



The professional society for health
economics and outcomes research

Improving healthcare decisions

505 LAWRENCE SQUARE BLVD SOUTH
LAWRENCEVILLE, NJ 08648

P +1-609-586-4981
F +1-609-586-4982

info@ispor.org

**2023–2024
Board of Directors**

July 5, 2023

President

Brian O'Rourke, PharmD
Brian O'Rourke
Healthcare Consulting
Ottawa, ON, Canada

Docket Number FDA-2023-D-0026

Dear FDA:

President-Elect

Eberechukwu
Onukwughu, PhD
University of Maryland
Baltimore, MD, USA

ISPOR – the professional society for health economics and outcomes research - is pleased to respond on behalf of its membership to your consultation entitled “Patient-Focused Drug Development: Incorporating Clinical Outcome Assessments Into Endpoints for Regulatory Decision-Making.”

Past President

Jan Elias Hansen, PhD
Genentech
San Francisco, CA, USA

ISPOR is a scientific and educational society with many of its members engaged in evaluating health technologies, including pharmaceuticals, medical devices, and other interventions. We have a large membership living and working in 110 countries globally, across a range of disciplines, including health economics, epidemiology, public health, pharmaceutical administration, psychology, statistics, medicine, and more, from a variety of stakeholder perspectives, such as the life sciences industry, academia, research organizations, payers, patient groups, government, and health technology assessment bodies. The research and educational offerings presented at our conferences and in our journals are relevant to many of the issues and questions raised in this request for information.

Directors

Dalia Dawoud, PhD
Associate Director (research),
The National Institute for Health
and Care Excellence (NICE)
London, England, UK

Beth Devine, PhD
University of Washington
Seattle, WA, USA

Marlene Gyldmark, MPhil
Roche Diabetes Care
Basel, Switzerland

Andrea Manca, MSc, PhD
University of York
York, England, UK

The response to this consultation was led by the Policy Outlook Committee of our most senior advisory body, the Health Science Policy Council. We solicited our general membership, as well as our Clinical Outcomes Assessment (COA) and Patient Centered special interest groups for comments. The attached document provides both summary and line-by-line responses based on their comments. We hope they prove useful.

Amy K. O'Sullivan, PhD
Sr. Vice President and
Chief Scientific Officer
Ontada
Boston, MA, USA

ISPOR would be happy to answer any questions about our response, and to participate in any follow-up consultations on the relevant program items mentioned within the report.

Dong-Churl Suh, PhD
Chung-Ang University
Seoul, Korea

Sincerely,

Lucinda Orsini, DPM, MPH
Vice President, Value and
Outcome Research
Compass Pathways
Skillman, NJ, USA

Robert Abbott
CEO & Executive Director
ISPOR

Treasurer (2023-2026)

Sean D. Sullivan,
BScPharm, MSc, PhD
University of Washington
Seattle, WA, USA

CEO & Executive Director

Rob Abbott
ISPOR
Lawrenceville, NJ, USA

Patient-Focused Drug Development: Incorporating Clinical Outcome Assessments Into Endpoints for Regulatory Decision-Making.”

Section I

The ISPOR community congratulates the Food and Drug Administration (FDA) on the release of the fourth Patient-Focused Drug Development (PFDD) guidance document draft. As researchers and methodologists in the health economics and outcomes research (HEOR) fields, including clinical outcome assessment (COA), we are eager to engage with FDA on this draft guidance to help shape the methods used and needed in COA endpoint development and use in regulatory decision making.

The COA field has advanced considerably in the twenty-plus years since the 2009 FDA Patient-Reported Outcome (PRO) guidance was being drafted, including much more explicit emphasis on the patient voice. In alignment with the previous guidances in the PFDD series, Guidance 4 places significant emphasis on patient engagement and experience. We applaud the FDA for its efforts to advance PFDD and champion patient centricity in medical product development to identify endpoints that represent what is meaningful to patients.

ISPOR members request that FDA maintain flexibility as the COA field integrates into daily practice the new terminology and/or approaches recommended in the guidance and that a transition period be expected for that integration and to accommodate the great number of trials containing COAs that were underway prior to draft guidance release. Some of the changes proposed would be facilitated through dialogue between FDA and the research community; ISPOR offers the FDA a forum for such interactions.

The ISPOR community of researchers appreciates that - as highlighted in both Guidances 3 and 4 - sound rationale based on evidence should be the foundation for assessing all aspects of COA work, including incorporating COAs into endpoints. We ask FDA to continue to use that bar as it assesses COA endpoints implemented prior to draft guidance release.

We applaud the FDA for recommending that sponsors consider innovative approaches, such as construction of personalized endpoints. This openness by FDA is welcome since the agency has resisted the approach of personalized endpoints in the past. However, taking this on would be a pioneering effort by a sponsor, one that offers many opportunities coupled with many risks. Thus, we ask FDA for additional guidance on such innovative approaches and examples of how these approaches have been and/or can be used, as well as opportunities for interacting and/or collaborating with FDA to bring novel approaches into the mainstream (eg, pilots or collaborative research agreements). Again, flexibility by FDA will be needed as the field moves toward novel methods where there is little implementation experience.

We appreciate the opportunity to provide below more detailed questions, comments, and recommendations on Guidance 4, offered through the lens of our members, many of whom will be users of the final guidance documents.

Section II

Overall, this guidance is timely and will enable Sponsors to develop COA strategies that meet the needs of selecting and reporting on concepts important to patients and the agency. While we acknowledge the historical development of endpoints with COAs in mind, we note that all too often endpoints are described as a change in COA versus a change in concept as measured by the COA. It is important to shift the thinking toward highlighting the concept/domain (derived from the patient) that is expected to change with treatment. Given that the progression from establishing the relevant concept or domain of interest, selecting, modifying or developing the appropriate COA, and then defining an endpoint is well characterized in PFDD Guidance 3, we hope that the agency will refer back to this guidance to reinforce this key approach to deriving appropriate COA endpoints.

Generally, in the discussion of Section II, we believe that greater clarity is required by use of consistent

terminology (eg, assessments and measurements), defining terms, and the inclusion of specific examples to ensure that sponsors are provided clear guidance on COA endpoint selection. Specifically, it would be useful for the agency to define “endpoint”, the level of evidence that would be considered acceptable to the agency regarding endpoints, and a reference to relevant resources which would be valuable for Sponsors. Indeed, some specific examples throughout this section would be of value to Sponsors to ensure greater understanding of the requirements of the agency. We would further encourage FDA to explore and highlight the domains or concept of interest within endpoints as opposed to only in the COA. We would recommend revisions in this section to the effect of “the core of any COA endpoint should be the concept of interest (what is being measured from the patient perspective).” Additionally, from a reporting perspective, some guidance on where the evidence for the endpoint should be presented (for example, in the clinical trial protocol, briefing book, both, or other) would be encouraged. From an interpretation perspective, we also believe that sponsors would benefit from further guidance and examples of how endpoints are evaluated with respect to the estimand framework¹, and more specifically the handling of intercurrent events in describing the endpoint.

We note that within Section II, in the discussion of ‘Endpoints constructed by dichotomizing COA scores’ the guidance appears contradictory at times, noting that sponsors should pre-specify only a single score threshold and provide evidence for this threshold while also suggesting additional analyses to explore treatment effects over a range of thresholds. We recommend FDA clarify expectations regarding score thresholds. For example, if a study team has evidence for a single score threshold, there should not be a need for exploratory analyses looking at different thresholds. If, however, a study team is not sure of the most appropriate score threshold then designing endpoints to analyze different score thresholds would be more appropriate.

The guidance highlights the preferred method for adjusting for baseline in the context of a statistical model but does not explain why rationale for this consideration should be included. The guidance implies that analyzing the PRO endpoint as a change from baseline should rarely be done, but it is not clear why that is the case. Furthermore, this seems to contradict parts of the guidance where a reference to analyzing change from baseline endpoints is made. We recommend either including a statement that endpoints should not be analyzed as change from baseline (with rationale) or qualify the existing text with a clarification that using a change from baseline endpoint is another acceptable approach. For example, consider discussing that change from baseline strategies that require subjects to have at least one baseline and follow up assessment can lead to selection bias, with the randomized groups no longer assumed to belong to the same population of subjects.

In line with current analytical methods, we recommend FDA include a discussion on the possible approaches to analysis of ordinal data including via ordinal logistic regression by treatment at the end of the study period. We would appreciate some comment on whether it is reasonable to describe changes on ordinal scales in the context of proportion of patients moving across categories along the scale (ie, shift tables).

Section II provides some information on missing data, suggesting a potential solution in site staff asking patients to complete the missing assessment. We note that this approach may not be feasible due to research staff workloads, recall of the assessment (as PRO data cannot be captured outside of the specified window) and the potential need for a patient to come back to the trial site to perform the assessment either on a site-based tablet electronic clinical outcomes assessment (eCOA) collector or a performance-based outcome (PerfO). We strongly encourage FDA to consider feasibility and patient and research staff burdens when considering approaches to missing data. We recommend FDA emphasize the importance of training patients and staff on proper data collection and highlight the value of greater engagement with patients (eg, patient engagement in protocol development, COA development, and mock clinical trial participation to test

¹ European Medicines Agency. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials 2020. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf. Accessed July 5, 2023.

burden prior to study start) to learn in advance what PROs/circumstances may be burdensome.

With respect to the statistical aspects of missing data, further guidance, and examples of imputation strategies for percent change-from-baseline when the COA baseline value is zero would be appreciated. Additionally, we recommend adding the example of data which is not normally distributed; we expect that if an outcome is not normally distributed, it would be recommended to transform the data in some fashion (ex. Logarithmic) then summarize as percent change. The guidance here also implies that only sensitivity analyses can be used to address missing data. We recommend FDA reword this section so that analyses for missing data can be considered primary analyses rather than limited to only sensitivity analyses.

Section III

The guidance implies that Section III will provide approaches to help with the interpretability of endpoints without the need for dichotomization. However, it is not clear that Section III achieves this. We encourage FDA to add further detail and guidance regarding the dichotomization of continuous/ordinal data, and the associated interpretability issues which may arise, and refer to that guidance within the Guidance 4 document. We recommend FDA provide examples regarding approaches to the use of in-trial data to support score threshold definitions. In many situations (rare disease, oncology, and novel PROs) it will not be possible to have prior data to inform or confirm thresholds. This may be an interim cut of data or use of a time point in the trial other than the primary endpoint.

The guidance recommends the use of personalized endpoints if there is heterogeneity in the manifestations of a condition. We welcome FDA's discussion of personalized endpoints. We note however there may be concerns regarding whether this is a practical way to assess the impact of a product on patient outcomes. We welcome additional guidance from the Agency on how a study may be powered to show a treatment effect if a personalized endpoint is used. If a clinical study cannot power a study to show treatment differences, we question what circumstances this approach may be useful in.

The discussion in this section regarding meaningfulness is important and helpful. The determination of meaningful treatment benefit is a critical topic; thus, it is very positive to have a detailed section on this in the guidance. Greater clarity would make this discussion even more helpful. In particular, some methods in the section regarding Meaningful Score Difference (MSD) and Meaningful Score Regions (MSR) are relatively novel and not all are in widespread use. Providing more detailed information would thus be useful to readers. In particular, we suggest clarification regarding the distinction between within- and between-group differences for meaningful score differences. Is FDA suggesting the same methods to assess each one? In addition, greater explanation of the more technical, field-specific terms would increase the accessibility of this guidance. Inclusion of a glossary that defines these terms and connects them to more broadly used concepts and terms could be helpful. Further clarification regarding the application of MSD and MSR would be beneficial.

Section IV

These specific points around issues arising in COA are very helpful. We do have a few specific comments/questions by section:

A.1. Masking. The presumption that lack of masking is likely to bias COA data seems overly general and not supported by research. While double-blind randomized designs are preferred, many situations exist in which they are unfeasible or unethical (eg, indications in which no standard of care or ethical alternative treatment is available) and we should not dismiss COA data as biased in these situations. Examples of publications

supporting this position include Atkinson et al 2016², Roydhouse et al 2019³, and Lord-Bessen et al. 2023⁴. FDA Oncology Center of Excellence (OCE's) 2022 workshop on the topic also highlighted the relevance of patient-reported outcomes in oncology, where many trials are not double-blind⁵.

An alternative may simply be to soften the bias declaration language to something like the following (replacing "will" with "may"): "Patients', clinicians', and/or caregivers' knowledge of treatment assignment (eg, in single arm trials, open label trials, open-label treatment extension periods) may influence how they report information on a PRO, clinician-reported outcome (ClinRO), or observer-reported outcome (ObsRO) measure, or how they engage with PerfO tasks (eg, amount of encouragement given to patients when measuring walking distance), may bias estimates of treatment effect."

A.2. Practice Effects. A practice effect does not have to be a "change," a practice effect is a bias. Practice and learning could prevent a score from worsening, therefore the bias being no change. For example, though Parkinson's symptoms are progressing the progression is not reflected in the PerfO because the participant has learned how to complete the task. We suggest revising to state: "A practice effect (sometimes also called a learning effect) is a bias that results from..."

In the discussion about timing of assessments and practice effects, one consideration is certainly how rapidly the change in the underlying construct is likely to occur, but for a given overall trial length, increasing timing between assessment implies decreasing their frequency, which introduces statistical considerations around the precision of the estimated effect. It seems that a balance of study concepts, feasibility, and post-design analytics (eg, practice effect bias, construct validity, precision, et al) - should go into the timing decision.

A.3. Use of Assistive Devices. Does this type of concern over outcome-relevant post-randomization influences apply only to assistive devices, or could it apply similarly to rescue medications. They can be another signal that the disease is not controlled but also affect the measured outcome. Do the same suggestions relative to the use of COA's apply? We recommend FDA consider broadening this discussion to an estimand framework, in particular how to include intervening events into the endpoint and analysis.

A.4. Considerations When Using a Nonrandomized Designs, External Controls, or Nonconcurrent Control. The use of external controls and "preexisting differences between the groups" issue discussed in this subsection are such major methodological issues involving selection/confounding and related biases that it deserves more attention. At the very least, it seems relevant to reference other recent FDA draft guidances, namely "Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products" and "Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products."

A.5. Analysis of Treatment Effects for Subgroups Based on Post-Baseline Scores. The above comment about employing the estimand framework also is relevant here.

A.6. Computerized Adaptive Testing (CATs). Please note that CATs may also be based on Rasch modeling. In addition, suggestions about the best timing for discussion and alignment with the appropriate agency (eg, ideally at phase II and at the latest 30- 60 days prior to closure of the study and unblinding) would be most welcome.

² Atkinson TM, Wagner JS, Basch E. Trustworthiness of Patient-Reported Outcomes in Unblinded Cancer Clinical Trials. *JAMA Oncol.* 2017 Jun 1;3(6):738-739. doi: 10.1001/jamaoncol.2016.3328. PMID: 27658006; PMCID: PMC5362349.

³ Roydhouse JK, Mishra-Kalyani PS, Bhatnagar V, Gutman R, Sridhara R, King-Kallimanis BL, Kluetz PG. Does knowledge of treatment assignment affect patient reported of symptoms, function, and quality of life in cancer trials? An evaluation in multiple myeloma. *Value in Health.* 2021;24(6):822-829.

⁴ Lord-Bessen J, Signorovitch J, Yang M, Georgieva M, Roydhouse J. Assessing the impact of open-label designs in patient-reported outcomes: investigation in oncology clinical trials. *JNCI Cancer Spectr.* 2023 Mar 1;7(2):pkad002. doi: 10.1093/jncics/pkad002. PMID: 36661326; PMCID: PMC10023242.

⁵ FDA Workshop:7th Annual Clinical Outcome Assessment in Cancer Clinical Trials Workshop - 06/29/2022

In addition, the guidance does not discuss relevant situations in which the Sponsor wants to demonstrate no difference between groups. This might be in the context of adding a novel drug to an established drug regimen, where there is a need to demonstrate no side effects that impact patient health. We encourage FDA to add considerations for design, analysis and interpretation in such scenarios. We note that this is a common situation in oncology and other complex chronic conditions.

B. Formatting and Submission Considerations.

Lines 1472-1509: The guidance refers readers to a variety of FDA guidances on formatting and considerations. However, these are high-level guidances and not specific to COA. Please describe the timing and format of submission of a COA within an NDA and through a PRO Dossier.

Lines 1517-1527: The guidance refers readers to several FDA guidances on electronic devices for completion of COAs. Some of these guidances focus on drug-device combination submissions and may not be relevant to eCOA collection devices. Please clarify the relevance of these guidances to eCOA collectors especially with respect to threshold analyses and human factors submissions related to the collectors.

We would like to acknowledge ISPOR members Bryan Bennett, Eleanor Perfetto, and Jessica Roydhouse for their assistance in assembling these comments, as well as ISPOR staff Richard Willke, Clarissa Cooblall, Kelly Lenahan, and Rachel Peoples.