysis supported by sensitivity (accounting for missing data) and/or supplementary analyses (accounting for intercurrent events). It is recommended to report an overview of relevant intercurrent	Intercurrent events can cause relevant PRO data to be unavailable for the analysis (e.g. after death). However, the way an intercurrent event is addressed in the analysis is linked to its interpretation and therefore to the

	cause. No distinction is made between intercurrent events and missing data.	events and reasons and frequencies for missing data.	objective. Missing data are unobserved data (e.g. questionnaire lost) and its impact on the results should be handled via sensitivity analyses. Missing data may bias results and affect uncertainty if not handled properly.
Completion rates and available data rates	PRO data may be unavailable for different reasons. The data actually used in analyses are often insufficiently reported. There is no standard measure addressing data quality nor consistent terminology.	Completion rates and available data rates should be reported for each assessment time point. For both, **the numerator** is set to the number of patients that completed the PRO assessment at that time point. **For the completion rate** the denominator is set as the number of patients with a scheduled PRO assessment at that time point. This denominator can change over time to account for e.g. deaths. **For the available data rate**, the denominator equals the number of patients randomized in the trial. This denominator will not change over time.	For the calculation of completion rates, a distinction should be made between failure to collect relevant data (leading to missing data) and the choice not to collect or use data due to an intercurrent event such as treatment discontinuation or progression of disease.
Handling missing data	Simple techniques such as single imputation, complete case analysis (only including patients with no missing data), or available case analysis (only including patients with no missing data at the time point of interest)	Single imputation, complete case analysis, or available case analysis to handle missing data are generally **not recommended.** A justification should be given if these approaches are used. As an alternative, multiple imputation techniques can be considered.	In many RCTs, attrition bias can occur when participants with specific attributes (e.g., worse physical status) are more likely to drop out than others. Simple techniques tend to be biased. Moreover, these techniques often ignore the uncertainty resulting

	are often used as they are easily understood.		from missing data may lead to a biased estimate.
Repeating cross-sectional analyses are not recommended in longitudinal analyses	Cross-sectional treatment effect estimates take into account data from only a single specific time point. This is not an efficient use of the PRO dataset. Moreover, such repeated estimates at consecutive time points are often presented as a longitudinal series. Time trends are then inferred but these may result in misleading interpretation.	It is **not recommended** to analyse data at each time point separately using multiple cross-sectional analyses. Longitudinal modelling is preferred.	There is a considerable loss of information by using cross-sectional analyses instead of modelling the full longitudinal profiles. Repeated cross-sectional testing results in multiple testing. It does not take into account data selection over time (due to patient attrition or missing data), or correlation between different observations of the same patient.
Single arm trials (WP3)	**Current standards**	**SISAQOL-IMI standards**	**Rationale for change**
Research objectives and estimands	In single-arm trials, PRO objectives are often unclear or not mentioned at all.	Single-arm trials should have pre-specified PRO objectives that should be translated into key clinical questions using the estimand framework.	Unclear or missing PRO objectives may lead to inappropriate analysis and ambiguous interpretation of results. A clearly specified objective is needed to define the research question and the corresponding estimand.
PRO objectives and the absence of a randomised control group	PRO objectives in single arm trials are usually descriptive. Naïve numerical comparisons are often made with external control data without	PRO objectives can be descriptive or confirmatory. The analysis strategy should be aligned with the research question using the estimand framework to address the question of interest. Comparisons can be made using change from baseline or a suitable	It is critical to align the analysis strategy with the research question of interest using the estimand framework question. In some situations, comparing PROs from a single-arm trial to

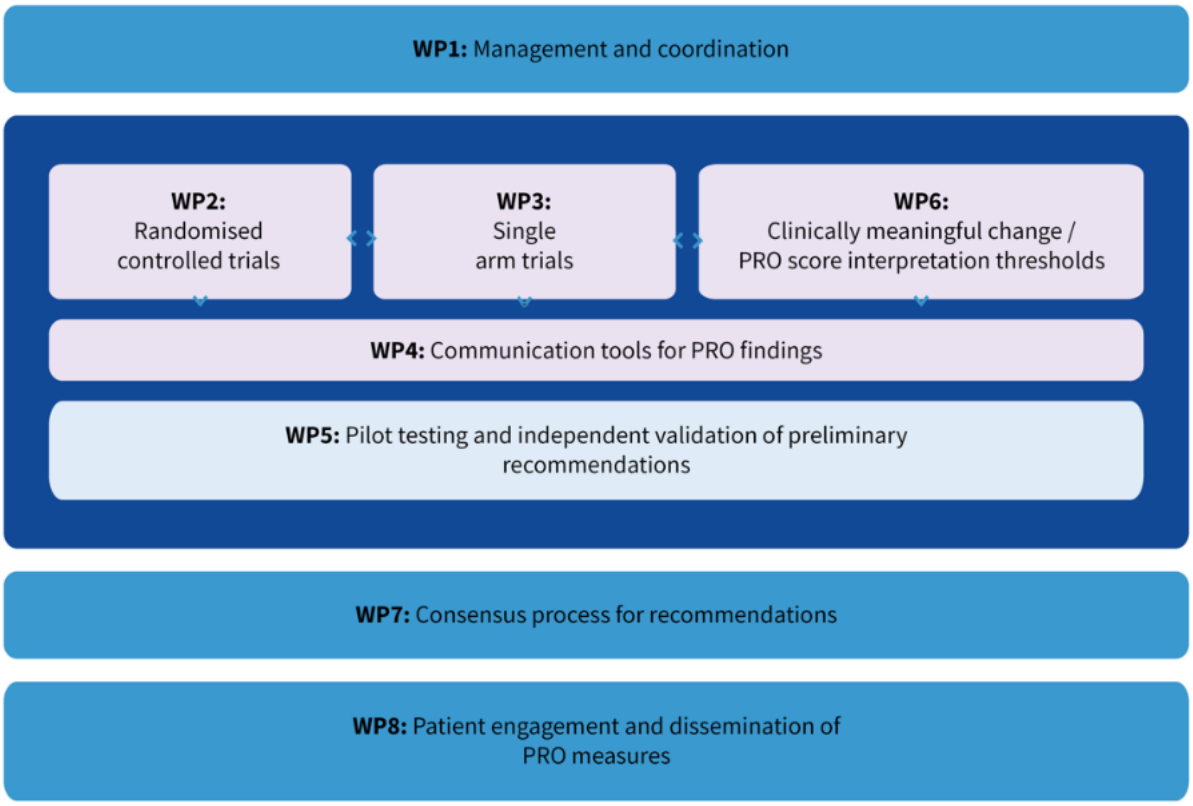
	considering the differences between data sources.	external control. Appropriate steps should be taken in the design and conduct to reduce bias and avoid misleading interpretations. The absence of randomisation should be addressed.	an external control can serve as confirmatory. However, without appropriate design and analysis considerations comparison with external control data may lead to erroneous conclusions.
Handling death in single-arm trials	Currently, there are no well-defined strategies for handling death in PRO analysis in single-arm trials.	There are different strategies to handle death in single-arm trials. The chosen strategy should be defined prior to analysis in line with the pre-defined PRO objective. For example, when describing PROs over time, the while-alive strategy is generally preferred. The population-level summary for this approach includes the PRO score of participants alive and descriptive statistics about death such as the proportion of patients still alive at the time point of assessment.	Different strategies can be considered for handling death. For each strategy, the underlying assumptions and resulting interpretation need to be considered. Using a "while-alive" strategy in single-arm trials aligns with a descriptive research objective (i.e., the intention to inform clinicians and patients about expected PRO after the start of treatment). To provide a comprehensive understanding, the expected PRO score at a specific time point should be accompanied by the estimated probability of survival at that same time,
Handling missing data vs intercurrent events	The distinction between missing data and data after intercurrent events is frequently overlooked, and assumptions made when	Researchers should clearly specify which strategies of the estimand framework are used for the intercurrent events and how missing values are handled. The plausibility of the underlying assumptions on which the analysis	It is critical to understand the underlying assumptions associated with each method. Sensitivity analyses should be conducted to assess the

	handling them are not specified.	method relies and whether the result is still in line with the intended estimand should be examined.	impact of assumptions made about missing data mechanisms and supplementary analyses accounting for handling intercurrent events. For example, methods such as linear mixed models or generalized linear mixed models implicitly impute values for expected outcomes after death when, in reality, PRO values cease to exist after death. These methods correspond to a hypothetical strategy, which may not align with the research question.
Communication of PRO findings using visualisation (WP4)	**Current standards**	**SISAQOL-IMI standards**	**Rationale for change**
Figures on main results	The PRO results presented in figures may not correspond to the pre-specified research objective and statistical analysis plan of the trial.	The figures should reflect the pre-specified PRO objectives and statistical analyses of the trial, particularly regarding the PRO domains and time points/frames presented. If figures are presented for additional exploratory/descriptive results, their purpose should be clearly indicated.	1) To enable immediate differentiation as to whether figures depict the main results of a confirmatory analysis or additional exploratory/descriptive results. 2) To avoid selective presentation (cherry picking) of results such as certain PRO domains and time points
Statistical significance	Figures representing results of exploratory/descriptive analyses may include	Either figures representing results of confirmatory analyses with predefined hypotheses should be the only ones to include	1) To promote targeted and transparent reporting of statistical test results. 2) To prevent readers

	information on statistical significance without clearly identifying them as exploratory/descriptive.	information on statistical significance; or, if figures depict results from statistical tests for exploratory/ descriptive purposes, this should be clearly indicated and a rationale given.	from concluding that exploratory/descriptive results from statistical tests provide the same level of evidence as confirmatory results.
Scaling in graphs	Inconsistencies may exist in the scaling applied to graphs within and across trial reports/ publications.	Use consistent scaling reflecting the full PRO score range whenever possible, particularly in graphs based on the same PRO score whenever possible.	To promote distortion-free and comparable representations of PRO results within and across trial reports/publications.
Sample size, intercurrent events and missing data	Inconsistencies may exist in the inclusion of numbers of observed patients in figures presenting PRO results.	Graphs should include the number of observed patients, missing data and intercurrent events at each assessment point.	Enhances transparency on 1) the number of patients on which PRO results are based and 2) how these numbers compare to the original sample size.
Directionalities of PRO scores	Despite existing standards on this topic, there are inconsistencies in whether labels are provided to support interpretation.	Graphs should include labels to support interpretation (e.g., which direction indicates a good/bad PRO score, which direction indicates an (improvement/ worsening).	Existing standards require emphasizing to facilitate interpretation of results.
Interpretation of PRO results (WP6)	**Current standards**	**SISAQOL-IMI standards**	**Rationale for change**
Harmonized terminology for thresholds for interpretation of PRO data	Different terms and definitions are used and inconsistently applied for conceptually similar thresholds for interpreting PRO data (e.g., minimal clinically important difference, clinically meaningful change).	A harmonised terminology has been established, providing clarification on terms and definitions for various types of PRO score interpretation thresholds.	The currently heterogeneous terminology poses challenges when selecting thresholds for a specific purpose and may lead to inappropriate or misleading application and interpretation.

Differentiation of patient- and group-level PRO score interpretation thresholds	The literature rarely distinguishes PRO score interpretation thresholds for patient-level scores (i.e., for within-patient change) and group-level scores (e.g., for between-group differences, within-group change).	Terminology is provided that differentiates patient- and group-level PRO score interpretation thresholds and different types of thresholds are linked to specific statistical analysis methods for correct implementation and interpretation.	Lack of distinction between patient- and group-level thresholds compromises interpretation of PRO data and sample size calculation. This can lead to invalid conclusions, for example, when responder thresholds for within-patient change are used for interpretation of mean differences between groups.
Key criteria for selecting PRO score interpretation thresholds	Different methodological approaches are used to establish PRO score interpretation thresholds for specific PRO measures, with anchor- and distribution-based methods being the most common.	Key criteria for selecting appropriate thresholds are provided.	The various methods used may result in threshold values that are not fit-for-purpose in specific settings. The established key criteria aim to support the application of valid, relevant thresholds as a cornerstone of PRO data analysis and interpretation.
How to report PRO score interpretation thresholds	Reporting PRO score interpretation thresholds (e.g., in clinical trial publications and protocols) is frequently insufficient and does not allow for an understanding and a critical evaluation of their appropriate implementation.	Reporting critical aspects of how thresholds are selected and applied in the analysis and interpretation of PRO data are encouraged.	Detailed reporting increases the clarity of PRO objectives and provides transparency for evaluating the selection of thresholds, their use in statistical analysis, and for appropriate interpretation of PRO results

Appendix 1 Organisation of work packages (WP) in SISAQOL-IMI



Appendix 2. Established voting rules for the SISAQOL-IMI survey

- Each SISAQOL-IMI organisation got one vote.
- For each statement, the possible responses for the agreement were “strongly agree”, “somewhat agree”, “neither agree nor disagree”, “somewhat disagree”, or “strongly disagree”.
- In addition, there were two other response options: “this statement is not applicable to my organisation” or “don’t know”.
- The respondents were encouraged to provide comments if responding “neither agree nor disagree, somewhat disagree, strongly disagree, this statement is not applicable to my organisation or don’t know”. Optionally, comments could also be provided when responding “strongly agree” or “somewhat agree”.
- In the analyses, “strongly agree” and “somewhat agree” were counted as agreement, while “neither agree nor disagree”, “somewhat disagree”, and “strongly disagree” were regarded as no agreement.

- Responses of "this statement is not applicable to my organisation" or "don't know" were regarded as "abstaining", and the organisation was not counted as having voted for that statement (removed from the denominator).
- A statement was ratified if at least 2/3 of voting organizations agreed on the statement (denominator may vary for each statement, as the number of organisations "abstaining" may differ).
- If a ratified statement did not have consensus in all stakeholder groups, the comments from organisations who disagreed were analysed and discussed. If only one or few stakeholder groups did not agree with a statement, we carefully reviewed their reasons (reviewed comments, discussed with involved parties, clarified misunderstandings) and explored possible compromises.

Appendix 3. Hierarchical list of references/sources used to define SISAQOL-IMI glossary terms

References/sources
1. The European Medicine Agency (EMA). International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use guidance documents (ICH) Topic E 9 Statistical Principles for Clinical Trials. CPMP/ICH/363/96. September 1998. and EMA ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. Step 5. EMA/CHMP/ICH/436221/2017. 17 February 2020. Available from: https://www.ema.europa.eu/en/ich-e9-statistical-principles-clinical-trials-scientific-guideline
2. The U.S. Food and Drug Administration Patient-Focused Drug Development Glossary. Content current as of: 06/08/2018. Available from: https://www.fda.gov/drugs/development-approval-process-drugs/patient-focused-drug-development-glossary
3. Dictionary of Quality of Life and Health Outcomes Measurement. Nancy E. Mayo (Ed.). ISOQOL. 2015 First edition. (See https://www.isoqol.org/ for more information) Approval for use of terms and definitions in the SISAQOL-IMI glossary was granted by ISOQOL in June 2024.
4. Pe M, Alanya A, Falk RS, Amdal CD, Bjordal K, et al. Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints in Cancer Clinical Trials-Innovative Medicines Initiative (SISAQOL-IMI): stakeholder views, objectives, and procedures, The Lancet Oncology, Volume 24, Issue 6, 2023, Pages e270-e283, ISSN 1470-2045, https://doi.org/10.1016/S1470-2045(23)00157 2.
5. National Cancer Institute. NCI dictionaries. NCI Dictionary of Cancer Terms.

Available from:
https://www.cancer.gov/publications/dictionaries/cancer-terms

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):483–494. doi:10.1001/jama.2017.21903.

7. The U.S. Food and Drug Administration. Oncology Center of Excellence. Patient-Friendly Language for Cancer Clinical Trials (Content current as of: 02/16/2021)

Available from: https://www.fda.gov/about-fda/oncology-center-excellence/patient-friendly-language-cancer-clinical-trials

Appendix 4. Level of agreement on ratified statements (≥2/3 organisations accepted), n = 146*.

Statement no WP2/RCT	Acronym	# agreed	# agreed + disagreed	% agreement§
1	EstFrame1_GEN	40	41	98
2	AnalSens1_GEN	39	41	95
3	Anal2_RCT	40	41	98
4	Pop1_RCT	38	41	93
5	Pop2_RCT	36	41	88
6	ICEdisprog1_RCT	36	40	90
7	ICEprodev1_RCT	37	41	90
8	ICEconc1_RCT	36	41	88
9	ICEdisprog3_RCT	32	40	80
10	ICEprodev3_RCT	38	40	95
11	ICEconc3_RCT	37	40	93
12	ICEdisprog2_RCT	30	40	75
13	ICEprodev2_RCT	37	40	93
14	ICEconc2_RCT	36	40	90
15	ICEdisprog5_RCT	37	41	90
16	ICEprodev5_RCT	38	41	93
17	ICEconc5_RCT	39	41	95
18	ICEdisprog4_RCT	37	41	90
19	ICEprodev4_RCT	39	41	95
20	ICEconc4_RCT	38	41	93
21	AnalSens2_GEN	41	41	100
22	ICEdeath1_RCT	38	41	93
23	ICEdisc1_GEN	32	38	84
24	Psum4_RCT	37	38	97
25	Psum5_RCT	40	41	98

26	Psum6_RCT	32	38	84
27	Design3_GEN	38	41	93
28	Anal1_GEN	37	38	97
29	ICE3_RCT	29	38	76
30	Design4_GEN	40	41	98
31	AnalMain5_GEN	35	38	92
32	PROvar2_RCT	32	38	84
33	AnalMain2_GEN	37	41	90
34	AnalMain3_GEN	38	40	95
35	AnalMain4_GEN	33	37	89
36	Assump1_GEN	41	41	100
37	AnalMain1_GEN	37	41	90
38	Psum3_GEN	37	37	100
39	Psum1_GEN	41	41	100
40	Psum2_GEN	41	41	100
41	AnalMain8_RCT	36	38	95
42	AnalMain7_GEN	36	39	92
43	AnalSens3_GEN	38	40	95
44	AnalMain6_GEN	38	40	95
45	Design1_GEN	39	40	98
WP3/SAT	**Acronym**	**# agreed**	**# agreed + disagreed**	**% agreement[§]**
1	EstFrame2_SAT	41	41	100
2	EstFrame3_SAT	40	41	98
3	EstFrame4_SAT	28	40	70
4	EstFrame5_SAT	34	41	83
5	EstFrame6_SAT	34	41	83
6	Pop3_SAT	35	41	85
7	AnalPres3_SAT	36	41	88
8	AnalMain9_SAT	39	41	95
9	ICEdeath3_SAT	28	40	70
10	ICEdeath4_SAT	29	41	71
11	Design2_GEN	39	41	95
12	ICE5_SAT	34	41	83
13	ExComp1_SAT	31	38	82
14	ExComp2_SAT	32	39	82
15	AnalPres2_SAT	38	41	93
16	ICEdeath2_SAT	31	41	76
17	Assump2_SAT	32	38	84
18	AnalPres1_GEN	39	40	98
19	ICE1_GEN	32	37	86
20	ICE2_GEN	30	38	79
21	ICE4_SAT	28	38	74

22	PROvar1_GEN	36	41	88
23	Design1_GEN	37	40	93
24	AnalSens4_SAT	40	40	100
25	AnalMain7_GEN	36	37	97
26	AnalSens3_GEN	36	40	90
WP4/Graphs	**Acronym**	**# agreed**	**# agreed + disagreed**	**% agreement[§]**
1	VizSci1_GEN	39	41	95
2	VizSci11_GEN	39	41	95
3	VizSci3_GEN	38	41	93
4	VizSci2_GEN	38	38	100
5	VizSci6_GEN	41	41	100
6	VizSci7_GEN	34	38	89
7	VizSci8_GEN	38	38	100
8	VizSci5_GEN	39	41	95
9	VizSci4_GEN	39	40	98
10	VizSci9_GEN	37	38	97
11	VizSci10_GEN	41	41	100
12	VizSci13_GEN	38	41	93
13	VizSciType1_GEN	34	38	89
14	VizSciType2_GEN	32	38	84
15	VizSciType3_GEN	37	41	90
16	VizSciType4_GEN	34	38	89
17	VizSci12_GEN	35	40	88
18	VizPlain1_GEN	39	41	95
19	VizPlain2_GEN	35	38	92
20	VizPlain3_GEN	38	38	100
21	VizPlain4_GEN	36	38	95
22	VizPlainType1_GEN	35	38	92
23	VizPlainType2_GEN	32	38	84
24	VizPlainType3_GEN	35	38	92
25	VizPlainType4_GEN	37	41	90
WP6/CMC	**Acronym**	**# agreed**	**# agreed + disagreed**	**% agreement[§]**
1	CMCselc2_GEN	38	41	93
2	CMCimp6_GEN	39	41	95
3	CMCimp1_GEN	40	41	98
4	CMCimp5_GEN	39	41	95
5	CMCimp4_GEN	38	40	95
6	CMCimp3_GEN	40	41	98
7	CMCimp2_GEN	37	40	93
8	CMCimp7_GEN	37	38	97

9	CMCimp8_GEN	33	38	87
10	CMCimp9_RCT	34	36	94
11	CMCselc7_GEN	38	41	93
12	CMCselc1_GEN	38	41	93
13	CMCselc5_GEN	36	40	90
14	CMCselc3_GEN	40	41	98
15	CMCselc4_GEN	37	38	97
16	CMCimp14_GEN	39	41	95
17	CMCimp13_GEN	41	41	100
18	CMCrep1_GEN	39	41	95
19	CMCimp10_GEN	41	41	100
20	CMCrep3_GEN	33	38	87
21	CMCrep4_GEN	36	38	95
22	CMCrep5_GEN	36	38	95
23	CMCimp11_GEN	40	41	98
24	CMCrep6_GEN	38	41	93
25	CMCselc6_GEN	32	39	82
26	CMCrep2_GEN	36	40	90
27	CMCimp12_GEN	35	40	88
WP8/patient engagement	**Acronym**	**# agreed**	**# agreed + disagreed**	**% agreement§**
1	VizPlainType5_GEN	36	36	100

WP: work package
RCT: randomised controlled trial
SAT: single-arm trial
CMC: clinical meaningful change (PRO score interpretation thresholds)
**22 adapted statements without voting (5 in RCT/WP2 and 17 in SAT/WP3)*
§ = the following responses are not included in denominator: missing votes, don't know, not applicable

Appendix 5. Statements with divergent views

The five statements with divergent agreement among stakeholder groups (i.e. at least one stakeholder group with < 50 % acceptance)

WP2/RCT	**Statement text**	**Range of % agreement by stakeholder group §**	**Action taken to address divergent views**

# 9: ICEdisprog3_RCT	When the goal of the PRO objective is to draw conclusions about clinical benefit (confirmatory objective) using time-to-deterioration within a time frame for a specific PRO domain: if a disease progresses before the PRO deterioration occurs, the main PRO analysis technique would be to use the PRO scores collected after disease progression to determine whether a PRO deterioration event occurred or not. Any deviations should be justified.	33-100	"Any deviations should be justified" was included in the statement text
# 12: ICEdisprog2_RCT	When the goal of the PRO objective is to draw conclusions about clinical benefit (confirmatory objective) using time-to-improvement within a time frame for a specific PRO domain: if a patient's disease progresses before a PRO improvement occurred, the main PRO analysis technique would be to use the PRO scores collected after disease progression to determine whether there was a PRO improvement or not. Any deviations should	0-100	"Any deviations should be justified" was included in the statement text, as well as an additional need to perform supplementary/ sensitivity analysis

	be justified. An overview of the relevant ICEs should be reported and discussed to assess to which extent the intercurrent event could have impacted the PRO results. This may be supported by supplementary/sensitivity analysis. Alternative ICE strategies might be considered as supplementary analyses to explore the robustness of outcomes and inform about the potential of bias.		
WP3/SAT			
# 3: EstFrame4_SAT	A PRO objective for single arm studies may be to make comparisons, either to baseline or to external controls. This is on condition that appropriate care is taken in the design and the conduct of studies to reduce bias, to avoid misleading interpretations due to the absence of randomisation and treatment blinding.	0-100	Both adjustment to the statement and explanations has been made to exercise caution to avoid undue bias
# 21: ICE4_SAT	If the PROMs collected after an ICE are not relevant to the research question of interest (such as	43-100	Both adjustment to the statement and explanations has been made to express the need to perform

	switching of treatment whilst assessing benefits), and furthermore, if the research question would consider a scenario in which the ICE would not occur (for example, evaluating patients who did not switch treatment), the PRO values after the ICE may be disregarded and addressed using a hypothetical strategy. To understand the impact of the model assumptions and to ensure the validity of analysis results, it is recommended to conduct sensitivity and supplementary analyses.		supplementary/ sensitivity analysis
WP4/ Graphs			
# 24: VizPlainType3_GEN	Icon arrays can be used for *plain figure* versions for visualization of proportion of patients (e.g., improvement, stable state or deterioration) at a specific time point.	0-100	Adjustment to the explanations has been made to ease the clarity

WP: work package
RCT: randomised controlled trial
SAT: single-arm trial
§ missing voting, don't know, and not applicable are not included (as above)

The statement text and its explanation can be found in the web tool and the guidebook. For further information about the diverging views see guidebook appendix.