# *STRengthening Analytical Thinking for Observational Studies (STRATOS):* RECOMMENDATIONS FOR THE DESIGN AND ANALYSIS OF PATIENT REPORTED OUTCOMES (THE SISAQOL-IMI PROJECT).

Saskia le Cessie 1,2,3 , Satrajit Roychoudhury 4 , Jammbe Z. Musoro 5 , Madeline Pe 5 Ahu Alanya 5 , Cecilie

Delphin Amdal 6,7 , Dries Reynders 3 , Doranne Thomassen 1 , Willi Sauerbrei 8 , Els Goetghebeur 3

1. Department of Biomedical Data Sciences, Leiden University Medical Center, the Netherlands.
2. Department of Clinical Epidemiology, Leiden University Medical Center, the Netherlands.
3. Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Belgium.
4. Pfizer Inc, US.
5. European Organisation for Research and Treatment of Cancer (EORTC) Headquarters, Brussels, Belgium.
6. Research Support Services, Oslo University Hospital, Norway.
7. Department of Oncology, Oslo University Hospital, Norway.
8. Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center - University of Freiburg, Germany.

In a previous article in Biometric Bulletin (2024, 41(3)), we described the involvement of the STRATOS initiative in the SISAQOL-IMI Consortium. The SISAQOL-IMI Consortium, which started in 2021, aims to develop standards for designing, analysing, presenting and interpreting health-related quality of life (HRQOL) and patient-reported outcome (PRO) data in cancer clinical trials (1) . Recently, the SISAQOL-IMI consortium delivered its final recommendations (2) . In this article we describe these results with an emphasis of the contribution of STRATOS members. SISAQOL-IMI is an international, multidisciplinary consortium, involving regulatory agencies, health technology assessment bodies, the pharmaceutical industry, and academic and professional societies alongside experts in statistics, PRO measurement, clinical oncology, and patient advocacy.

The project is led by the European Organization for Research and Treatment of Cancer (EORTC) and the pharmaceutical company Boehringer Ingelheim AG. SISAQOL-IMI work was organized into several work packages, including methodology for cancer randomized controlled trials (RCTs), feasibility of recommendations for single-arm trials (SAT), guidance on clinically meaningful change in cancer trials, and communication tools for PRO findings in cancer trials. STRATOS contributed primarily in the work package 3 (WP3) on SAT. The WP3 was led by Saskia le Cessie (STRATOS member) and Satrajit Roychoudhury (Pfizer Inc.), Els Goetghebeur and Willi Sauerbrei actively participated in this WP. Four junior researchers attached to STRATOS TG7 members played an important role in the development of the work and became (joint) first authors on at least one of the papers.  Doranne Thomasson recently became a new STRATOS member herself.

SATs play an increasingly important role in cancer research. However, SATs face specific challenges

due to the absence of a randomized control group, which may limit the interpretation and conclusions. SISAQOL-IMI WP3 investigated the feasibility of using PROs in SATs and developed recommendations for the use of PROs in non-randomized cancer studies with a specific focus on SATs. The work resulted in four publications.

The first paper contained the results of an extensive literature review focusing on design and reporting of PRO in SATs (3). It showed that PRO objectives in SATs were often unclear or not mentioned at all. Moreover, the methods to address missing data or intercurrent events were not properly specified or not justified. Finally, PROs were rarely collected after patients stopped treatment. All together this may generate vague results and unclear conclusions. Based on these

findings we recommend using the ICH-E9 (R1) estimand framework for SAT as well (4). This framework allows PRO objectives to be translated into key research questions of interest, including pre-specified strategies to handle intercurrent events, and providing an appropriate summary measure for precise description of treatment effect.

By means of a case study, the second paper outlines the meaning and impact of the estimand framework in the analysis of longitudinal PRO data in a SAT (5) . It demonstrated that different strategies for handling intercurrent events can lead to substantially different results and conclusions, even in a descriptive setting. In particular, addressing death should be carefully considered in advance, because PROs after death are not defined. The chosen strategy should be defined prior to analysis in line with the pre-defined PRO objective.

The third paper focused on imputation of the missing data in PRO measurements in SATs (6). Missing PRO data are challenging because missingness may be related to a patients' disease status and the corresponding patient experience. Information on intercurrent events (e.g., death, disease progression) may therefore support the reconstruction of unobserved PRO values. The paper developed imputation models for repeated PRO measurements that incorporated information from intercurrent events. A key message of this paper was that the missing-data model should be separated from the analysis model, when missingness is related to intercurrent events. Sensitivity analyses should be conducted to assess the impact of assumptions made about missing data mechanisms.

The fourth paper examined the use of external control data in SAT (7). Use of appropriate external data may address some of the concerns of SATs where a control group is lacking. However, choosing a relevant estimand for the comparison, and appropriate accounting for confounding and differing study drop-out remain challenging. The paper focused on settings with substantial mortality and emphasized the importance of a two-dimensional estimand, consisting of survival over time and mean PROs while alive. Two different methods for estimation under 'no unmeasured (time-varying) confounding' have been considered: a regression-standardization approach and a double re-weighting approach to account for differences in baseline-case-mix and censoring over time. Both performed well in simulations and on a case study comparing the treated SAT with the control arm of an existing RCT. A sensitivity analysis comparing the SAT with the treated arm of the RCT was consistent with the assumptions made. Key messages were that 1) the two-dimensional causal estimand is meaningful in the presence of a terminal event and 2) adapted sensitivity analyses are warranted when estimating treatment effects from SATs with an external control. After several rounds of formulating recommendation statements, followed by annual consensus discussions and revisions, WP3 formulated a final set of recommendations for SATs. This was done in close collaboration with the other work packages, particularly the work package on RCTs, because many recommendations apply to both SATs and RCTs. Accordingly, recommendations for RCTs and SATs were aligned and harmonized wherever appropriate. In total there are 43 WP3 recommendations and space allotted to this article does not allow us to discuss all recommendations in detail. However, we highlight the following key points:

1. SATs should have pre-specified PRO objectives that should be translated into key clinical questions using the estimand framework.

2. PRO objectives in SATs 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 external control. Appropriate steps should be taken in the design and conduct to reduce bias and avoid misleading interpretations. The absence of randomization and blinding should be addressed and alternative identifying assumptions discussed.

3. 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.

4. 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 method relies and whether the result is still in line with the intended estimand should be examined.

Alongside with the pivotal SISAQOL-IMI publication (2), which explains how the recommendations were developed and presents the key scientific results, the SISAQOL-IMI website offers practical tools to help users easily find and apply the recommendations that best suit their needs (https://www.sisaqol-imi.org/). For example, the online interactive table allows stakeholders to identify relevant recommendations based on the type of trial (RCT vs. SAT), the objective of the trial (confirmatory vs. descriptive), and the type of PRO endpoint (e.g., responder analysis vs. time-to-event analysis). The online interactive glossary provides scientific and plain language definitions of key terminology, to facilitate understanding across diverse audiences. Finally, an interactive guidebook provides guidance on how to navigate both online tools, while also providing additional information on the goals and methodology of SISAQOL-IMI. In summary, this collaboration across many different disciplines has produced a set of clear, well-considered recommendation statements with broad acceptance for the use of PROs in cancer studies. We hope that these recommendations will contribute to improved design, analysis and reporting of studies involving patient-reported outcomes.

1. 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-e83.

2. Amdal CD, Falk RS, Alanya A, Schlichting M, Roychoudhury S, Bhatnagar V, et al. SISAQOL-IMI consensus-based guidelines to design, analyse, interpret, and present patient-reported outcomes in cancer clinical trials. Lancet Oncol. 2025;26(12):e683-e93.

3. 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-e206.

4. Agency EM. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. 2020 [Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf]

5. Thomassen, D., Roychoudhury, S., Amdal, C. D., Reynders, D., Musoro, J. Z., Sauerbrei, W., Goetghebeur E ,le Cessie, S. (2024). The role of the estimand framework in the analysis of patient-reported outcomes in single-arm trials: a case study in oncology. BMC Medical Research Methodology, 24(1), 290.

6. Thomassen, D., Roychoudhury, S., Amdal, C. D., Reynders, D., Musoro, J. Z., Sauerbrei, W., Goetghebeur E, Le Cessie, S. (2025). Handling missing values in patient-reported outcome data in the presence of intercurrent events. BMC Medical Research Methodology, 25(1), 56.

7. Reynders D, Thomassen D, Roychoudhury S, Amdal CD, Musoro J Z, Sauerbrei W, le Cessie S, Goetghebeur E (2025). Evaluating treatment effects on longitudinal outcomes with attrition due to death: Methods for a two-dimentional estimand with a case study in Quality of Life. arXiv preprint arXiv:2509.25548 and submitted.