WP5 (D5.4): Independent validation (summary)

Introduction

One of the work package 5 (WP5) tasks was to evaluate the feasibility of implementing the SISAQOL-IMI recommendations developed by the scientific working groups (WP 2, 3, 4 and 6), with respect to writing the study protocol and the statistical analysis plan (SAP) for randomized controlled trials (RCTs) and single-arm studies. The task was achieved through semi-structured interviews (cognitive debriefing) and independent validation process by blind members. This summary presents the findings from the independent validation of the feasibility of implementing the SISAQOL-IMI recommendations in the development of study protocols and SAPs for RCTs and single-arm studies (Table 1). The summary contains list of trials used in the validation process, profiles of blind members, criteria used to evaluate the feasibility of implementation the recommendation in protocols and SAPs, perception of blind members about the difficulty or ease of implementing the recommendation and summary of results.

Methods

The feasibility of implementing the preliminary recommendations developed by the WPs 2, 3, 4, and 6 was assessed through an independent validation process. This was conducted by independent validation members, "blind members (n=12)" from academia, industry, and patient organizations, with varying expertise in the design, analysis and interpretation of clinical trial data (clinicians, researchers, statisticians) and were not involved in the recommendations' initial development (Table 2).

For a set of finalized clinical trials incorporating patient-reported outcomes (PROs), the blind members were asked to utilize the recommendations when completing PRO-related sections of study protocols and SAPs. Additionally, they were asked to assess the usefulness of the recommendations used. The PRO sections were assessed for the correctness of implementation (by developers of the recommendations). Key outcomes of the validation included the ability of the blind members to implement the recommendation statements sufficiently, and their perception of the recommendations' utility when completing the protocols and SAPs.

Two types of documents were prepared for the independent validation:

- the protocol and SAP with blanked out parts for "blind members" to complete during the validation exercise, and
- the gold standard reference documents which include correct responses for the blanked-out parts. The correct responses were provided by the WP leaders based on the SISAQOL-IMI recommendations that they developed.

All blind members were asked to complete the protocol, but only statisticians were tasked to do both protocol and SAP.

Table 1: Trials included in the validation

Trial	Tumor type	Disease stage / incidence	Treatment	PRO objective / endpoint used
EORTC 62091/RCT	Soft tissue sarcoma	Advanced / rare	Chemo vs. other chemo	Time to event
EORTC 22033/RCT	Low grade glioma	Early / rare	RT vs. chemo	Magnitude of change
BI LUC-Lung 3/RCT	NSCLC	Advanced / common	TT vs. chemo	Responder analysis

Pfizer PROFILE 1005/single arm	NSCLC	Advanced / common	TT	Single arm descriptive
Pfizer PROFILE 1007/RCT	NSCLC	Advanced / common	TT vs. SOC	Time to event

Table 2: Blind members demographics

Blind member	Stakeholder group	Profession	PRO research experience
#1	Academic	PRO researcher	3 years
#2	Academic	Biostatistician	14 years
#3	Industry	Statistician	missing
#4	Industry	Statistician	5 years
#5	Industry	Statistician	13 years
#6	Academic	Clinical & PRO researcher	5 years
#7	Non-profit/cancer organization	Statistician	0.5 years
#8	Industry	Statistician	5 years
#9	Dropped out		
#10	Academic	Research manager	8.5 years
#11	Academic	Clinician	3 years
#12	Academic	Statistician	14 years
#13	Academic	Statistician	10 years

Table 3: sufficient or insufficient implementation of the recommendation statements by blind members

Reference protocol	**Info to be completed**	**Gold standard answer**	**Blind member answer**	**Conclusion**
xxx	Describe the **population level summary** for the confirmatory PRO objective, including the time frame	mean magnitude of change in role functioning from baseline until 36 months of follow-up	the mean magnitude of change in role functioning from before randomization until 36 months of follow-up	Sufficiently implemented **(≥ 67%)**
yyy	define the **time windows**, (e.g., one week before and one week after the scheduled assessment and describe whether this is similar for the two treatment arms)	defined as one week before and one week after the scheduled assessment and are similar for the two treatment arms	three months from treatment for both arms, until 36 months	Not sufficiently implemented **(< 67%)**

Table 4: blind member's perceptions of usefulness of the recommendation statements

Example recommendation	Blind member answer	Conclusion
WP6 Survey 2, Rec. #2: Anchor-based clinically meaningful change (CMC) thresholds should be based on anchors that are meaningful to patients, and if other anchors are used, this should be justified explicitly.	Not at All useful (0/11=0.0%) A little useful (0/11=0.0%) Quite a bit useful (5/11=45.5%) Very much useful (6/11=55.5%)	Useful (≥67%)
WP6, survey 2, Rec. #3: Use of clinically meaningful change (CMC) thresholds should account for possible differences between improvement and deterioration of PROM scores. If no such distinction is made, the reason should be stated.	Not at All useful (1/11=9.1%) A little useful (3/11=27.3%) Quite a bit useful (4/11=36.4%) Very much useful (3/11=27.3%)	Not useful (<67%)

Results

Results showed that many recommendation statements were successfully implemented and deemed useful. However, specific areas, such as intercurrent events, time windows, visualization guidance, and the interpretation of clinically meaningful changes (CMCs), required further refinement. To address these gaps, additional examples, glossary updates, and templates were added to the recommendations, and supplementary training materials are under development.

WP2: Randomized Controlled Trials (RCTs)

The majority of the recommendation statements from WP2 were sufficiently implemented and were reported to be useful in RCT protocols and SAPs. However, topics relating to intercurrent events and strategies, time windows, completion and available data rates, and confirmatory analysis (methods, variables, time in the model) were not sufficiently implemented. These gaps were addressed by providing more clarity through the examples and explanations, and enhancements to the glossary.

WP3: Single Arm trials

The majority of the recommendation statements from WP3 were sufficiently implemented and were reported to be useful in single arm protocols and SAPs. However, recommendations on missing data strategy, population-level summary, and characteristics of external control groups were less fully implemented. These gaps were addressed by adding clearer examples and explanations, revising the statement on missing data, and updating the glossary.

WP4: Communication Tools (Visualisations)

The majority of the recommendation statements from WP4 were not sufficiently implemented in RCT SAPs. Feedback indicated that the recommendations lacked guidance on which data to visualize and how to select appropriate graphs for different contexts. To address this, WP4 leaders provided templates that clearly specify which data to visualize and recommend appropriate graph types for each context.

WP6: Clinically Meaningful Change (CMC)

The majority of the recommendation statements from WP6 were sufficiently implemented and were reported to be useful in RCT protocols and SAPs. However, topics related to the definition of CMCs (group level vs. patient level), the development of CMCs (such as characteristics and the distinction between improvement and deterioration), and the interpretation of CMCs were noted as challenging to implement. These issues were addressed by enhancing clarity through additional examples and explanations, adding an introductory paragraph to clarify development methods, and drafting a publication to provide further guidance.

Conclusion

This validation underscores the general feasibility of the SISAQOL-IMI recommendations while highlighting opportunities for enhancement in areas requiring improved clarity and guidance. The insights from this report will inform ongoing refinements to ensure the practical application of these recommendations in clinical trial protocols and SAPs.