This is a repository copy of *Moving forward toward standardizing analysis of quality of life data in randomized cancer clinical trials*.

White Rose Research Online URL for this paper: http://eprints.whiterose.ac.uk/135870/

Version: Accepted Version

**Article:**
Bottomley, A, Pe, M, Sloan, J et al. (38 more authors) (2018) Moving forward toward standardizing analysis of quality of life data in randomized cancer clinical trials. Clinical Trials. ISSN 1740-7745

https://doi.org/10.1177/1740774518795637

© 2018, The Author(s). This is an author produced version of a paper published in Clinical Trials. Reprinted by permission of SAGE Publications.

**Reuse**
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Moving forward towards standardizing analysis of quality of life data in randomized cancer clinical trials

Andrew Bottomley¹, Madeline Pe¹, Jeff Sloan², Ethan Basch³, Franck Bonnetain⁴*,
Melanie Calvert⁵, Alcyn Campbell⁶, Charles Cleeland⁷, Kim Cocks⁸, Laurence Collette¹, Amylou C Dueck³, Nancy Devlin¹⁰, Hans-Henning Flechtnert¹¹, Carolyn Gotay¹², Eva Greimesl¹³, Ingolf Griebesch¹⁴, Mogens Groenvold¹⁵, Jean-Francois Hamel¹⁶, Madeleine King¹⁷, Paul G Kluetz¹⁸, Michael Koller¹⁹, Daniel C Malone²⁰,
Francesca Martinelli¹, Sandra A Mitchell²¹, Carol M Moinpour²², Jammbe Z Musoro¹,
Daniel O’Connor²³, Kathy Oliver²⁴, Elisabeth Piault-Louis⁶, Martine Piccart²⁵,
Francisco L Pimentel²⁶, Chantal Quinten²⁷, Jaap C Reijneveld²⁸, Christoph Schürmann²⁹, Ashley Wilder Smith²¹, Katherine M Soltys³⁰, Rajeshwari Sridhara¹⁸,

¹European Organisation for Research and Treatment of Cancer (EORTC), Brussels, Belgium
²Alliance Statistics and Data Center, Mayo Clinic, Rochester, MN, USA
³Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC, USA
⁴Methodology and Quality of Life Unit in Cancer, INSERM U1098, University Hospital of Besançon, Besançon, France
*deceased May 20, 2017
⁵Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, College of Medical and Dental Sciences, University of Birmingham, UK
Genentech, a member of the Roche group, San Francisco, CA, USA

Department of Symptom Research, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Adelphi Values, Bollington, Cheshire, UK

Alliance Statistics and Data Center, Mayo Clinic, Scottsdale, AZ, USA

Office of Health Economics, London, UK

Clinic for Child and Adolescent Psychiatry and Psychotherapy, University of Magdeburg, Magdeburg, Germany

School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada

Obstetrics and Gynecology, Medical University Graz, Graz, Austria

Boehringer-Ingelheim, Frankfurt, Germany

Department of Public Health and Bispebjerg Hospital, University of Copenhagen, Copenhagen, Denmark

Methodology and Biostatistics Department, University Hospital of Angers UNAM, Angers, France

School of Psychology and Sydney Medical School, University of Sydney, Sydney, NSW, Australia

US Food and Drug Administration, Silver Spring, MD, USA

Center for Clinical Studies, University Hospital Regensburg, Regensburg, Germany

College of Pharmacy, University of Arizona, Tucson, AZ, USA
Acknowledgments of research support for the study

EORTC received an unrestricted education grant from Boehringer Ingelheim GmbH to initiate this work and additional financial support was provided by Fonds Cancer (FOCA) from Belgium.
Corresponding Author
Andrew Bottomley, Ph.D., Quality of Life Department, European Organization for Research and Treatment of Cancer, 83/11 Avenue E. Mounier, 1200 Brussels, Belgium; Tel: +32 (0) 2 774 16 61; andrew.bottomley@eortc.be

Running Head
PRO analysis in cancer trials

Total number of words: 2363

Note
This publication reflects the views of the individual authors and should not be construed to represent official views or policies of the US Food and Drug Administration, US National Cancer Institute, Medicines and Healthcare Products Regulatory Agency, Institute for Quality and Efficiency in Health Care (IQWIG) or Health Canada.
Abstract

Background

There is currently a lack of consensus on how health-related quality of life and other patient-reported outcome (PRO) measures in cancer randomized clinical trials are analyzed and interpreted. This makes it difficult to compare results across RCTs, synthesize scientific research, and use that evidence to inform product labelling, clinical guidelines, and health policy. The Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data for Cancer Clinical Trials (SISAQOL) Consortium aims to develop guidelines and recommendations to standardize analyses of PRO data in cancer RCTs.

Methods and Results

Members from the SISAQOL Consortium met in January 2017 to discuss relevant issues. Data from systematic reviews of the current state of published research in PROs in cancer RCTs indicated a lack of clear reporting of research hypothesis and analytic strategies, and inconsistency in definitions of terms, including “missing data”, “health-related quality of life”, and “PRO.”

Based on the meeting proceedings, the Consortium will focus on three key priorities in the coming year: developing a taxonomy of research objectives, identifying appropriate statistical methods to analyze PRO data, and determining best practices to evaluate and deal with missing data.

Conclusion
The quality of the Consortium guidelines and recommendations are informed and enhanced by the broad Consortium membership which includes regulators, patients, clinicians, and academics.

**Keywords:** guidelines, standards, cancer clinical trials, health related quality of life, patient-reported outcomes
**Background**

The patients' voice is increasingly part of the evaluation of risks and benefits of cancer therapies. As such, data on patient-reported outcomes (PROs) that quantify how a patient feels and/or functions, are frequently collected in cancer clinical trials.\(^1\) However, the lack of standards and clear guidelines on how these patient-reported data should be analyzed and interpreted diminishes their added value and make it difficult to compare results across different trials.\(^2\) This hinders research findings from informing important processes such as clinical decision making, product labelling, clinical guidelines, and health policy.\(^3\)

To explore the perspectives of multiple stakeholders, the European Organisation for Research and Treatment of Cancer convened a multidisciplinary international consortium focusing on “Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data for Cancer Clinical Trials” (SISAQOL). This manuscript summarizes the Consortium’s work to date and provides a critical backdrop for future recommendations.

**Methods and Results**

The European Organisation for Research and Treatment of Cancer’s kick-off meeting in January 2016 solicited attendees’ views on the need for developing standards, guidelines, and recommendations for PRO analysis in trials. There was a clear consensus that standards and best practices for PRO data analysis are lacking, such guidance is urgently needed, and a multidisciplinary team of experts is crucial to ensure technically correct, comprehensive, and balanced
recommendations. Based on this input, SISAQOL moved forward. A summary of this initial meeting has been previously reported.\textsuperscript{3}

The SISAQOL Consortium’s second consensus meeting was convened a year later to discuss concrete strategies regarding standardizing PRO analysis, with the end goal being to produce internationally recognized guidelines. Participants were leading PRO researchers and statisticians and representatives from international oncological and medical societies, advisory and regulatory bodies, academic societies, the pharmaceutical industry, cancer institutes, and patient advocacy organizations (see author list).

Perspectives

Regulators/advisory bodies

Regulators and advisors from the European Medicines Agency network, the US Food and Drug Administration, Health Canada, and the Institute for Quality and Efficiency in Health Care discussed the current role of PROs in their organizations’ decision-making processes.

It was clear that these groups recognize the importance of the patient’s experience and perspective and their added value in the benefit-risk assessment of cancer treatment, and efforts are underway to identify methods to best incorporate the patient’s voice into their programs.\textsuperscript{4–7}

However, it was also evident that regulators have reservations about the conclusions drawn from PRO data to date. Poorly defined research objectives (and hypotheses) and lack of rigorous standards in analyzing PRO data in regulatory submissions have hampered the usefulness of such data for regulatory decision-making. To assess the potential added-value of patient-reported data in trials, one key criterion is to establish international standards in data analysis.
It was emphasized that throughout a patient’s cancer journey, clear communication between the patient and the stakeholders involved in treatment on risks, benefits and potential side effects is crucial. Patients need to be heard regarding side effects, their feelings about their treatments, and how they are functioning physically, mentally, emotionally, and socially. Such information needs to be collected and synthesized across patients to increase the knowledge base about patient experiences in a way that will be useful for future patients. Identifying the best ways to involve patients and survivors in initiatives such as SISAQOL, which focus primarily on technical research issues, is challenging. The discussion of missing PRO data provided a clear opportunity for possible patient participation. Missing data is a critical issue in any trials missing data present difficulties in analysis and drawing robust conclusions about treatments. Minimizing avoidable missing assessments is critical. While researchers have identified many factors that contribute to avoidable missing data, patients themselves generate PRO data, and SISAQOL provides an opportunity to work with patients to get their ideas about how to minimize the amount of avoidable missing PRO data in clinical trials and to communicate the importance of providing complete data.

Five systematic reviews provided a summary of the current quality of hypothesis reporting and analysis of PRO data in published trials in locally advanced and metastatic breast cancer, advanced non-small cell lung cancer, small cell lung cancer, as well as two reviews on methods for dealing with missing data. For the purposes of this report, three key findings from these reviews are highlighted.
Hypothesis. Clear research objectives and a priori hypotheses are needed prior to statistical analysis. Otherwise, statistical analyses are exploratory, and no conclusions can be drawn. In the systematic reviews for metastatic breast\textsuperscript{11} and advanced non-small cell lung cancers,\textsuperscript{12} findings showed that only 7% of the articles (metastatic breast: 4 of 58 articles; advanced non-small cell lung cancer: 2 of 27 articles) reported specific a priori PRO research hypothesis. In a systematic review evaluating the quality of PRO reporting in trials published between 2002-2008, only around 50% of the 794 trials reported a PRO hypothesis.\textsuperscript{16, 17} These findings imply that although PRO data are being included in trials, statistical analyses are often being conducted without clear reported PRO research objectives and hypotheses. This causes uncertainty regarding whether the results reported are based on: a) a priori hypotheses with an a priori statistical analysis plan that allow conclusions to be drawn, or b) exploratory analyses intended to generate future hypotheses, but where findings from this trial remain inconclusive.

Statistical methods. In the three systematic reviews, preliminary findings showed that at least 10 different statistical methods were used to evaluate PRO data.\textsuperscript{11–13} This is a problem, since the variety of statistical techniques employed makes it challenging not only to compare findings across trials, but also to build on previous work to make the results more generalizable and conclusive.

Another problem is the failure to correct for type 1 error (or alpha adjustment) for multiple testing. This problem is particularly relevant for PRO data due to the possibility of calculating scores for an entire measure, subdomains and/or at a range of time points. If multiple scales and/or assessment points are tested independently from one another, and the alpha level is not adjusted for multiple testing (e.g. it remains at 0.05 for each of the tests), the probability of observing at least one
significant result simply due to chance is inflated. This then leads to findings that are difficult to interpret. This was a limitation found in this literature. All three reviews\textsuperscript{11–13}, less than 40% of the articles controlled for type 1 error when it was needed (metastatic breast cancer: 40%, 23 of 58 articles; advanced non-small cell lung cancer: 4%, 1 of 27 articles; small cell lung cancer: 27%, 9 of 33 articles).

Missing data. Missing data is a common problem in PRO analysis in trials. How missing data are considered in analysis, especially when the amount of missing data is substantial, may bias the analysis and critically influence the conclusions that can be drawn. For this reason, reports need to specify the analytic approach used to address missing data.\textsuperscript{18, 19} In the systematic reviews for metastatic breast cancer\textsuperscript{11} and advanced non-small cell lung cancer,\textsuperscript{12} only 24% (14 of 58 articles) and 19% (5 of 27 articles) of the articles, respectively, reported how the analysis addressed missing data. Furthermore, the statistical methods across reports ranged from simple imputation (e.g. last observation carried forward) to model-based methods (e.g. pattern mixture modelling). These findings demonstrate the lack of standardization on how to handle missing PRO data.

Implications

\textbf{Developing hypothesis}

The systematic reviews show a lack of clearly reported research hypotheses. New guidelines for protocol development (i.e., SPIRIT PRO)\textsuperscript{20, 21} and PRO reporting (i.e., CONSORT-PRO)\textsuperscript{16} also recognize this issue. It was proposed that three components are necessary to specify in an a priori research hypothesis, specifically:

\begin{itemize}
  \item the domains of interest;
  \item how the reference arm is expected to behave within the time frame of interest;
  \item how the treatment arm is expected to behave relative to the reference arm.
\end{itemize}
A rationale and evidence-based arguments informed by clinical and patient experience are needed to support these components of the hypothesis.

To address standardized classification of such hypotheses, the Consortium agreed to develop a taxonomy of PRO objectives, including underlying assumptions. This taxonomy has the potential not only to help researchers to be more precise in hypotheses in protocols, but also to allow comparison of objectives and findings across trials. The taxonomy is currently under development.

**Statistical methods**

The systematic reviews\(^{11-13}\) demonstrate that the current trials literature does not provide a good foundation to determine which statistical method is recommended for a specific research objective. Not only is there a lack of clearly reported research objectives, but there is also no consensus on which statistical methods to use. Rather than recommending a specific statistical method, it was agreed that a more useful approach is to define essential statistical properties for analyzing PRO data. For example, an important statistical property is adjusting for covariates. Covariate adjustment is a common practice in trials for stratification, controlling for potential imbalance between treatment arms, or improving precision of the treatment effect (especially when the covariate has an important influence on the outcome).\(^{22, 23}\)

The Consortium will compile a systematic list of statistical properties, with a recognition of the importance of balancing feasibility and accuracy. Following consensus on identifying essential statistical properties, the Consortium will determine statistical methods that fit these criteria, which can then be matched with research objectives identified in the previously mentioned taxonomy.

SISAQOL also emphasized the importance of developing criteria for descriptive statistics (including visualization) that can provide more complete documentation of
patient reports. For example, it is common practice to report the mean (or median) levels of a PRO measure per treatment arm over time. However, although this summary statistic may be useful, it is not sufficient to use it alone. Rather, this should be accompanied by a measure of variability to provide an indication of the diversity of responses. For example, an average score of “3” in a possible range of scores from 1 to 5 could mean that all participants reported a “3” or that half of the participants reported “1” and the other half reported a “5”. A measure of variability can capture this difference, whereas the average would not. SISAQOL Consortium members will work toward developing guidelines to standardize descriptive analyses and visualization approaches across all trials.

Missing data

Before undertaking statistical analysis, the researcher needs to be certain that the dataset is valid for analysis. Guidelines often indicate that a substantial amount of missing data can invalidate any analysis\(^{18}\). The Consortium questioned the definition of substantial, given that this is not consistent in the literature. The Institute for Quality and Efficiency in Health Care standard approach (e.g. Regofaranib,\(^{24}\) p. 3) is to consider valid any analysis from a dataset that includes baseline data with at least one follow-up from at least 70% of patients. However, this criterion is not used consistently across the literature. Different definitions of missing data and their calculation may lead to varying practices and results and call out for guidelines.

It is not currently clear if it will be possible for international consensus on a fixed threshold that defines an acceptable percentage of missing data. For example, in a hypothetical situation where 65% of PRO data are missing, some investigators would agree that drawing conclusions on treatment efficacy based on these patient reports would be futile. However, others may argue that analyzing the 35% of patients for
whom data are available could be useful to understand more about patient well-
being in this subgroup, although generalization to the larger trial population would
not be possible. Exploring the potential to identify a fixed threshold for an acceptable
percentage of missing data to have a valid analysis and robust findings is a priority
question for the SISAQOL Consortium.

Another SISAQOL goal is to develop and validate a set of macros, an automated
way to systematically examine missing data patterns and the impact of different
imputation methods on findings. An initial pilot test of macros developed by the Mayo
Clinic team was performed on a Mayo trial dataset. Capabilities of these macros
include producing percentages of missing values over time and providing more
detailed information on missing data patterns. Moreover, these macros also
implement and test the effects of several imputation methods, which could then be
used for sensitivity analysis.

The macros (or others) may prove useful following further testing and validation with
other clinical trial datasets and guidelines on the appropriate use and interpretation
of findings from these missing data macros are needed.

**Terminology**

An evidence-based review on the history on terminology of patient-reported
indicators (such as quality of life, health-related quality of life and PRO) in the
context of cancer and trials demonstrate the relatively recent emergence of terms
(see Table 1). Indeed, widespread consensus on the exact meaning of these terms
is not yet set, and new terminologies continue to surface: e.g. patient-generated
health data, patient experience and patient-centered outcome.

Currently, definitions have been offered by regulatory bodies\(^5,6\)), and academic
societies (e.g. International Society for Quality of Life Research\(^25\)). Although not all
definitions are the same, health-related quality of life is generally seen as a subcategory within the broader PRO construct, which may include other patient-reported variables. Currently, as seen in Table 1, the most citations and research information are based on “quality of life” and “health-related quality of life” endpoints than for the broader “PRO” concept. It is not within the remit of the Consortium to find consensus on these non-statistical terminologies. Regardless of the terminology used, Consortium members cited likely considerable overlap in data analytic approaches for all PROs, given that all come from the same source (the cancer patient).

Conclusion

Based on discussions and evidence extracted from systematic reviews of published literature, the SISAQOL Consortium has confirmed the priority need to develop guidelines and standards in analyzing PRO data in trials. The Consortium is focusing on three key priorities: developing a taxonomy of research objectives, identifying appropriate statistical methods to analyze PRO data, and determining how best to evaluate and deal with missing data. SISAQOL’s work will provide a toolbox for analysis of PRO outcomes in trials that is urgently needed and will advance the international research agenda now and into the future.
Acknowledgments

We would like to thank Linda Dirven for her comments on the manuscript.

Writing support services were provided by John Bean (Bean Medical Writing).
Declaration of conflicting interests

Andrew Bottomley reports grants from Boehringer Ingelheim for the SISAQOL project and from Merck for the Reference Values data project (where Andrew Bottomley is the Principal Investigator); and unrestricted education grants to EORTC from Pfizer and BMS for conferences. Madeline Pe’s fellowship was funded by Boehringer Ingelheim for the SISAQOL project. Melanie Calvert reports receiving honoraria from Astellas Pharma and Fening Pharma; consulting role for Astellas Pharma and Fening Pharma; and has travel, accommodations or expenses from Fening Pharma and Astellas Pharma. Alicyn Campbell reports that she is an employee of Genentech. Kim Cocks reports being an employee of Adelphi Values LTD; consulting or advisory roles with AMGEN, ENDOMAG LTD, ORTHOX LTD, CREOMEDICAL LTD. Nancy Devlin reports having consulting or advisory roles for the Association of the British Pharmaceutical Industry (ABPI), Pfizer, Lilly Global and Astellas Europe; and have received research funding from the Association of the British Pharmaceutical Industry (ABPI), Pfizer, Lilly Global, Roche and Astellas Europe. Michael Koller reports travel, accommodations or expenses paid or reimbursed by Biofrontera. Ingof Griebsch reports to be an employee of Boehringer Ingelheim and has stock or other ownership interest to disclose with BMS, Roche, Astra Zeneca, Celgene and Lucyte. Dan Malone reports receiving honoraria received from Sanofi; and consulting or advisory roles with Sanofi, Amgen and Pharmacies. Kathy Oliver reports receiving honoraria from BMS, AbbVie, GSK, Novartis; consultancy or advisory roles with BMS, AbbVie, GSK, Novartis; participated in a speakers’ bureau with BMS. Elisabeth Piault-Louis reports being an employee of Genentech; and has stocks or other ownership interest with Genentech; and has travel, accommodations or expenses paid by Genentech. Francisco Pimentel
reports being an employee of BlueClinical Phase 1; consulting or advisory roles with OM Pharma SA; travel, accommodations expenses paid by OM Pharma SA. JC Reijneveld reports receiving travel, accommodation or expenses from Roche Nederland NV. Martin Taphoorn reports consulting or advisory roles with Hoffmann-La Roche. Galina Velikova reports receiving honoraria from Roche and EISAI; participating in a speaker’s bureau for Roche; and travel accommodations or expenses paid or reimbursed by Roche. All other authors declare no competing interests.

References


Table 1. Citations on quality of life related terms found by searching PubMed

<table>
<thead>
<tr>
<th>Term</th>
<th>1st mention</th>
<th>&quot;Critical mass&quot; (N)**</th>
<th>2015/16 (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptom</td>
<td>1939</td>
<td>1975 (79)</td>
<td>1,846</td>
</tr>
<tr>
<td>Quality of life</td>
<td>1968</td>
<td>1979 (79)</td>
<td>4,603</td>
</tr>
<tr>
<td>Health-related quality of life</td>
<td>1989</td>
<td>1999 (90)</td>
<td>681</td>
</tr>
<tr>
<td>Patient-reported outcome</td>
<td>2003</td>
<td>2013 (81)</td>
<td>182</td>
</tr>
<tr>
<td>Patient-centered outcome</td>
<td>2004</td>
<td>NA (25 total)</td>
<td>9</td>
</tr>
</tbody>
</table>

Note. as of January 22, 2017.

N = number of citations

NA = not available

** Based on qualitative visual examination of upward trajectory maintained over time